



City Research Online

City St George's, University of London

Citation: Millan Cifuentes, J. D. (2017). Curiosity driven search experiences. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/18085/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Curiosity-driven Search Experiences

Juan David Millan-Cifuentes

Thesis
City University London
2015

Para mi amado padre

Acknowledgement

A PhD is not 100 meters race, but it's more like a marathon where endurance and perseverance are more important than speed. Sometimes a PhD feels like lonely journey. However, I have been extremely fortunate to have been surrounded by people who have given guidance and support throughout my time as a research student. Without these people, this thesis may not be in existence and it certainly would not exist in its current form.

I would like to thank my supervisors Dr. Andrew MacFarlane and Profs. Ayse Göker who have been simply wonderful. They have offered me inspiration, support, patience, motivation, advice, freedom and encouragement. For this I will always be grateful. I would also like to thank Hans Myrhaug because of his useful comments and support during my employment at Ambiansense Ltd.

The members of the computer and information science department of City University have been a great help too for my research. In particular Dr. Stephann Makri, and Dr. David Haynes who gave me their time and feedback while I was transferring from MPhil to PhD.

To my friends and colleagues throughout the years, thank you for your patience and attentive ear. Jaime Chavarriaga, thank you for sending me a link that changed my life and for helping while I was writing my proposal to start this long journey. To Alvaro Pachon and Monica Lora who have always supported my professional career.

To my family, who have been extremely positive and supportive throughout my PhD. My brother, my true friend and mentor. My father who taught me the value of education and the joy of being in a classroom with a black marker. My mum who taught me what is most important in life. To all of you, thank you!!

Finally, to Nadia, my wife and mi hermosa mona, who has been my unending source of strength and joy. Thank you for everything. For all of your love and support. I wouldn't have been able to do this without you. Specially, when I was giving up, you believed in me, this kept me going.

Abstract

Casual-Leisure Search describes any behaviour that allows people to express and satisfy hedonistic needs rather than information needs as part of the information-seeking process. For example, individuals who search their social media universe for hours after a long day at work may do so out of curiosity, to relax or for fun (e.g. exploring for the experience). Studies have shown that classical information seeking (IS) and interactive information retrieval models (IIR) have failed to represent them because they were created observing people in work related scenarios, and assuming that search is always a rational decision making process and with an extrinsic utilitarian value. The research described in this PhD work investigates IIR from the perspective of the psychological curiosity and leisure information seeking behaviour. Traditional search engines focus the user experience on satisfying users with topically relevant information (i.e. quick lookup search and then moving on), but they are limited supporting the discovery of unknown information because they fail to entice and engage users exploration as proxy to seek enjoyment both in leisure and work scenarios. The research described increases understanding of the role that curiosity plays in IIR and investigates the merits of incorporating the characteristics and function of human curiosity in the design of IIR systems. The research is grounded by the theoretical understanding of how human curiosity works. A review of appropriate psychological curiosity literature offers a means to critique existing IIR tools and a basis from which to start designing novel curiosity driven search tools. In the first experimental work, this research compared IIR behaviour between a standard query-response paradigm and a curiosity driven search map prototype using social media content, and attempts to learn lessons from the behaviour that people show in everyday casual-leisure search scenarios. In the second experiment, this research contrast IIR behaviour between standard query-response paradigm and a curious adaptation of query-response paradigm using search notifications or recommendations for news reading in a social media leisure search scenario. The tools are evaluated to determine the usefulness of incorporating curiosity in the design of IIR systems, to learn about the effect in user engagement, how users exploration is increase when motivated by a hedonistic

need, and then elaborate a set of design recommendations to enhance the search experience in leisure scenarios.

Contents

Abstract	6
1 Introduction	19
1.1 Problem Statement	21
1.2 Overall aims and goals	22
1.2.1 Objectives	23
1.2.2 Research Questions	23
1.3 General Research Design	24
1.3.1 Social Media Scenario	25
1.4 Scope and Limitations	27
1.5 Dissemination	28
1.6 Outline Thesis	28
2 Literature Review	31
2.1 Curiosity	32
2.1.1 Curiosity-drive theory	32
2.1.2 Optimal arousal theories	33
2.1.3 Interest/Deprivation Model of Curiosity	34
2.1.4 Wanting/Liking Model of Curiosity	35
2.1.5 Curiosity Arousal Principles	37
2.1.6 Other Types of Curiosity and Discussion	38
2.1.7 What is Curiosity for this research?	39
2.2 Models of Information Behaviour	40
2.2.1 Dervin's Sense-Making Model	40
2.2.2 Information Foraging	42
2.2.3 Wilson's Information-Seeking Behaviour Model	42
2.2.4 Ellis's behavioural model	43

2.2.5	Information Search Process Model	44
2.2.6	Work Task Information Behaviour	45
2.2.7	ELIS Model	46
2.2.8	Serious Leisure	47
2.2.9	Casual-Leisure Behaviour Model	48
2.2.10	Information Behaviour Literature Discussion	49
2.3	Information Retrieval Interaction Models (IIR)	51
2.3.1	Anomalous State of Knowledge (ASK)	53
2.3.2	IS&R Nested Model	54
2.3.3	Stratified Model of Information Retrieval Interaction	55
2.3.4	Marchionini's Information-Seeking Process Model	56
2.3.5	Berrypicking Model	59
2.3.6	IIR Literature Discussion	60
2.4	IIR and Evaluation	62
2.4.1	Relevance	62
2.4.2	User-centred approach	63
2.4.3	Classical Measures of IIR	64
2.4.4	Search Experience	64
2.4.5	User Engagement	65
2.4.6	Serendipity	65
2.5	Context and IR	66
2.6	IIR and Search Interfaces	68
2.6.1	Formulating the query	69
2.6.2	Examining Results	69
2.6.3	Controls and Facets	73
2.7	Casual-Leisure Search	75
2.7.1	Definition	75
2.7.2	Motivation	76
2.7.3	Limitations and Challenges of IIR for Casual-Leisure Search	77
3	Exploring the effects of curiosity driven design in the search experience	81
3.1	Social Media Search	81
3.2	Motivation	82
3.3	Curiosity driven search application (Ambiecities)	84
3.3.1	Curiosity in Information Systems and IIR	84

3.4	Designing for Curiosity	86
3.5	Methodology	92
3.5.1	Data Collection	92
3.5.2	Procedure	93
3.5.3	Search Scenario	95
3.5.4	UES for leisure	97
3.5.5	Participants	98
3.5.6	Data Preparation	98
3.6	Results	99
3.6.1	Reliability analysis of UES	99
3.6.2	Relationship between the UES sub-scales	102
3.6.3	Total scores of UES	103
3.6.4	Browsing and Searching Behaviour	105
3.6.5	Relevance judgments in a Casual-leisure Scenario	110
3.6.6	Task Perception	112
3.7	Discussion	113
3.8	Limitations	118
3.9	Implications	119
3.9.1	Designing for the utilitarian and hedonic aspects of information searching	119
3.9.2	Measuring social media search experiences in a casual-leisure experimen- tal setting	121
3.10	Conclusions	125
4	Curious query-response design for leisure news search	127
4.1	Reflection on experiment 1	128
4.1.1	Lessons Learned	128
4.1.2	Limitations	130
4.1.3	Motivation	132
4.2	Newslabels. A ‘curious’ news search application based on social media	134
4.2.1	Requirements	134
4.2.2	Software design, architecture and development of Newslabels	135
4.2.3	Mapping curiosity arousal principles into the query-response paradigm . .	142
4.2.4	‘Curious’ companion software design and architecture	148
4.3	Methodology	149
4.3.1	Variables and constructs	149

4.3.2	Search scenarios	155
4.3.3	Data Collection and Procedure	156
4.3.4	Participants User Study	158
4.3.5	Data Preparation	160
4.4	Results	162
4.4.1	Reliability analysis of UES	163
4.4.2	Search Scenario Perception	165
4.4.3	Relationship among different constructs and total time on the task	167
4.4.4	Personal Interest Variable Analysis	169
4.4.5	System Assignment Variable Analysis	170
4.4.6	Analysis Between Subjects for each Tasks	170
4.4.7	Search Notification Perception and UX	173
4.4.8	Total Dwelling Time and Guessed Total Dwelling Time	177
4.4.9	Curiosity as a trait	178
4.4.10	Analysis of Interaction Logs	180
4.5	Discussion	197
4.5.1	Does curiosity-driven search design increase users' engagement during casual-leisure search?	199
4.5.2	Does curiosity-driven IIR-system increase users exploration during casual-leisure search?	200
4.5.3	Why did Focused Attention (FA) sub-scale was not different between system assignment groups as Ambicities?	202
4.6	Limitations	203
4.7	Implications	204
4.7.1	On-line News Search on Leisure Scenarios	204
4.7.2	Affective IIR-systems	205
4.8	Conclusions	206
5	Conclusion	209
5.1	Analysis of Research Questions	209
5.1.1	Does curiosity-driven search design increase users' engagement during casual-leisure search?	210
5.1.2	Does curiosity-driven IIR-system increase users exploration during casual-leisure search?	212

5.1.3	How could casual-leisure search behaviour be studied in a formal experimen- mental setting?	213
5.1.4	What do the findings mean respect to the search experience design? . . .	214
5.2	Limitations	217
5.3	Contributions	218
5.4	Future Work	221
5.4.1	Understanding the Role of Curiosity in IIR	221
5.4.2	Affective IIR-systems. Designing, Implementing and Evaluating Curios- ity Driven Search Tools	222
5.4.3	IIR evaluation in leisure scenarios	223
5.5	Conclusion	224
A	Publications	237

List of Figures

1.1	Research Questions and related areas on knowledge	24
1.2	Overview Research Design	26
2.1	Litman [1] Wanting / Liking Curiosity Model	36
2.2	Dervin's Sense-Making Model	41
2.3	Wilson's Information-Seeking Behaviour Model [2]	43
2.4	Ellis's Information Seeking Behaviour Model illustrated by Wilson [2]	44
2.5	Kuhlthau's Information Search Process Model (ISP) [3]	45
2.6	Byström and Järvelin's Work Task Information Behaviour Model [4]	46
2.7	Original Casual-Leisure Behaviour Model	49
2.8	Standard information retrieval interaction in the information seeking process [5] .	52
2.9	Anomalous State of Knowledge	53
2.10	IS&R cognitive nested model	55
2.11	Stratified Model of Information Retrieval Interaction	56
2.12	Information-Seeking Process Model	57
2.13	Exploratory Search (Lookup, Learn and Investigate)	58
2.14	Bates' Dynamic Search Model	59
2.15	User Context Model [6]	67
2.16	Amazon Search Box	68
2.17	Amazon Search Suggestion for query formulation	68
2.18	Amazon's SERP	70
2.19	Amazon's misspelling auto-correction	70
2.20	Amazon's SERP	71
2.21	Scentbar suggestion query component. Screenshots are taken from [7]	72
2.22	Amazon's facets for cycling clothes	73
2.23	Timeline Slider Facet in ACM's Digital Library for the query "IR"	73

2.24 Pivotal Path Screenshot for the query “Ayse Goker” (see http://mariandoerk.de/pivotpaths/demo/)	74
2.25 Casual-Leisure Search Model	76
3.1 Ambiecities Mobile Map and List View	84
3.2 Ambiecities Software Architecture and Interaction Design.	87
3.3 GUI Screenshots. London, UK Map and real-time Twitter markers. Each marker contains a Tweet with links and photos	88
3.4 GUI Screenshots. London, UK Map after 3 seconds, and new Tweet appears. The red marker have changed location	88
3.5 GUI Screenshots. Regensburg, Germany Map just one Tweet after 2 minutes	89
3.6 GUI Screenshots. Japan Map after 4 seconds	89
3.7 Ambiecities vs. Twitter Experiment Setup	94
3.8 UES Total Scores vs. Total Time On The Task	108
3.9 FA sub-scale Scores vs. Total Time On The Task	108
3.10 NO sub-scale Scores vs. Total Time On The Task	109
3.11 Which of the following features made the information displayed by the App enjoyable for you?	111
3.12 Task Perception Summary (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”)	113
4.1 ‘Cold’ vs ‘curious’ query-response information interaction	136
4.2 Architecture Newslabels.com	137
4.3 Newslabels Prototype. Standard graphical layout. Results for the query “cat” using the filter “week ago”	139
4.4 Newslabels Prototype. Results for the query “barcelona” using the filter “24 hours ago” the 22th of April 2015	140
4.5 Newslabels Prototype. Timeline facet for the query “barcelona” using the filter “20 days ago” the 22th of April 2015	141
4.6 Sharing functionality	143
4.7 Flash notification message in Booking.com when searching a hotel for Istanbul, Turkey	144
4.8 Linkedin using partial exposure in the information interaction and interface design to evoked curiosity	147
4.9 Search Notifications	150

4.10 Search Notifications	151
4.11 Search Notifications	152
4.12 Protocol for the user study	159
4.13 Example of the recorded interaction logs	161
4.14 Normality ShapiroWilk Test all scales. The red color cells indicate that ShapiroWilk Test accepted the assumption of Normality	164
4.15 Task Perception between tasks (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”). Task A is brown and Task B is blue	166
4.16 Total Dwelling Time (TDT) for Task A and Task B	167
4.17 Scatterplot of total Dwelling Time (TDT) vs UES for both Task A and Task B. Grouped by system assignment: Cold is blue and ‘Curious’ is green	171
4.18 Boxplot of total dwelling time (TDT) vs personal interest (PI) across tasks. Grouped by system assignment: ‘Cold’ is green and ‘Curious’ is blue	174
4.19 Boxplot of error guessed total dwelling time (EGTDT) vs personal interest (PI) across both tasks. Grouped by system assignment: ‘Cold’ is green and ‘Curious’ is blue	178
4.20 Descriptive statistics for I-type and D-type scales of curiosity as trait [8, 9]	179
4.21 Descriptive statistics for each action group by topic	181
4.22 Mean number of actions group by type of action and system assignment	182
4.23 Total unique queries grouped by system assignment	183
4.24 Stream graph representing the relation between different actions on both systems group by date	187
4.25 Stream graph representing the relation between different actions in a sample of search sessions	187
4.26 A line graph representing the total amount of sessions for each IIR-system group by date	188
4.27 A line graph representing the total amount of time in a search session (Seconds) for each IIR-system group by dates	188
4.28 Matrix Scatter Plot between DTD (Minutes) and Click events	191
4.29 Matrix Scatter Plot between Query and Click events	191
4.30 Matrix Scatter Plot between Query and Click events	192
4.31 3D Scatter Plot between TDT(Minutes), Click and Query events	192
4.32 Word cloud based on the content of the Tweets saved during all search sessions. .	193

4.33	1st to 6th most common saved Tweets. Order from left to right, and top to bottom.	196
4.34	7th to 10th most common saved Tweets. Order from left to right, and top to bottom.	198

List of Tables

3.1	Curiosity driven scenarios evoked by novelty and uncertainty	91
3.2	User Engagement Scale Questionnaire for the casual-leisure scenario	97
3.3	Ambicities UES	100
3.4	Ambicities UES with 25 questions	100
3.5	Twitter UES	100
3.6	Twitter UES with 24 questions	101
3.7	UES Factor's Correlation	102
3.8	Ambicities UES total scores and standard deviation for each question	103
3.9	Twitter UES total scores and standard deviation for each question	104
3.10	Total Scores for all the 28 participants	105
3.11	Time Spent Searching (Hours: Minutes: Seconds)	105
3.12	Detail Participants Total Time on Task, Estimated Total Time on Task, and Absolute Error (Hours: Minutes: Seconds)	107
4.1	User Engagement Scale Questionnaire and Serendipity items for the casual- leisure news scenario	153
4.2	Questionnaire for PIN, DIV, UNEX and EXPO constructs	154
4.3	Litman et al. questionnaire for I-type Curiosity (IC) and D-type Curiosity (DC)	154
4.4	Variables and metrics summary based on previous research	155
4.5	Session 1 UES with 31-items	163
4.6	Session 2 UES with 31-items	163
4.7	Session 1 Other Scales	164
4.8	Session 2 Other Scales	164
4.9	General correlation between constructs and TDT in Task A	168
4.10	General correlation between constructs and TDT in Task B	169
4.11	Between subject analysis for system assignment in Task A	172

4.12	Between subject analysis for system assignment in Task A	172
4.13	Between subject analysis for system assignment in Task B	172
4.14	Between subject analysis for system assignment in Task B	172
4.15	Total sessions grouped by system type. The system with “Curious” search notification (enable) and the “Cold” search system (unable).	180
4.16	Total sessions grouped by topic (Financial, General, Sports).	180
4.17	Descriptive statistics for total actions group by system assignment, topic, and type of action	181
4.18	Descriptive statistics for each action group by system assignment	182
4.19	Top 50 most common queries grouped by system.	184
4.20	Descriptive statistics of total time in a search session group by system assignment.	185
4.21	Top 10 most common news sources in the saved Tweets	185
4.22	Top 10 longest sessions.	186
4.23	Top 10 longest sessions.	186
4.24	Correlations within the recorded actions and TDT	189
4.25	Top 10 most common click search notifications.	193
4.26	Click search notifications group by the screen location.	193
4.27	Duration of Search Notification in the screen (Milliseconds)	193
4.28	Most frequent 26 terms int the saved Tweets	194
4.29	Top 10 most common news sources in the saved Tweets	195
4.30	Top 10 most common saved Tweets	197
4.31	Descriptive statistic for the position of documents saved during all search sessions.	197

Chapter 1

Introduction

“Happiness is the absence of the striving for happiness” Chuang Tzu (Taoism, 389-286 BC)

Interactive Information Retrieval (IIR) acknowledges search as an interactive conversation between information systems and individuals trying to solve their information problem. In a traditional IIR scenario, everything starts when the user becomes aware of an information need, then depending on the individuals searching skills and particular context decides to iteratively interact with the information system until the needed information is found and used.

Most IIR-systems (e.g. search engines) design their interactive communication as a ‘cold’ query-response interaction problem. For example, finding something when users know what they want and have words to describe it, finding something users have seen before. Search is always a rational decision making process with an extrinsic utilitarian value. Thus, IIR-systems are often related to a magic search box and a rank list of results (SERP¹). Even though query-response interaction design has been apparently successful in work-based scenarios, many have highlighted the importance of moving beyond this paradigm when the work-task is exploratory or involves learning [10, 11].

IIR-systems are limited not just because of their ranking algorithms or technology stack, but by the interaction paradigm and the assumptions made about the searcher [12, 13]. Query-response interaction paradigm main focus is satisfying users with information (e.g. quick search, some topically relevant rank results and then moving on) instead of enticing, engaging and guiding users to discover and explore the unknown.

Also, this paradigm has directed the attention of IIR on the destination of the search (i.e. the results) with metrics such as precision and recall instead of the search experience itself [14] (e.g engagement, exploration, flow, etc). But IIR should do much better than accept situational judgments as unique possible source evaluation because ‘I liked what I got from

¹Search Engine Results Page

my search results’ could also imply ‘I am not worried about what I did not’[15]. Thus, the importance of designing the search experience to entice exploration and search engagement (i.e. making people curious or ‘worry’ about what they ignore).

However, casual-leisure information seeking studies have shown that users who have a strong hedonistic² motivation (e.g. curiosity, boredom, relaxing) could engage interacting with a query-response information system for more time than what they have thought and search without a particular extrinsic objective. During this leisure seeking scenarios, users could get pleasure from exploring and interacting with the IIR system. For example, an individual might search social media for hours because he is bored or out of curiosity rather than because he wants to fill an information need. Accordingly, Elswailer et al. [16] have suggested the need to reconsider information behaviour theory and the design of IIR-systems for casual-leisure.

Drawing from psychology, human curiosity is the most important driving force of cognitive development and exploratory behaviour which fundamental goal is to bring pleasure rather than simply acquire knowledge [1]. Curiosity is often related to a intrinsic motivation that leads to positive subjective experiences and personal growth because produces an openness to unexplored experiences, building a foundation for better opportunities to experience discovery and joy. For example, Steve Jobs co-founder of Apple and Pixar in a commencement speech at Stanford University explained how his curiosity was vital to the success of his career and entrepreneurship (e.g. going to a calligraphy class proved to be valuable later when creating typeface in the Mac). He said: “[...] much of what I stumbled into by following my curiosity and intuition turned out to be priceless later on. [...] Stay Hungry. Stay Foolish”[17].

Although curiosity is an internal initiation process and personality trait, there are common arousal mechanisms of curiosity (e.g. novelty, complexity, uncertainty, conflict) [18]. Curiosity could be experienced with a greater probability by designing products and services following these arousal mechanisms.

Therefore, this research aims to apply psychological curiosity theory [1, 18, 19, 20, 8, 21] to the design of IIR search experience, study the effects of curiosity driven design in the user’s search behaviour (e.g. exploration, engagement, relevance judgments) in social media casual-leisure search scenarios and how to design autotelic³ search experiences (e.g. users find irrelevant information but are happy with their search experience, then when they are fully absorbed and engaged they explore in depth and may discover unexpectedly useful information or just have fun).

²The word hedonistic is used in a context relating to pleasure and is not intended to be derogatory.

³Autotelic: having a purpose in and not apart from itself [22]. The word comes from to Greek roots: auto (self), and telos (goal) [23].

1.1 Problem Statement

Which is more important: the journey or the destination? Sometimes the journey proves to be as important as reaching a particular destination. For example, in 1492 Christopher Columbus who was trying to find a western sea route from Europe to Asia (i.e. the Orient or “the Indies”) when he arrived at the Bahamas in the New World (i.e. the Americas). Although he did not reach his final destination, this journey has changed the course of mankind’s history. Could this also be happening in IIR?

Classical IIR based on work-task scenarios usually puts the emphasis on the destination of the search (the results) with metrics such as precision and recall, rather than the search journey (the user experience). Although IIR evaluation partially includes characteristics of the search journey by capturing the users interactivity (e.g. number of queries, the mean size of the query, dwelling time, usability measures, etc) is biased towards the classical system based approached and pure utilitarian IIR model. The focus of IIR evaluation is to assess the efficiency or effectiveness of the IIR-system in a particular work scenario and not the user’s “emotional response to various stages of the interaction” [24]. Thus, it can be argue that in some sense classical IIR have missed the ‘true’ complete picture of the searching. Should IIR be interested in studying, understanding and supporting the search experience (i.e. the journey) beyond a rational decision making process with an extrinsic value in a leisure scenario?

In 1995, Tefko Saracevic in a SIGIR paper wrote that the most significant question for IIR would be “How does all this information, and associated information technology and information systems affect our [...] leisure, society, culture? How do IR and related applications reorder life?” [25]. For example, Elswailer et al. [16] have reported search behaviours that break traditional IIR models after gathering user’s self-reported search experience on social media and studying people’s information seeking behaviours when watching television. They mentioned that social media sites, mobile devices and other pervasive technologies have made information accessible to people in leisure scenarios and open up casual-leisure⁴ search behaviours motivated by hedonistic needs such as having fun, or relaxing instead of a well-defined information need. During search sessions users might find topically irrelevant information but they may keep exploring because the IR system satisfies their current leisure need.

Therefore, if a casual-leisure search user has an undefined information need or a hedonistic motivation, why should the search experience be based on an interaction model that assumes

⁴Casual-leisure is a name coined by sociologist Robert A. Stebbins in his theory of Serious Leisure that refers to a “the practice of doing what comes naturally” in order to experience fun or joy. Serious leisure theory defines other two categories: Serious leisure and Project leisure. Serious leisure will be discuss further in the next chapter.

the existence of this information need and the user’s capacity to communicate it? [12, 13]. How should the search experience be designed when people just want to have fun or explore to kill time? How would this influence the relevancy or perception of “the results”? Elswailer et al.’s [16] observation of search behaviours like “Needless Browsing” or “Exploring for the Experience” highlights the importance of designing the search experience over and above satisfying users with topically relevant results (i.e. quick lookup search and then moving on) because when users have strong hedonistic motivation, they could engage exploring beyond what they planned and with more probability discover unknown information while having fun. However, their research was limited to the users self-reported experience, and they did not assess the user’s engagement, exploration and interaction with an IIR-system.

Accordingly, there are several unanswered questions. How should the search experience be designed to encourage exploration and engagement? What are the effects of designing the search experience based on curiosity arousal mechanisms? Could the search interaction be designed to entice user’s exploration and engagement even without presenting topically relevant results? As Karen Spark-Jones mentioned in 1990, IIR should do much better than accept situational judgments as unique possible source evaluation because ‘I liked what I got’ from my search results could also imply ‘I am not worried about what I did not’[15]. The main hypothesis of this PhD work is that incorporating curiosity in the design of IIR systems could increase exploration and some attributes of user’s engagement which could determine the way users evaluate their own search experience.

1.2 Overall aims and goals

This research aims to understand search beyond a finding or lookup problem [14] . The main goal is to gather evidence of the benefits of applying psychological curiosity theory to the design of IIR search experience (i.e. curiosity-driven search experience) and identify a set of practical design principles that researchers and practitioners could used to design a better search experience [18].

A second goal of this research will be to comprehend how the experience of the search process (the journey itself) could become as important as the results or outcome (the destination) [16] (e.g. “this is not topically relevant, but it’s very funny, I can’t stop exploring”, “there are a lot of not interesting information, but I want to keep watching what new information appears”) when users’ information searching process is evoked by curiosity and not an knowledge gap [1, 19].

For this research, a “better search experience” can only be captured by using both pragmatic and hedonic features of information searching. Thus, this PhD work adopts user’s engagement [26] (e.g. absorption, felt involvement, novelty, etc) and exploration [27] (e.g. length of the search session, amount of information found) constructs to study and evaluate information interaction with IIR-systems.

1.2.1 Objectives

The three main objectives of the thesis are as follows:

- Objective 1: To identify evidence of an increment in the users engagement [26] and users exploration [27] of searchers when interacting with an IIR system design based on curiosity arousal principles
- Objective 2: To find a suitable experimental set up to study casual leisure search “in action.”
- Objective 3: To suggest and investigate possible explanations for observed search behaviours based on the curiosity principles mapped in the IIR-systems.

1.2.2 Research Questions

The purpose of this PhD research is summarized by these questions.

RQ1 Does curiosity-driven search design increase users’ engagement during casual-leisure search?

RQ2 Does curiosity-driven IIR-system increase users exploration during casual-leisure search?

RQ3 How could casual-leisure search behaviour be studied in a formal experimental setting?

Figure 1.1 illustrates the relationship between the research questions and the areas of knowledge cover by this PhD Work.

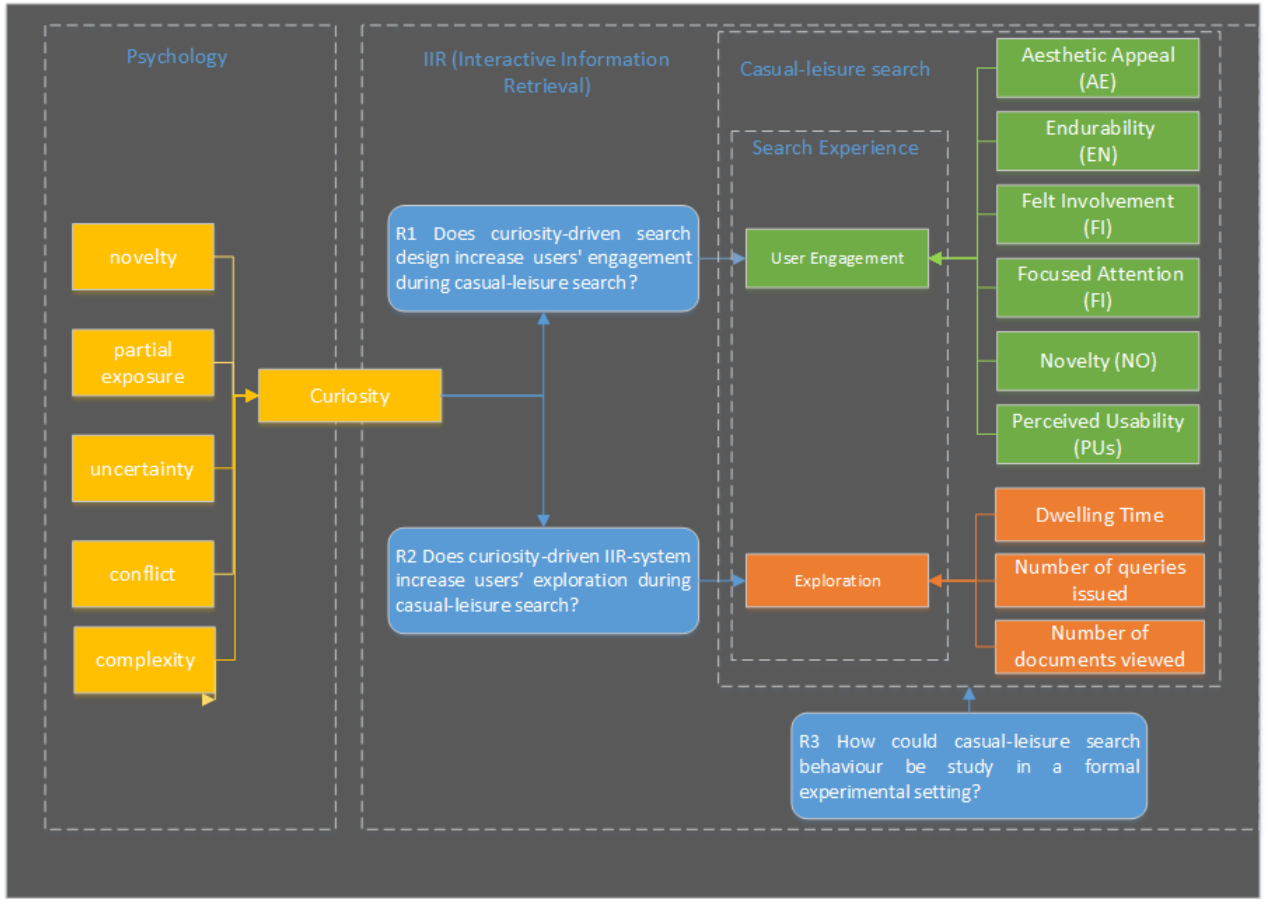


Figure 1.1: Research Questions and related areas on knowledge

1.3 General Research Design

The overall research design followed an user-centred approach both in the laboratory and naturalistic scenarios [28]. Figure 1.2 depicts the main building blocks of this PhD work.

The qualitative research began by determining a body of Information Behavior (IB), Interactive Information Retrieval (IIR) and Psychology (e.g. curiosity, flow theory). For the literature review two potential strategies: (a) a broad literature search in electronic databases (ACM, IEEE, Science Direct, Springer) and (b) a more focused approach that starts with articles that are known to be relevant in the topic of Casual-Leisure Search or Fun Information Interaction (FII) and then their bibliographies. The second one was selected because it promised better precision and a higher gain over the invested time .

After the identification of a pool of relevant articles, the analysis began. Based on titles and abstracts, some new articles were added from Human Computer Interaction (HCI) and other related fields.

In parallel with construction of the literature review, the search scenarios⁵, the tools and the setup for the first user study were created. The software development followed partially the XP methodology [30] and the idea of quick prototypes⁶ to test the assumptions coming from the literature analysis.

The first user study was done during the spring of 2014 where some participants interacted with the newly develop software application based on concepts taken from physiological curiosity theory and others with the Twitter search service, then they followed the experiment protocol and answered the questionnaires assessing their search experience. The third chapter contains a detailed description of the set up, with results and findings.

The second users study took place during the summer of 2015 where some participants interacted with a news search application based on the principles of physiological curiosity and the lessons learnt from the first user study. The fourth chapter details the set up of the laboratory, the application, the results and findings.

Both studies comprise the main empirical evidence for this PhD. For data analysis and collection of data, this research has used SurveyGizmo⁷, Excel and statistic programing packages to follow a more in depth analysis over the variables, concepts and relationships discovered during the user studies

At the end, this research provides contributions to the theory of casual-leisure search regarding the the evaluation and the methodology to study this casual-leisure search scenarios in action with “high ecological validity”[31]. Based on the empirical evidence, practical recommendations are given concerning the importance of mapping curiosity theory concepts to the search experience design specially for casual-leisure search scenarios.

1.3.1 Social Media Scenario

Microblogging content from Twitter has been used as the main data source and use case scenarios because previous casual leisure behaviour researchers have mentioned microblogging as a very important scenario for casual leisure search behaviour [16, 31]. Theses studies have shown the existence of “real” and “common” casual-search behaviours after user’s share their search experience in a corpus of 2.4M tweets. Thus, their research was limited to the self-reported tweets, and they did not asses the user’s engagement, exploration and interaction with Twitter’s search service or other IIR-system.

⁵In IIR, search scenario is a short cover story that describes the situation leading to the information need based in Borlund’s Simulated Work Task [29]. However, this PhD does not address work task context rather extends Borlund’s definition of simulated search task to the leisure context

⁶The word prototypes is used to describe incomplete versions of the software program being developed

⁷<http://www.surveygizmo.com>

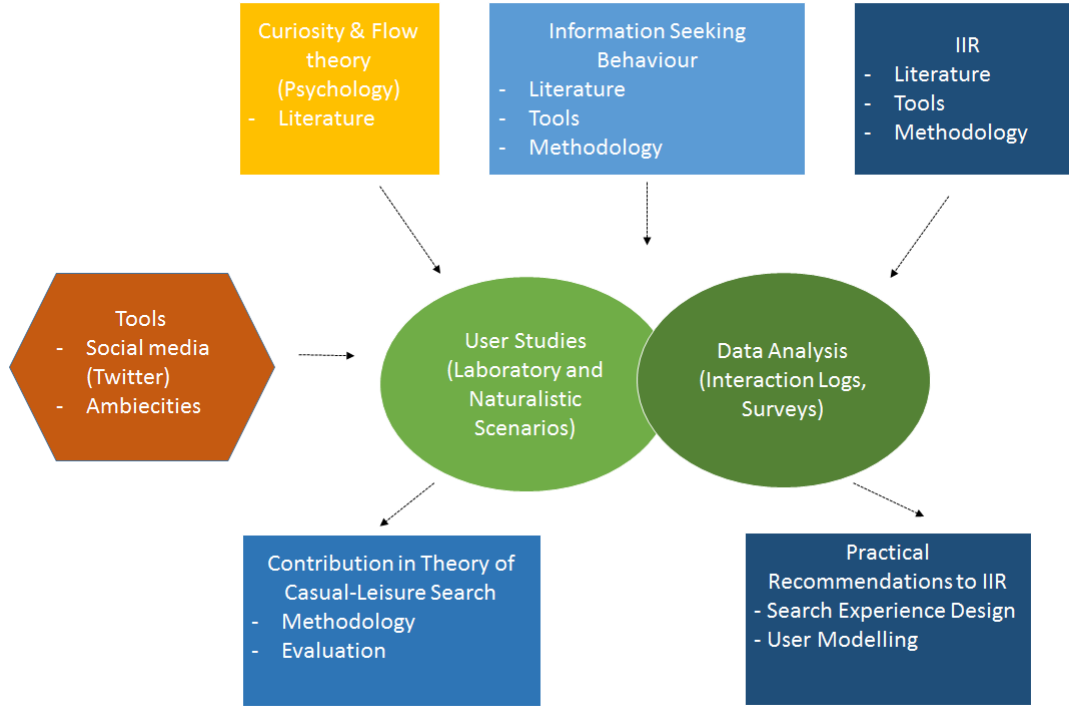


Figure 1.2: Overview Research Design

Therefore, this research focuses on studying casual-leisure search interaction in action (i.e. during and after the search session) [32], and how to improve user’s engagement (i.e. the quality of user experience that describes a positive human-computer interaction) [26] and exploration of a IIR-system by evoking curiosity⁸ driven search sessions. Instead of designing the search experience only with an utilitarian and pragmatic view of information seeking (i.e. to quickly solve information needs and then move on), this research gather evidence of the benefits designing the search experience (and IIR-systems) using curiosity arousal principles and theories in various leisure search scenarios where hedonic aspects of information searching could be more important than the perceived utilitarian value.

This PhD work employs user experience (UX) as a lens for researching IIR behaviour in social media leisure scenarios [26, 33] and traditional IIR evaluation [28]. Does curiosity-driven search experience affect focused attention (i.e user’s absorption, and temporal dissociation), felt involvement engagement’s factor (i.e. users feelings of being drawn in, interested, and having fun during the search), endurability engagement’s factor (i.e. users overall evaluation of the experience, and its perceived success), novelty engagement’s factor (i.e. users level of interest in the task and curiosity evoked by the system and its contents), perceived usability

⁸Human curiosity is referenced assuming the premise that curiosity main purpose is not amass knowledge but bring pleasure or joy.

engagement's factor (i.e. users affective and cognitive responses to the system) [26] during casual-leisure search on social media? Does the length of the search session (i.e. dwelling time) of a casual social media searcher increase when exposed to a curiosity-driven search experience? Does curiosity-driven search experience design pique the curiosity of participants even when content is not topically relevant? Does curiosity-driven search experience design encourage a more diverse and in depth exploration of participants? Does curiosity-driven search experience affects user's relevance judgements?

The research collected throughout the course of this PhD work project is used to determine whether the findings lead to a rejection of the main null hypotheses. As mentioned in the previous section, the principal source of data for this research comes from user studies where participants, their use of IIR-systems and user's experience are observed, recorded and analyzed.

These are the current null hypothesis based on the research questions and the causal-leisure search 'simulated' scenario for social media.

H₀₁ Participants having a curiosity-driven search experience do not feel more user's engagement [26] than those using classical query-response paradigm.

H₀₂ Participants using a curiosity-driven search experience design and the query response paradigm expend equal time exploring [28].

H₀₃ The search behaviour of participants is equally related to both curiosity-driven search experience design and query-response paradigm.

H₀₄ Participants using a curiosity trigger search experience and the query response paradigm equally asses their search experience.

The first null hypothesis are associated with RQ1; the second with RQ2; third and the fourth with RQ3. RQ1 are taken from the User's Engagement Scale (UES) by O'Brien and Toms [26]. The third null hypothesis is related with the study of casual-leisure search and the observation of search behaviours similar to "Needless browsing" or "Exploring for the experience" [16, 31].

1.4 Scope and Limitations

The selected empirical research design tries to balance qualitative evidence and quantitative interaction signals to understand and identify casual-leisure search behaviour within the scope of the simulated search scenarios.

Although this empirical approach and the social media scope might have an impact on the generalisability of the results and reproducibility of the experiments, there must be a trade off

between understanding “real user’s behaviours”, the probability of discovering “general laws” and other adjacent limitations around this PhD project (e.g. budget, location, users, personal bias).

1.5 Dissemination

During five years, this research has published the following three papers(Appendix A):

- Millan-Cifuentes, J. D., Göker, A., Myrhaug, H., & MacFarlane, A. (2014, August). Curiosity driven search: when is relevance irrelevant?. In Proceedings of the 5th Information Interaction in Context Symposium (pp. 279-282). ACM.
- Millan-Cifuentes, J. D., Göker, A., & MacFarlane, A. (2014, August). Designing autotelic searching experience for casual-leisure by using the user’s context. In Proceedings of the 5th Information Interaction in Context Symposium (pp. 348-350). ACM.
- Millan-Cifuentes, J. D. (2013, September). Context modelling for serendipitous discoveries in exploratory mobile search over social media. In Proceedings of the Fifth BCS-IRSG Symposium on Future Directions in Information Access (FDIA 2013)

1.6 Outline Thesis

The remaining chapters are as follows:

Chapter 2. Literature Review shows relevant theories and studies zooming from a ”macro-level” of information behaviour and other physiological constructs to the “micro-level” of Interactive Information Retrieval. While different important contributions and models are presented, the section questions the classical query based interaction paradigm, and highlights the role of emotion (rather than cognition) in the searching process specially when the motivation has an hedonistic background. This chapter also exhibits the research gaps identified and addressed by this PhD project.

Chapter 3. Casual Leisure Social Media Search contains the first user study done in this research with two main novel contribution. First, participants using curiosity-driven search had longer search sessions and experience more focused attention engagement’s factor than those using query-response paradigm even-though the content presented by the newly develop application was not topically relevant. Supporting the premise that the interaction design itself (without topically relevant content) was enough to generate

curiosity, encourage participants to explore for more time than they planned and engage in their search journey. This finding also suggests that users with high focused attention and curiosity assessed their search experience using other metrics beyond topical relevance (e.g. the feeling of temporal or spatio-temporal disassociation, the level of curiosity, the feeling of being drawn and involved in the exploration). Second, using the hedonic value of curiosity this research was able to trigger and observe search behaviours such as needless browsing and exploring for the experience in a control environment. The section describes the construction of a prototype of social media search using curiosity arousal mechanisms and the settings of the user study based on the extracted concepts from the literature review.

Chapter 4. ‘Curious’ query-response design for leisure news search captures the second experimental setup of this PhD research. At the beginning, of this chapter summarises the lessons learned, and discusses limitations from the first study to then provide a motivation for this second leisure news study. Next, the chapter describes how novelty and uncertainty curiosity arousal principles are applied in the information interaction and the software design of the search prototype created for the leisure news search study. This section specially discusses the first version of the curious companion software module. The prototype uses Twitters social media API to get relevant news and comes with two flavours of information interaction: cold query-response and curious query-response. Finally, the chapter reports the experimental design, the data collection, the search scenarios as well as a general demographic description of the participant; then presents the results of this work to determine if taking curiosity arousal principles into account in the design of IIR systems is advantageous in a leisure scenario from both utilitarian and hedonic perspectives of the search experience. This chapter also explores the role of personal interest and curiosity as trait in the searching process.

Chapter 5. Conclusion This chapter returns to the research questions introduced in Section 1.2. The research contributions are outlined, along with an argument for the analysis approach. Some time is spent reflecting on the implications for academic research Interactive Information Retrieval. By critically engaging with the limits of this research as well as its contributions, a research agenda is also identified. The chapter concludes by stating the thesis argued from this research.

Bibliography An extensive list of reference material and sources gathered during the course of this PhD project.

Appendix The appendix includes three published articles.

Chapter 2

Literature Review

IIR unified computer science and library and information science (LIS) to define the study of interactions of people, information, and IR systems [5]. IIR focuses in the Searching Information Behaviour that is *“the ‘micro-level’ of behaviour employed by the searcher in interacting with information systems of all kinds. It consists of all the interactions with the system, whether at the level of human computer interaction (for example, use of the mouse and clicks on links) or at the intellectual level (for example, adopting a Boolean search strategy or determining the criteria for deciding which of two books selected from adjacent places on a library shelf is most useful), which will also involve mental acts, such as judging the relevance of data or information retrieved”* [34].

Although this research focuses mainly in the field mentioned in the previous paragraph, this section firstly highlights some important studies of curiosity to provide a glimpse of the rich and long research tradition of this subject in the field of psychology, and show the potential of using psychological curiosity theory for designing the search experience and understanding exploratory behaviour. Secondly, the research presents a small sample of information behaviour models and frameworks to provide a holistic picture of the relationship, and the historical background between interactive information systems and an individuals information-seeking behaviour. Certainly, to design positive search experiences requires understanding of how people seek information [35]. Thirdly, an overview of information retrieval interaction models, methodologies, evaluation measures and approaches is presented to set the background for this research. Fourthly, literature related to context, and IIR is examined to show the vital role of context in the decision making process of searching and how has been used so far to enhance IIR-systems. Finally, this work discusses and reasons on current limitations of the IIR field by using casual leisure search research (i.e. search for fun). Overall this literature review demonstrates that information searching is a complex, meaningful and sometimes irrational decision making process where emotion (rather than cognition) and other contextual factors could have

a large impact in the user's search experience and satisfaction. Also this section questions the classical query-response paradigm of IIR and the oversimplified view of digital search as only a finding process that satisfies a short-term information need expressed by one or two terms.

2.1 Curiosity

Researchers in Psychology have associated curiosity with both emotion (e.g. fear, fun, boredom, anxiety, humor) and exploratory behaviour. Similar to other desires (e.g. for sex, or food), curiosity is related with approach behaviour and experiences of reward [18, 36].

Curiosity is often represented as positive emotion, intrinsically rewarding and highly pleasurable [21]. However, discovering may also be rewarding because it dissipates undesirable states of uncertainty or illiteracy rather than stimulates one's interest [18, 36]. These views are acknowledged by various theoretical models of curiosity such as the curiosity-drive theory, optimal arousal theory, I/D models of curiosity, and Wanting/Liking model of curiosity [1]. This section presents a sample of curiosity theory models to build a motivation for designing products and service, beyond IIR or computer science, to evoke human curiosity and, perhaps in the future, to create machines that incorporate this important feature reflected in both animals and humans alike.

2.1.1 Curiosity-drive theory

Curiosity-drive theory, initiated by Daniel Berlyne, defines curiosity as a subject's response to a novel, complex or ambiguous stimuli, which trigger a relatively unpleasant experience of uncertainty and motivates an individual's exploration, and acquisition of new information [18]. The individual is rewarded with positive emotions after the exploratory behaviour has reduced the unpleasant feeling of uncertainty. For example, relating experiments on curiosity and memory, Berlyne found that answers for questions rated as more complex and difficult were better remembered, indicating that learning was reinforced by the degree of curiosity that was reduced [18].

Berlyne categorizes two types of curiosity: perceptual and epistemic. Perceptual curiosity is evoked by complex or ambiguous patterns of sensory stimulation such as sights and sounds, and motivates behaviors such as visual inspection in order to acquire new information [18, 37]. Epistemic Curiosity (EC) is a uniquely human desire for knowledge (or "drive to know"), evoked by inquisitiveness and experimentation, that motivates people to learn new ideas, and solve intellectual problems [18].

Another contribution of Berlyne was his differentiation of curiosity having direction between specific and diverse. Specific exploration was defined as a response to either epistemic or perceptual stimuli with high degree of novelty and complexity. For instance, the desire to close an information gap. On the contrary, diversive exploration was triggered by boredom, or un-stimulating environment, and describes the behaviour of actively seeking stimulation.

Later, Loewestein (1994) [36] described curiosity as consequence of an unpleasant sensation of “deprivation” called “gap in our knowledge” which is decreased by exploratory behaviour. He posits that knowledge acquisition is the primary goal of information seeking. Although, Loewestein recognises that pure interest could trigger information seeking behaviour, he argues that such motivation should not be named curiosity.

Loewestein main idea is that as the perceived degree of knowing becomes stronger (e.g. tip of the tongue), knowledge gaps will appear smaller, and states of curiosity will grow as individuals feel close to eliminating their knowledge discrepancy [36]. Therefore, his main contribution was the relationship between the magnitude of knowledge gap stimulation and curiosity.

2.1.2 Optimal arousal theories

However, the fact that individuals would start exploring before any given stimulus, or they seem to look for opportunities to have their curiosity aroused cannot be easily explain by the curiosity drive reduction theory. Other limitation of the curiosity drive theory was the observation that some individuals would experience more curiosity than others in a similar environment. Thus an alternate view of curiosity was needed.

Using Wundt’s hedonic theory [38], Spielberger & Starr (1994), among other previous researchers, argued that animals and humans are motivated to maintain a pleasurable state named “optimal level of stimulation” where being over- or under-aroused is unpleasant [39]. Then, curiosity is represented as a non-linear relationship between pleasant states of curiosity and aversive conditions of anxiety that are aroused by a novel stimuli or situation (i.e. the Wundt’s Curve). Spielberger & Starr (1994) explicitly addressed the limitations in Berlynes work (1960) by including the state trait distinction of emotion and personality traits; and considering curiosity as an emotion associated with pleasure and enjoyment.

In contrast to the drive-theory, optimal arousal model assumes curiosity induction is rewarding, and involves feeling of interest instead of uncertainty [39]. For instance, if individuals are bored, they will be motivated to increase their arousal to reach an optimal level, and will explore an environment in search of a stimuli that may excite their curiosity and generate positive feelings. After new information has been acquired, and boredom quickly reappeared,

individuals might be inclined to pursue new stimulation once again [19]. Thus, Spielberger & Starr described exploration as aiming to increase pleasurable states evoked through curiosity.

Also, their inclusion of curiosity and anxiety as personality traits explain the influence of personal difference in the arousal of parallel emotions states and behaviours. Based on an empirical study, they observed that people with high levels of trait curiosity and anxiety presented more information seeking behaviour in nonthreatening conditions [39].

Nonetheless, changes in the “arousal” curiosity do not always demonstrate the expected Wundt’s curvilinear relationship with efficiency in task performance, suggesting that what defines “optimal” level of arousal may depend primarily on the situation. Besides, optimal level of stimulation theory fails to explain why individuals would want to learn new information if this ultimately leads to boredom.

2.1.3 Interest/Deprivation Model of Curiosity

Litman & Jimerson’s (2004) [19] initial contribution integrated the curiosity-drive theory (e.g. Loewenstein, 1994) [36] and optimal level theory (e.g. Spielberger & Starr, 1994) [39] by considering that both the satiation and activation of curiosity could be rewarding.

In Litman & Jimersons viewpoint, interest (I) induction and deprivation (D) elimination consider distinct types of curiosity that correspond to very different motives for acquiring new information [19]: I-type curiosity involves the anticipated pleasure of new discoveries (CFI - feeling-of-interest). For example, an activity done for the joy or inherent interest in doing it, where individuals recognize opportunities to discover something entirely new such as reading a juicy gossip fashion magazine, or casually watching advertisement with an amusing anecdote. Whereas D-type curiosity is concerned with reducing uncertainty and eliminating undesirable states of ignorance (CFD - feeling-of-deprivation). For instance, an activity done to achieve an external outcome, such as pay or career. D-type curiosity is stimulated when people lack information they intend to incorporate into an existing body of knowledge.

CFI is a pleasurable state about the amusing means of acquiring information which can increase learning and well-being [1]. In a similar way, Kashdan et al. [21] noticed the benefits of the feeling of interest (or openness to experience) when developing their Curiosity and Exploratory Inventory to measure curiosity. They research the effect curiosity in exploratory behaviour and the feeling of cognitive absorption (e.g losing track of the time). Taking the absorption concept, he explains overlapping features between curiosity and Csikszentmihalyi’s [23] ‘flow’ theory, the idea of an “optimal level of experience” in which a person is fully immersed by perceiving the correct balance between the challenge and the skills to face it.

In contrast to CFI state, individuals who seek new information or knowledge to solve an unmet need or uncertainty have an inclination to experience feelings of anger or anxiety [1, 19]. Therefore, for D-type curiosity is rewarding to acquire new information because it reduces negative feelings attributed to uncertainty.

Litman (2008) [8] summarizes four studies where trait measures of I-type of curiosity and D-type curiosity have been found to load statistically different factors and have correlated with measures of positive or negative emotions respectively. Drawing from these studies, Litman posits that D-type curiosity (or “need to know”) corresponds with qualitatively more intense expressions of curiosity and exploration than I-type. However, he also mentions that I-type curiosity resembles a relaxed and pleasant “take it or leave it” feeling towards new knowledge; acquiring new information is viewed as potentially pleasurable but not a necessity. When I-type curiosity is evoked, situations characterized by uncertainty are viewed positively, and chances to resolve that incertitude are considered as potentially pleasurable. Accordingly, Litman observed as well that I-type curiosity tendencies measurements were negatively correlated to negative affective tendencies such as anger, anxiety, and depression.

2.1.4 Wanting/Liking Model of Curiosity

Despite Litman & Jimerson (2004) I/D model explains both reduction and induction theories of curiosity, they did not explain how curiosity can feel like “interest” (I-Type) or like “deprivation” (D-Type), how these qualitative differences could be related with observed difference in information seeking behaviour [19].

Litman (2005) [1] theorized that I-type and D-type curiosity could be understood by varying levels of wanting and liking neurobiological systems [40]. He explains that wanting and liking affect appetitive motivation and the resulting feelings of pleasure through dopamine and opioid activation. Wanting and liking systems have been implicated in various studies of different kinds of stimuli such as food, water, drugs, and sensory stimulation in both humans and animals [40].

Wanting mirrors the intensity of “appetite” for new information (i.e., desire of new information), whereas liking referenced the “joy” (or degree of pleasure) anticipated from the acquisition of new information (e.i. expected and derived pleasure of of new information) [1]. Although wanting and liking systems might be activated contiguously (i.e. stimuli that are wanted are also likely to be liked), they are two physiological systems operating through a different neural circuit. Thus, they can be independently activated. For instance, addicts who crave drugs (high wanting) might not experience much pleasure (low liking) from taking them.

Litman described the four combinations of wanting and liking (Figure 2.1 depicts this cu-

<i>Wanting</i>		
	<i>Low level</i>	<i>High level</i>
<i>Liking</i>		
High level	Curiosity as a feeling of “interest” (Aesthetic appreciation)	Curiosity as a feeling of “deprivation” (Perceptual/conceptual/fluency)
Low level	Ambivalent disinterest or boredom (Spontaneous alternation or novelty seeking)	Need for uncertainty clarification (Need for cognitive closure; morbid or lurid curiosity)
I-D, interest-deprivation.		

Figure 2.1: Litman [1] Wanting / Liking Curiosity Model

riosity model [1]). First, “sub-rational liking” (high liking and low wanting) corresponds to I-type curiosity where pleasurable affective experiences are associated with an enjoyed (but not sought) stimulus. Analogously, food can be enjoyed even when an individual is not hungry. Regarding exploratory behaviour, Litman explains that individuals enjoy browsing and searching new pieces of information that stimulate their curiosity even if they may not have any specific question. In a similar way to aesthetic appreciation, in which the liked stimulus provides a hedonic value based on novelty and complexity qualities, but not a utilitarian value [18]. Second, “intensive craving” (high liking and high wanting) matches D-type curiosity because individuals who experienced a knowledge gap are in a need-state (“need to know”) that strongly motivates the acquisition of pertaining and relevant information until satisfaction. Third, “ambiguity tolerance” (low liking and low wanting) can be linked to ambivalent disinterest or boredom. For instance, individuals recognize an information gap, but are neither experiencing much pleasure from new information and knowledge, nor they are worried by the absence of information, or their uncertainty state. According to Litman, drive reduction, optimal stimulation state and I/D theories have overlooked this behaviour where individuals realise their lack of information but their curiosity is not evoked (i.e. “I don’t know, and I don’t care”). Four, “irrational wanting” (low liking and high wanting) is portrayed by Litman as a strong desire for information to reduce uncertainty (i.e. unpleasant state), even when the information is anticipated to have limited value in stimulating interest or improving understanding. For example, morbid curiosity motivates individuals to seek information with gruesome or vulgar violence and death scenes that they might dislike and find unpleasant.

An important limitation of Litman’s methodology [1], when testing and creating the wanting-liking model, is that tendencies to experience emotional-motivational states were examined rather than the actual states and following behaviors. Thus, the evidence provides indirect proof of the wanting-liking model.

2.1.5 Curiosity Arousal Principles

Berlyne [18] identified four collative variables or external factors that govern various forms of stimulus selection. They are novelty, uncertainty, conflict, and complexity. Each collative variable has the quality of arousal potential, the ability to affect the intensity of arousal. Berlyne [18] used the word ‘collative’ because to decide how similar or different a stimulus pattern is, the evaluation must compare or collate information from two or more sources. Based on his theory, a newsreader might choose an article depending on how novel the headline is, to what degree it arouses uncertainty, to what extent it arouses conflict and how complex it is. The same can be said from any information (e.g. Tweets, Web Pages).

Novelty refers to something new. The stimulus can be characterized as completely new (complete novelty), recently acquainted (short term novelty) or seldom experienced (long term novelty). Based on the aforementioned observations, Berlyne [18] proposed three criteria to measure novelty: the extent to which a stimulus is novel will be inversely related to (1) how often the stimuli have been experienced before (frequency), (2) how recently the stimuli have been experienced (recency), and (3) how similar stimuli have been previously experienced (similarity). For example, a social media consumer would be attracted to read an old Tweet if the text and media content offers a unique value, or a new Tweet if the information has been recently shared.

Novelty is often related to other supplementary properties in a given stimuli pattern such as change, surprisingness, and incongruity. These properties could influence exploratory behaviours. Change might reveal some insights in the exploratory directions (e.g. a Tweet changes its re-tweet status from 0 to 200 in 1 minutes). Surprisingness appears when there is a stimulus that produces an expectation, and meets the expectation (e.g. a Tweet with an unexpected funny photo and a link to more information). Incongruity indicates an expectation that is not met by the same stimulus (e.g. searching in social media for Cartagena, Colombia but then getting results from Cartagena, Spain in top of the page).

Uncertainty arises when an individual finds difficult to classify and respond a stimulus.

According to Berlyne [18], the likelihood of uncertainty caused by a stimulus can be estimated (1) formulating an inventory of stimuli that might occur (as response to the stimulus a given stimulus), (2) partitioning into classes, (3) assigning a probability to each class, and (4) using Shannons equation to compute the entropy.

Conflict emerges when a stimulus generates two or more incompatible responses. Berlyne [18] argued that conflict is positively associated to (1) the equality in strength of competing responses, (2) the absolute strengths of competing responses, (3) the number of competing responses, and (4) the degree of incompatibility between competing responses.

Complexity generally is defined as the amount of variety or diversity in a stimulus pattern. The three most obvious properties that determine the degree of complexity are (1) the number of distinguishable elements in a stimulus, (2) the dissimilarity between these elements, and (3) the degree to which several elements are perceived and responded to as a unit [18].

In summary, Berlyne's theory explains that there are some common properties (collative variables) in a stimulus causing or eliciting curiosity. Collative variables are quantitative, and they could be used to potentially measure the stimulation level of curiosity of a given stimulus. Therefore, potentially, information systems can be designed and evaluated to increase the stimulation level of curiosity.

2.1.6 Other Types of Curiosity and Discussion

As shown in this section, curiosity is the intrinsic desire for new information that will stimulate interest or reduce uncertainty. Moreover, other kinds of curiosity could be specified if one takes into account the type of information that evokes the curiosity state. There are at least two more known types of curiosity: 1) Litman & Pezzo (2007) defined interpersonal curiosity (IPC) as the intrinsic motive to seek people-information such as their public and private behaviors, and also their thoughts and feelings [20]. They identified three IPC factors (curiosity about emotions, spying and prying, and snooping) which influence the arousal of D-type curiosity rather than I-type curiosity. 2) Litman et al. (2005) explained sensory curiosity (SC) as an interest in engaging in adventurous and thrilling experiences (i.e. novel and unusual sensory experiences) that involve relatively little imminent danger (e.g. "riding a horse", "hiking through a remote rain forest", "sailing around the world") [41].

For the purpose of this PhD work, the previous section provides enough background on the complex construct of curiosity by showing that curiosity research holds a rich and long tradition

in psychology. However, scholars from fields like computer science, library information and science, and others have been slow to embrace the concept, even though, curiosity is often the motivation behind exploratory behaviours and learning, and contributes to the well-being of individuals.

Although curiosity can be understood as a personality trait, previous research has shown that there are potential properties in a given stimulus (i.e. collative variables such as novelty, uncertainty, complexity, surprise and conflict) that could be modified to evoke with more probability a state of curiosity. In the next chapter, this research provides examples of previous applications of curiosity into information and computer science; and initial example of how by mapping curiosity arousal concepts in the search experience could evoke a positive feeling of interest and entice people to explore more time.

According to Litman & Jimerson [19], and Litman [1], when I-type curiosity is triggered, situations characterized by uncertainty are viewed positively, prospects to resolve that uncertainty are considered as potentially enjoyable and learning is reinforced. Therefore, this finding provides a strong motivation and for glimpse of the potential benefits of applying curiosity theory in the design of IIR-systems.

2.1.7 What is Curiosity for this research?

Curiosity has been often associated to knowledge-acquisition in a learning scenario, for which the accomplishment is understood on the basis of an objective and demonstrable proof. In other words, this trait has been defined based on its extrinsic value [36, 1].

However, recent psychological curiosity models such Wanting/Liking and Interest/Deprivation show that curiosity explains behaviours when individuals enjoy exploring new pieces of information that stimulate their curiosity even if they may not have any specific question, or goal [1, 8]. For example, “irrational wanting” category of Wanting/Liking Curiosity model posits that curiosity motivates individuals to seek information with vulgar violence and death scenes that they might dislike and find unpleasant [1]. Thus, curiosity is triggered even in absence of an extrinsic value.

Therefore, this PhD understands curiosity as an emotional-motivational state that could be activated by the intrinsic enjoyment of exploring as well as the extrinsic value of knowledge-acquisition and the reduction of uncertainty. The main focus of this research is to understand how this “intrinsic enjoyment of exploring” could be activated by using curiosity arousal principles [18].

In the following section, this research will introduce a sample of information behaviour

models to provide a historical context between interactive information retrieval systems and an individuals information-seeking behaviours.

2.2 Models of Information Behaviour

Wilson [34] defines Information Behaviour as “the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking, and information use. Thus, it includes face to-face communication with others, as well as the passive reception of information as in, for example, watching TV advertisements, without any intention to act on the information given”.

Models are simplified abstractions of complex aspects of the ‘real world’. Models describe important attributes and their relations from a particular perspective, and they are usually represented as graphic. Also, models can be defined at different levels of abstraction to deal with the complexity of the real phenomenon and then help researchers to reason on it. After systematic research, a model can become a theory.

Information Behaviour is a complex phenomenon. Therefore, researchers from LIS have created many models to describe different characteristics of Information Behaviour at different levels of abstraction (e.g. General Information Behaviour, Information Use, Interactive Information Retrieval, Information Retrieval) and context (e.g. Libraries, Internet, Work Scenarios, Leisure, etc). This section presents some influential models in Information Behaviour and their impact in LIS research. The following models will be discussed to provide a general understanding of information behaviour research and their historical biased towards job related scenarios for information seeking and use.

2.2.1 Dervin’s Sense-Making Model

Since 1972, Dervin and her associates have developed a field to comprehend the phenomenon making or unmaking sense of information by using the metaphor of gap-bridging (Figure 2.2).

Dervin’s Sense-Making model assumes that a person’s information behaviour can be understood as situation-evoked process in which an individual feels a ‘knowledge gap’(i.e. information need) [42]. When this happens, the ‘gap’ should be bridged by information that makes sense to the person’s goals, current situation, emotions and existing knowledge.

The Sense-Making model is related to a “paradigm shift” in LIS research from information sources to information users [42]. Learning from Dervin’s Sense-Making, LIS methodologies and research move from a simple classification process of information objects to a more holistic study

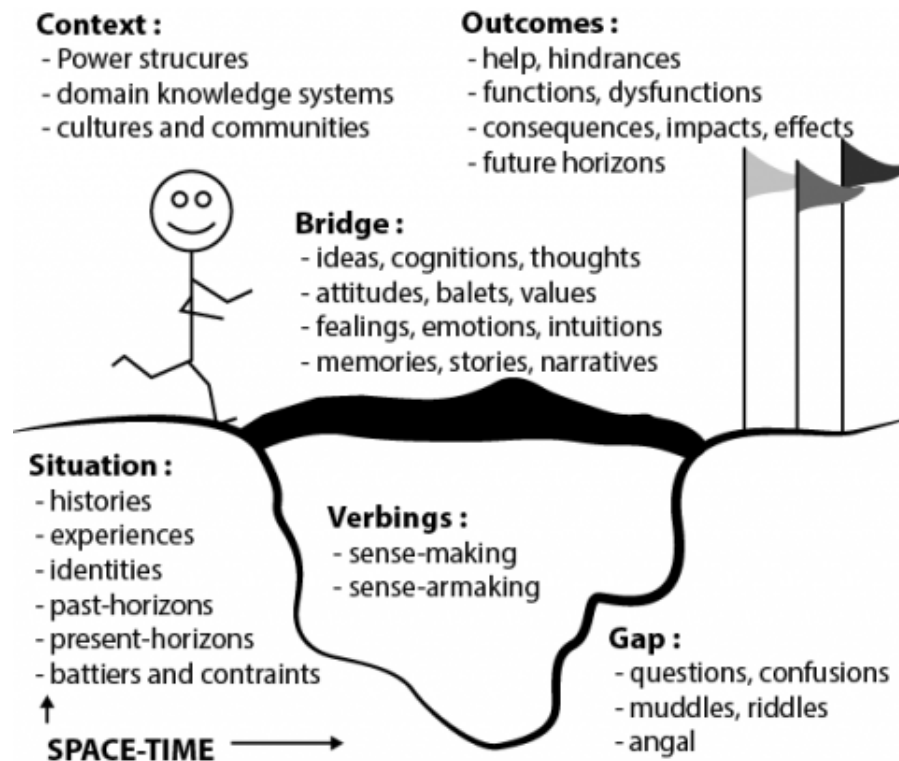


Figure 2.2: Dervin's Sense-Making Model

of the process by which people become informed [43]. Therefore, Dervin's Sense-Making has been frequently cited to describe and explain situations, gaps, and needs related to information seeking and use within a specific context.

Another characteristic of Sense-Making is the emphasis in the individual as a constructive actor rather than a passive receiver of information. By focusing on the user, Dervin's Sense-Making model recognises the importance of both cognition and emotion to every step of the gap-bridging process.

According to Pirolli and Card [44], the most common ways in which IIR systems support sense-making are the shoebox (i.e. potentially relevant documents into a single collection), the evidence file (i.e. snippets from the shoebox) and schema (i.e. meta representation of the information such as an ontology) [14].

Although Dervin's Sense-Making is biased by the LIS background, as a methodology offers a general and comprehensive way of studying the 'verbs' rather than the 'nouns' of information in "everyday" life.

2.2.2 Information Foraging

Since 1966, MacArthur & Pianka [45] developed the Optimal Foraging Theory to explain how animals decided what to eat. This theory understands animal's complex decision making process (e.g. hunting, gathering) using cost-benefit analysis. For example, Charnov [46] "marginal value theorem" explained that an animal should remain in a patch (i.e. a section of the animal's environment with potential food sources) as long as the resulting gain in time per unit of food exceeds the loss.

Although the human decision making process is by far more complex than bears or other animals, optimal foraging theory has been successfully applied to describe information seeking of scholars and web-browsing behaviour [47, 48].

In 1999, Pirolli and Card [49] proposed a framework called Information Foraging Theory in which they compare animals seeking behaviour to peoples information seeking in the web. They noticed that people measure "information value" from "proximal cues" known as information scent to guide them towards their seeking activity.

The information scent concept has been extensively applied by IIR-system designers to explain and predict web search. For example, there are well known important web search interfaces' techniques to guide users in their seeking process: descriptive titles, hit highlighting, query auto-complete, instant results and clear labeling [14].

Optimal Foraging Theory has many practical IIR applications and has provided tools to understand on-line information seeking within different contexts by modelling both the information seeker (e.g. perceptions, strategies) and the information environment. Nonetheless, Optimal Foraging Theory as research methodology it narrows the landscape of information searching behaviour to a rational and goal-driven act.

2.2.3 Wilson's Information-Seeking Behaviour Model

In 1981, Wilson proposed a model of needs and information seeking behaviour, and another of information use. Later in 1996, he corrected these models and merged them based on important developments in fields such as LIS and Psychology [2] (Figure Figure 2.3). Wilson's model characterized human needs in three groups (i.e. physiological, emotional and cognitive) and showed that these influence information behaviour. He indicates that instead of speaking about "information needs", it might be better to understand "information seeking towards the satisfaction of need".

Interestingly, Wilson argues that needs by themselves are not enough to trigger information seeking behaviour. He locates the information user as a "person in context" and suggests

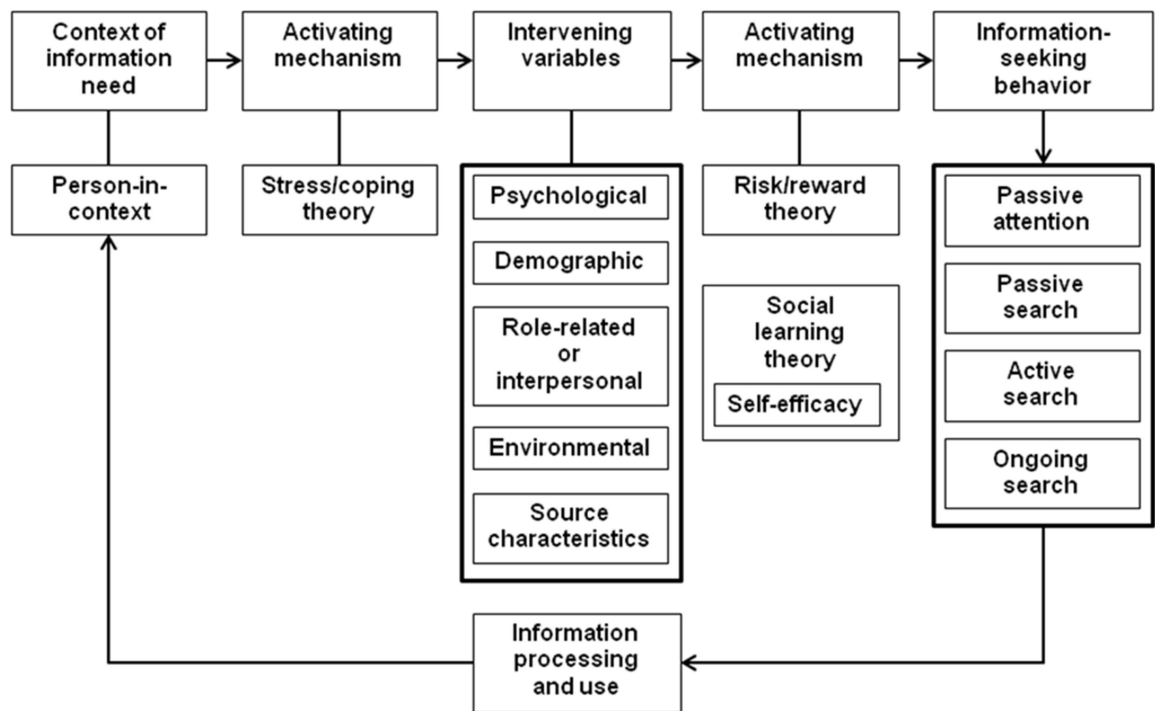


Figure 2.3: Wilson's Information-Seeking Behaviour Model [2]

that information seeking behaviour can be explained by risk/reward, stress/coping and social learning theory within a specific context (e.g. demographics, environment, psychological, etc).

Wilson's holistic model also includes both active and passive information behaviour, and the idea of a "feedback loop" that reflects the idea of learning whilst seeking for information.

2.2.4 Ellis's behavioural model

In 1989, Ellis started to build a model of information behaviour and IR from observing chemists, engineers, social and natural science researchers' information seeking patterns. However, Ellis's 'behavioural' model did not take in to account cognitive or affective aspects of the users information seeking [43].

Ellis model describes six common informational 'features', rather than 'stages', for information seeking even though the groups studied were from different backgrounds. The features are:

- Starting
- Chaining
- Browsing
- Monitoring

- Extracting

These categories, constructed by adopting ‘Ground Theory’ methodology, and their interactions constitute in essence the behaviour model itself. Although Ellis makes no casual claims and does not group them in sequence features, the model is widely referenced within the information behaviour field because as a methodological choice this model highlights that the study of search “behaviour offers a more tractable focus of study than cognition” [43]. Therefore, Ellis’s model does “not directly suggest hypotheses to be tested” but indirectly “a comparison of findings across several studies may suggest causative factors to be explored” [4].

In 1999, Wilson used Ellis’s features to represent information seeking as a process model where he grouped browsing, chaining and monitoring as a search moves within the information seeking selection/exploration stage whilst extracting is located in the collection of information sources [2] (see Figure 2.4).

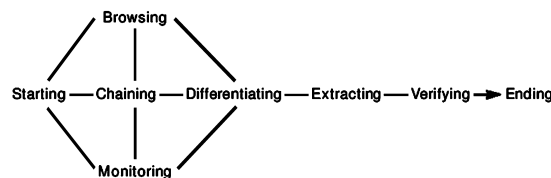


Figure 2.4: Ellis’s Information Seeking Behaviour Model illustrated by Wilson [2]

Ellis’s model has been used as a great methodological tool to zoom in and out between different levels of information behaviour (i.e. information seeking to information retrieval).

2.2.5 Information Search Process Model

Kuhlthau’s Information Search Process Model (ISP) [3], published in 2004, was the consequence of two decades of qualitative user studies of adolescent and college students. Kuhlthau’s model associated feelings, thoughts, actions and information tasks to different stages or sub-process of information seeking. Therefore, unlike Ellis’s model, ISP is based on the premise that information seeking can be grouped in sequences of behaviour. Figure 2.5 illustrates the ISP 2-dimensional model.

Kuhlthau’s ISP, grounded on personal construct theory, characterises six stages of the seeking process:

- Initiation
- Selection

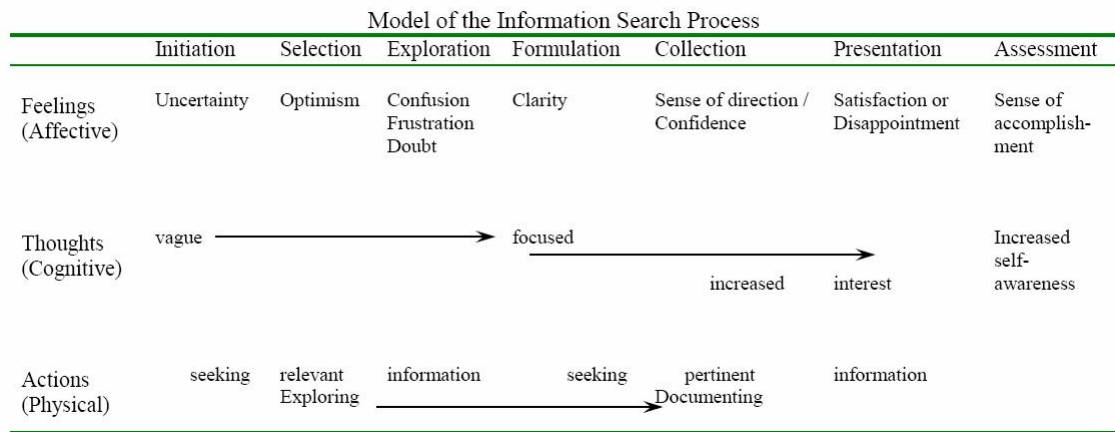


Figure 2.5: Kuhtlthau's Information Search Process Model (ISP) [3]

- Exploration
- Formulation
- Collection
- Presentation

The most fundamental contribution of Kuhtlthau's model is the combination between affective and cognitive approach. This methodological choice is expressed by the "uncertainty principle" where she defines uncertainty as a cognitive process that triggers feelings of anxiety and lack of confidence. Furthermore, the uncertainty concept divide the ISP in six components: process, formulation, redundancy, mood, prediction, and interest.

However, information and the movement through different stages of the information meaning process do not always reduce the user's uncertainty. Thus, Kuhtlthau, based on "the zone of intervention" concept, has suggested to practically guide and support the users at different phases of their seeking process assuming uncertainty as natural characteristic in the design of IR-systems. ISP encourage research beyond the query "to discover ways to enhance intellectual access that leads to learning, creativity and innovation" [43].

2.2.6 Work Task Information Behaviour

In 1995, Byström and Järvelin [50] created a model of information activities in the context of professional work by assessing task complexity as an independent variable of the information seeking process.

In this "feedback loop" model type, the seeking process starts with the identification of an information need based on a given work task, but she also mentioned four contextual precon-

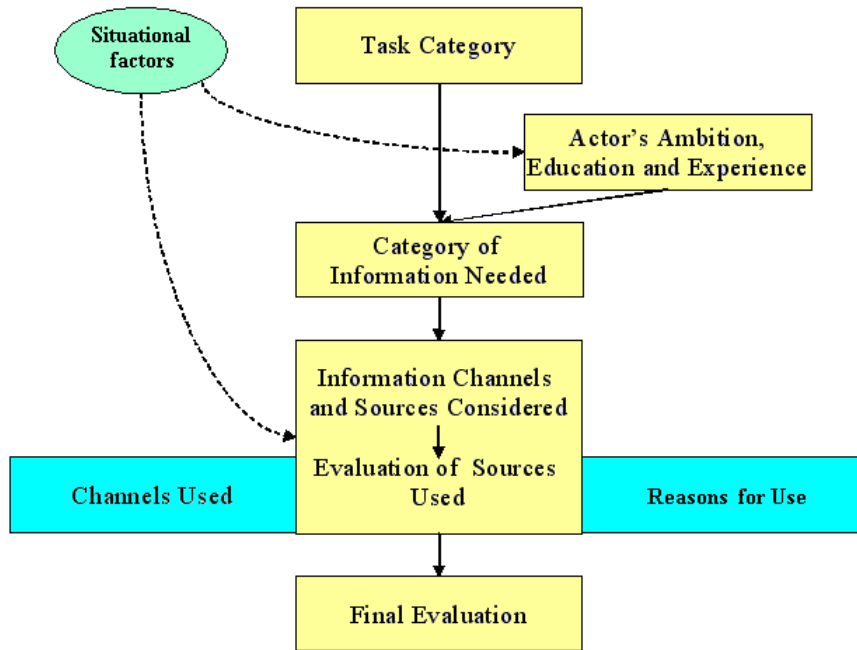


Figure 2.6: Byström and Järvelin's Work Task Information Behaviour Model [4]

ditions: personal factors (e.g. knowledge), situational factors, organization and information seeking style. Then to address the information need within the working context, the professional elaborates, executes and evaluates a plan of actions until the information gathered is good enough for the work task. Figure 2.6 represents the main components of this model.

As research method and theory, the model has connected research to real-life professional scenarios and bridged the gap between information seeking and information retrieval. Although work task information behaviour model uses only knowledge (i.e. novice vs experts) as a characteristic of the searcher and the professional field, it has been often quoted in the field of IIR as a practical framework to understand searchers' behaviour using the work task unit of analysis instead of job-level [4]. Byström's definition of task is general enough to study different contexts of information seeking including leisure scenarios [51].

2.2.7 ELIS Model

Since 1995, Savolainen built the Everyday Life Information Seeking (ELIS) model to study people's information behaviour in non-work contexts using empirical studies upon diverse ethnographic groups [52]. ELIS has explained daily use and selection of information using the sociological concept of "habitus" and more recently, the theory of "practice of everyday life" [53].

The main proposition of ELIS is that an individual's preference and use patterns are condi-

tioned by social and cultural factors. Savolainen describes information seeking as a continuous interaction between Way of Life (“how things are at this moment”) and Mastery of Life (“how they should be”).

Although ELIS fails to establish a causal hypothesis and fails to explain how a person seeks information in a particular situation, this model represents an initial step towards the study of leisure information, and non-purposeful seeking scenarios. ELIS emerges in LIS as macro-conceptual framework to think in everyday life information seeking behaviour.

2.2.8 Serious Leisure

Since 1973, sociologist Stebbins started working in leisure activities ‘...that people want to do and can do at either a personally satisfying or a deeper fulfilling level [54]. Serious Leisure Perspective (SLP) was developed by joining ethnographic field studies of numerous leisure contexts such as amateurs (e.g. musicians, actors, baseball players, football players, etc), hobbyists (e.g. cultural tourists, kayakers, snowboarders, etc) and career volunteers.

Stebbins defines leisure as “uncoerced activity engaged in during free time” [55]. Then, he defines three categories of leisure [54]:

- **Serious leisure** is a free-time enduring activity centered upon acquiring and expressing specialized skills, knowledge, or experience [56]. The usage of the word ‘serious’ suggests concentration and dedication, rather than coldness and austerity. The sub-types of serious leisure are: amateurism, volunteering, and hobbies; the last sub-type seems to be the most common of the three.
- **Project-based leisure** is “a short-term, moderately complicated, either one-shot or occasional, though infrequent, creative undertaking carried out in free time” activity [56]. These projects might demand considerable planning, effort and sometimes skill, nonetheless they do not require the enduring and lasting pursuit of serious leisure.
- **Casual leisure** is “an immediately, intrinsically rewarding, relatively short-lived pleasurable core activity, requiring little or no special training to enjoy it” [56]. In other words, casual leisure is hedonic and results in sensations of pleasure and enjoyment. There are six sub-types of casual leisure: play, relaxation, passive entertainment, active entertainment, sociable conversation, and sensory stimulation.

Serious leisure is neither an information behaviour model nor defines research methods but it is an interdisciplinary framework that has helped research of everyday life information seeking

and use. SPL was introduced to LIS by Hartel [57] who studied the information practices of hobby cooking.

Before Hartel's work, there was little research interest on linking LIS to leisure and everyday life scenarios. Some well-known exceptions are Savolainen's way of life information model and Ross [58] empirical user study to discover the relation between pleasure reading and serendipity. Therefore, serious leisure opens up an unexplored universe for LIS.

2.2.9 Casual-Leisure Behaviour Model

In 2011, Elswailer et al. [16] published a casual-leisure information behaviour model based on SLP, and empirical user studies of television information behaviour and the casual information behaviours described by users of Twitter.

In the television information behaviour study, using grounded theory, Elswailer et al. developed two coding schemes [59]: one describing a comprehensive set of information needs related to viewing television and a second grouping factors that might influence or motivate those needs. From this study, they noticed that the motivation given in diary entries of the user study were more related to the experience instead of the information found (e.g. kill time, relaxing, entrainment while they performed monotonous task, etc).

In the second study Wilson & Elswailer [31] collected a large Twitter corpus of behaviours containing search oriented keywords (e.g. search, browse, explore). From the analysis of tweets, they identified two general groups of casual behaviours: needless browsing and exploring for the experience. Also it seems many tweet authors were on state of flow [23] and engage while they described their experience rather than the information found.

After comparing both studies, they highlighted four different points of casual-leisure information behaviour which did not fit standard information behaviour models [16]:

1. Casual-leisure information behaviour tasks are triggered by being or wanting to achieve a particular mood or state.
2. The finding of information is often less important than experience of finding in casual-leisure situations
3. Casual-leisure needs are often linked with an ill-defined or absent information need.
4. Casual-leisure behaviours are not measure by the information found, but how the hedonistic needs of users are met

Figure 2.7 depicts Elswailer et al. casual information model as an extension of Ingwersen & Järvelin's [60] nested model and shows the four main differences listed by them in both user

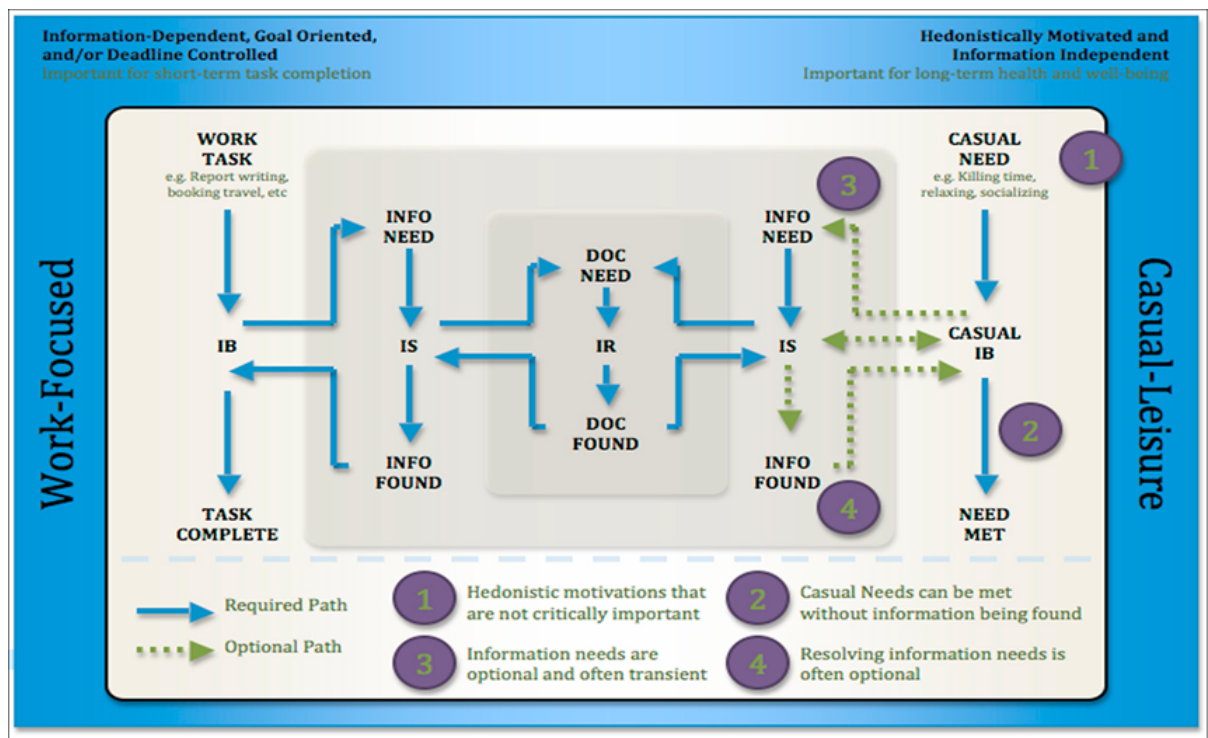


Figure 2.7: Original Casual-Leisure Behaviour Model

studies. However, Elsweiler et al. did not study leisure search in action and directly (e.g. an experimental setting where participants are asked to search for fun or in a given leisure scenario, then they evaluate their user experience while their behaviour is recorded by a researcher). For instance, during the social media study, they collected self-reported casual leisure search experience from Twitter (i.e. tweets with keywords like search or browse) to provide evidence for their behavioural model, but they were not able to show what particular aspects of the search experience engaged users and encouraged this exploratory behaviour because they did not record interaction data, nor ask their participants about particular aspects of the Twitter search service such as interactivity, or aesthetic appeal. Therefore they did not make any practical recommendation regarding how to improve the search experience for casual leisure information seeking scenarios.

2.2.10 Information Behaviour Literature Discussion

Although the concepts, models and theories presented are a small sample of relevant contributions in the active and versatile field of information behaviour, this literature expresses some of the principles and premises by which this research has been motivated. These identified contributions and concepts

- a) Users, rather than systems, are the most important and complex component in the process

of seeking and using information. For example, Sense-Making and Information Foraging practical applications have played a vital role improving the search experience of million of web users [14].

- b) From Devin's model onwards, the field of information behaviour has recognized the importance of understanding affective or emotional factors in the seeking process as well the cognitive and physical (e.g. Kuhtlthau' ISP [3]).
- c) Information behaviour might be trigger by both informational and hedonistic needs. However, the information behaviour motivated by a hedonistic need can be completely different from the behaviour generated by an information need. For instance, it seems those motivated by a hedonistic need could quickly reach a state of 'flow' [23], cognitive absorption and high levels of engagement while searching social media websites in their "everyday life". Also, individuals motivated by a hedonistic need described their experience rather than the information found. This observation suggests that user relevance judgments change when individuals are motivated by their desire to have fun or achieve a particular mood or state because the finding of information is often less important than the experience of finding.
- d) A particular set of casual-leisure information behaviours presented by Elswailer et al. [16] could be explain using the wanting-liking model of curiosity by Litman [1]. For example, 'needless browsing' and 'exploring for the experience' can be explained as examples of either I-type curiosity ("sub-rational liking") or irrational wanting because both absorption and exploratory behaviour are related to curiosity [21].

These are some of the discovered information gaps:

- a) Most information behaviour research has explored job-related scenarios, because the background of LIS. Thus, the reviewed literature shows the need to study leisure information behaviour scenarios which seem to be more frequent than ever before due to the ubiquity of mobile technologies. Although some information behaviour models and methodologies in work scenarios could be migrated to leisure information behaviour (e.g. task as unit of analysis [51]).
- b) Although emotional factors have been incorporated in some models of information behaviour, most ignored this important dimension of information seeking as shown by the presented sample of models

- c) Most information behaviour models, even those including emotional factors, failed to capture the role of positive emotions and hedonistic needs in the information seeking and use behaviour. For instance, Kuhlthaus “uncertainty principle” does not capture getting pleasure from searching, presumably because her studies were of school children doing an assignment, rather than searching for fun. In contrast, Litman & Jimerson [19], and Litman [1] portrayed of I-type curiosity concept show that individuals driven by this type of curiosity view situations characterized by uncertainty as positive, and opportunities to solve the incertitude as potentially enjoyable.
- d) Within Stebbins’s casual-leisure definition [54], Hartel’s hobby cooking information practices [57] and Elsweler’s casual-leisure information model [16], this PhD research finds its structure, scope and philosophical approach to connect LIS to leisure scenarios. However, neither Hartel nor Elsweler et al. studied leisure search in action, developed a methodology to study searching for fun, investigated how individuals curiosity (as an hedonistic need) could generate users engagement and exploration, and tested any of their practical recommendation regarding how to improve the search experience for casual leisure information seeking scenarios.

In conclusion, the presented sample of information behaviour has been presented as evidence of the adequacy of investigating information behaviour trigger by a hedonistic need in ‘everyday life’ leisure scenarios. Although the information behaviour models in the sample do not include directly the concept of curiosity, they show that information behaviour field have been biased by their historical background of LIS and the need of gaining awareness of other common information behaviour scenarios beyond job related context where experiencing some positive emotions and other human factors, such as curiosity, presumably might be more important than finding a particular document.

2.3 Information Retrieval Interaction Models (IIR)

Research in the area of IIR combines research from information retrieval (IR), information behaviour, and human computer interaction (HCI) to build a unique discipline that has the aim of helping people to explore, resolve, and manage their information problems via interactions with information systems.

Indeed, search can be considered as a communication problem between information systems and individuals. In an over-simplified IIR model, everything starts when the individual becomes aware of their information needs, then depending on the individuals searching skills and

particular context decides to iteratively interact with the information system until the needed information is found and used [5] (Figure 2.8).

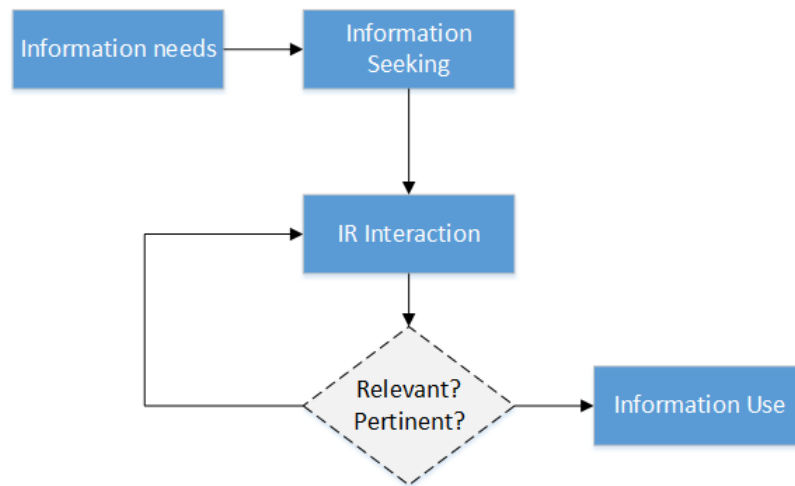


Figure 2.8: Standard information retrieval interaction in the information seeking process [5]

This section contains a representative sample of IIR information models. Although IIR seems a quite new field of research, Bourne and Hahn's [61] chronological account of on-line information services suggests that users and interactions were an important subject by the mid-1960s. They noticed that during this period of time several graphical and interactive techniques were introduced to guide users in their searching activities (e.g. the relevance feedback, on-line thesaurus, and save queries for later use).

By 1970, there were many important research projects related to IR. For example, Salton et al. [62] built the SMART system focused in core IR features such as indexing and retrieval, but it also made a great contribution to IIR because they identified that the user's original query was often wrong and some kind of users interaction with the IR-system was desirable. Salton et al. defined two types of interactive techniques: pre-search (e.g. construction of the query) and post-search (e.g. visualization of search results, relevance feedback).

Salton's work meant a division of IR research in two groups: firstly, the study of formal retrieval models and evaluation of systems based on these models (Cranfield Paradigm); secondly, information seeking and use in LIS. However, they also acknowledge the difficulty of evaluating IIR-systems and the limitations of the Cranfield evaluation paradigm (e.g. information needs are dynamic, relevance judgements depend on an individual's previous search expertise and knowledge, the success of information seeking situation is not based on the results of a single query).

Also in 1977, THOMAS system, created by Oddy, [63] was designed to help users with ill-defined information problems by constructing a conversation between the IR-system and users browsing interactive behaviour (i.e. assessing each document at a time and then giving a judgement upon various aspects of the document). Systems like THOMAS were ahead of their time and despite technical limitations followed the philosophy that search is interactive, it always a dialog.

In a similar way to the previous section (but using a more detail abstraction layer, the information searching lens), the IIR models to be presented provide evidence: 1) the motivation and adequacy of looking beyond the utilitarian value of information searching and the work task context as this research has done 2) the inherent nature of search as an iterative conversation between human and information system, often missed by standard IR.

2.3.1 Anomalous State of Knowledge (ASK)

In 1977, the idea of anomalous state of knowledge was published by Belkin [64] . Challenging the most important premise of IR-systems, he used the word anomaly to describe that the user's state of knowledge about a particular topic is somehow inadequate with respect to the individual's goal or task at hand (Figure 2.9). Later, Belkin et al. [65] called this a problematic situation suggesting that anomalous states may be of different kinds.

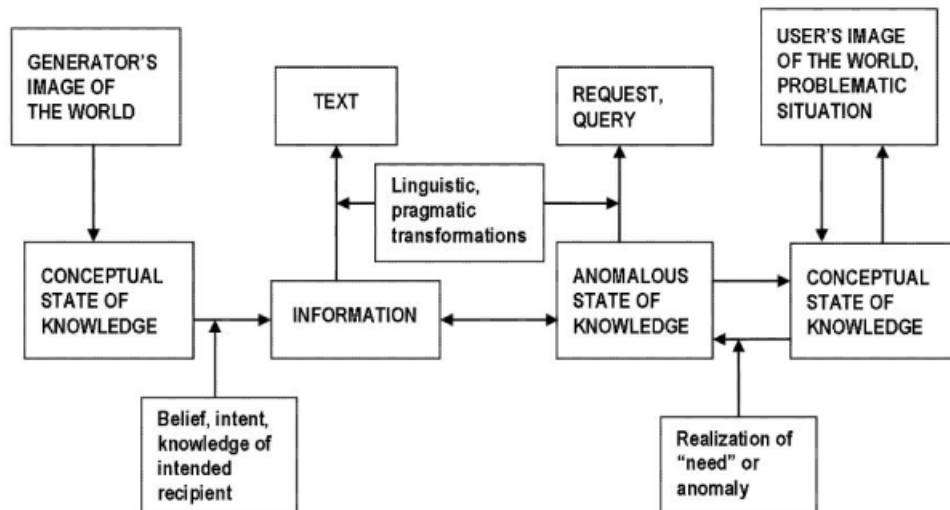


Figure 2.9: Anomalous State of Knowledge

The problematic situation described by Belkin had strong connections to Taylor's Question-

Negotiation model introduced in 1968. Taylor defined four levels of information needs (the first two levels represent Belkin's problematic situation). The four levels are:

- Q1** - the actual but unexpressed need of information (the visceral need)
- Q2** - the conscious, within-brain description of the need (the conscious need)
- Q3** - the formal statement of the need (the formalized need)
- Q4** - the question as presented to the information system

ASK followed the idea of cognitive viewpoint and the communication system model (i.e. a human generator that produces text and a user that reads the text), where any communication is an interaction between different states of knowledge. Therefore depending on the type of ASK, the IR-system must adopt appropriate interaction techniques to modify and resolve ASK. Belkin's episodic model of IIR has been partially implemented and tested in empirical experiments [66]. Belkin's extensive contribution in the field of information science has meant a turn from system orientation to users.

2.3.2 IS&R Nested Model

In 2005, Ingwersen & Järvelin's [60], integrated nested model from information seeking and IIR, was published in the book 'The Turn'. Following the cognitive approach, they were able to represent both general information seeking and concrete IIR scenarios by modeling the types of cognitive entities or actors (e.g. authors, indexing designers, interface designers, program committee members). IS&R nested model represents the task context as a set of four layers from the micro-level (IR) to macro-level (Culture). Figure 2.10 depicts IS&R model, with its five general components and eight relations.

According to Ingwersen & Järvelin every actor is influenced by two contextual forces: the social-cultural context and the information spaces (IT, Information Objects and Interface). Thus, each contextual universe plays a role in the other, and in order to study a particular component every other component must be taken into account because any component within the model could influence another. Also, there is a temporal nested factor namely historic context (actors previous experience) that affects systematic and socio-cultural contexts as well as actors' information seeking and use of information [43].

IS&R framework guides researchers of IIR, rather than information behaviour, to consider a wide range of factors that might influence real-life situations of information searching behaviour and emphasizes empirical research by using both qualitative and quantitative analysis methods.

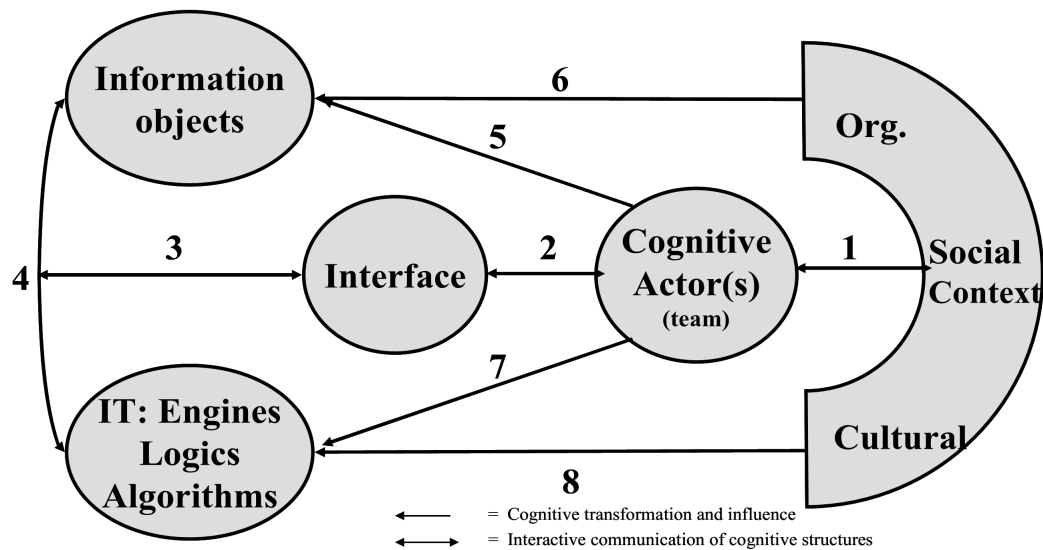


Figure 2.10: IS&R cognitive nested model

IS&R main weakness is brought by the assumption that all information behaviour could be delimited within the users' cognitive space. Although they mentioned that cognitive space for this model also covers emotional or affective factors.

2.3.3 Stratified Model of Information Retrieval Interaction

Since 1996, Saracevic introduced the Stratified Model of Information Retrieval Interaction by adopting stratification theory [67]. He defines an IIR as a conversation between user and computer by means of an interface with the main objective of affecting the cognitive user's state and then solving the problematic situation (Figure 2.11).

Both user and system are modelled by a group of layers that outline the dialogue process: the system by content, processing, and engineering layers; and the user by cognitive, affective, and situational layers. The main idea of these stratified models of interaction is that when one component is failing on either user or system side, the value of the entire IR-system is hampered. For example, the individuals assessment of the results might be seriously affected by their emotional state during the search task.

However, a possible weakness with this approach comes from ignoring temporal aspects by failing to mention the effect of time and iterative interactions.

Similarly to IS&R model, the stratified model assumes that the IIR process is "connected with cognition and then situational application" related with relevance [68]. A contextual component delimits the effect of social and cultural factors that may impact or trigger adaptations in various layers.

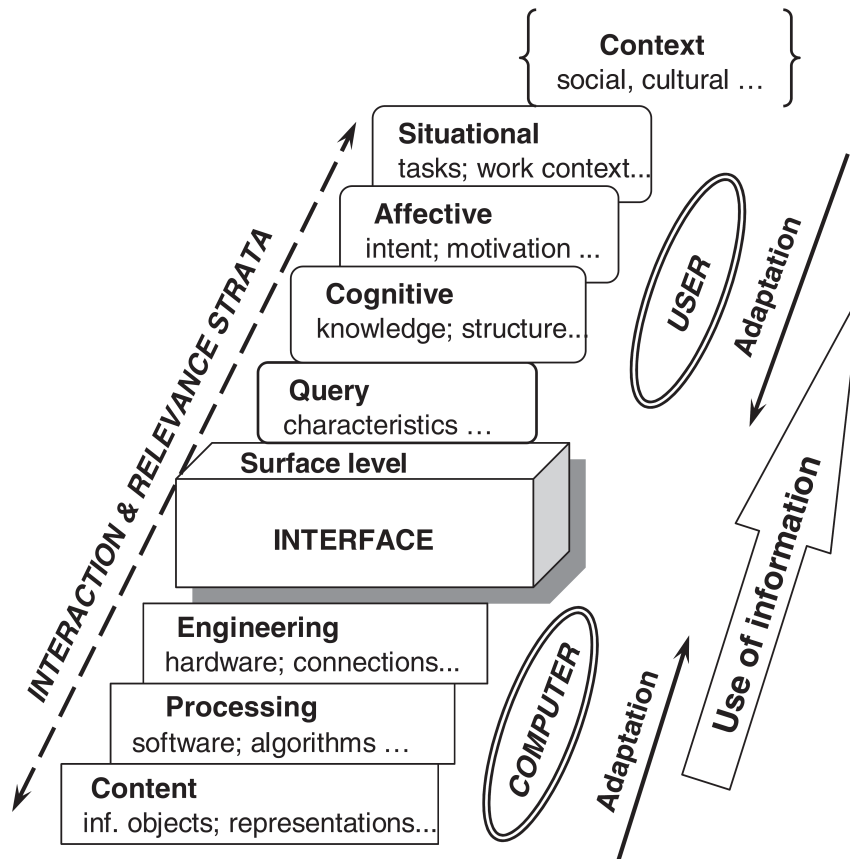


Figure 2.11: Stratified Model of Information Retrieval Interaction

Saracevic's model is often cited because as an IIR model it gives insight into other more complex types of relevance that should be modelled by an IIR-system (not just topical relevance) and the emphasis on context to understand users information seeking and searching behaviour. For instance, Saracevic's model and the concept of 'adaptation' show that the IR-system, the information, and the evaluation of the IR-system should be iteratively evaluated within a particular context [25]. Thus, the need to study IIR in other everyday life information scenarios, investigate users 'adaptations' and changes in their relevance judgments in order to perhaps create a better IIR-systems and understand the influence of other factors in the searching process.

2.3.4 Marchionini's Information-Seeking Process Model

In 1995, Marchionini's Information-Seeking Process (ISP) [69], based on a decision making and problem solving approach, describes a holistic model of information needs and the search process in electronic environments (Figure 2.12).

ISP model defines six variables within the search process: Information seeker, search task,

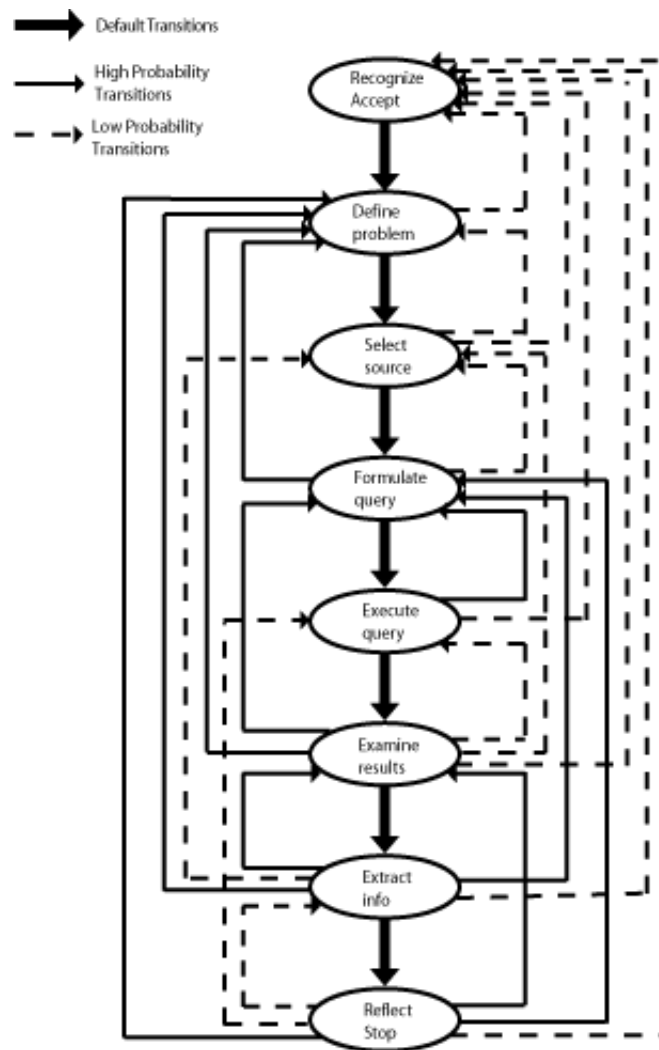


Figure 2.12: Information-Seeking Process Model

IR-System, search settings, and outcome of the search. The search process is broken into eight parallel sub-processes and their transition possibilities among those sub-processes (each transition has a probability).

The process is started by recognizing and accepting the need for new information. After the need has been accepted, it must be defined and understood. This critical step will influence sequentially the individual's choice of a system, the search query, and the execution of the search (e.g. typing the query, following links). However, a recognized problem might be ignored. Then the individual inspects the search results of the query execution, and makes relevant judgements over the extracted information from the results. Finally, the searcher must reflect upon the information and decide whether the information need has been satisfied or the searcher needs to be iterate.

Influence by physiological cognitive models [70], Marchionini suggests that interactions with

the IR-system are guided by a mental model (i.e. the individuals' internal representation of how a system works constituted on previous knowledge and abilities). ISP defines the search process beyond simply a findability information problem; search is associated with higher level cognitive processes such as learning and problem solving.

Marchionini's model reflects upon search as a holistic experience which help the IIR field to understand and interpret broader patterns of behaviour or interaction modes [14]. In 2006, Marchionini presented a taxonomy of search activities (Figure 2.13 [10]):

- Lookup
- Explore
- Investigate

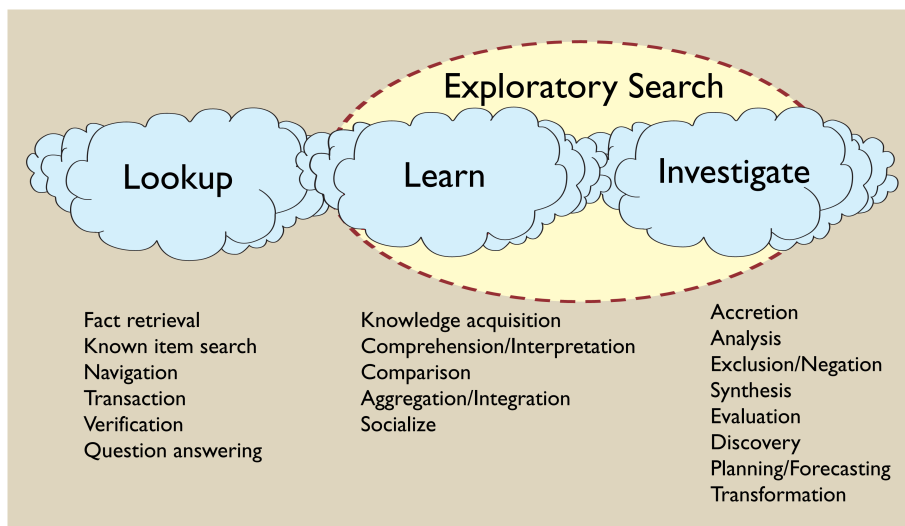


Figure 2.13: Exploratory Search (Lookup, Learn and Investigate)

This holistic view of the search process has benefited the IIR field by allowing practitioners and researchers to design and build IIR-systems beyond the query-response interaction paradigm [11]. However, Marchionini's model is limited by two assumptions: 1) the searcher's information needs remain unchanged throughout an iterative search process, 2) although Marchionini created his model based on the fields decision making and problem solving, ISP model represents searching and browsing as always rational making process where positive emotions and hedonistic desires are not captured.

2.3.5 Berrypicking Model

In 1989, Bates published the dynamic search model or berrypicking [71]. Drawing on her experience, she illustrated searching information as the task of picking berries in the forest (Figure 2.14). Bates' work is based on empirical studies of the search behavior of researchers who were experts in particular field (e.g., engineers, chemists, social scientists). Bates' model

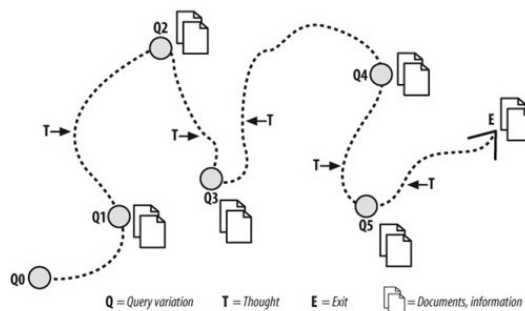


Figure 2.14: Bates' Dynamic Search Model

key contribution is that IIR is a nonlinear and evolving process in which information needs are neither static, nor represented by single query or satisfied by unique ideal group of documents.

Another significant contribution of Bates’s work is the definition of a system of search strategies and tactics by observing information seeking patterns that professionals employ in the diary work routine. She describes four levels of granularity within the search activity: [72]:

- **Move:** A specific step or action, such as keystroke, mouse click or finger touch in a mobile screen.
- **Tactic:** a set of moves such as TRACE (examining already found information for additional terms to continue the search), CHECK (monitoring).
- **Stratagem:** combination of moves and/or tactics (e.g. some examples include performing a citation search or following a footnote).
- **Strategy:** An overall plan for an entire search (e.g. pearl-growing, building blocks, successive fractions, interactive scanning, relevance feedback, trial and error, brief search).

Indeed, Bates’ work portrays that “search is not a quest for the perfect document but a conversation that helps us understand the right question to ask” [14]. Nonetheless, classical query-response interaction paradigm design seems to encourage the opposite behaviour (e.g. person writes the query, gets some ‘relevant’ results, and then moves on, after clicking in three or four results). The success on finding ‘the right question’ remains completely upon the searcher

experience and knowledge, the IIR-system just match the ‘perfect’ set of documents to a given query.

2.3.6 IIR Literature Discussion

The proposed models are examples historical development in IIR, and there is not a single model that represent all models [43]. The categories of IIR model by their approaches are:

- cognitive models (e.g. Ingwersen & Järvelin’s [60]).
- process model (e.g. Kuhlthau [3], Marchionini [69]).
- episodic models (e.g. Belkin [66])
- stratified models (e.g. Saracevic [68]).
- strategic models (e.g. Bates [71]).

The traditional IR model of interaction (i.e. the query-response paradigm) has been successful when looking for known information, where a single query represents the user’s information needs, and a perfect set of documents within the collection match the query. However, the reported IIR models show incompleteness of the query-response and Cranfield paradigm view of search because this interaction paradigm (designed for look-up search within a job-related context) oversimplifies the user as an only rational decision making organism and misses out the intrinsic iterative nature of searching. For instance, if a leisure search user has an undefined information need or a hedonistic motivation, why should the search experience be based only upon an interaction model that assumes the existence of this information need and the user’s capacity to communicate using a query? [12]; if the searcher is in a learning or exploratory scenario, how does the standard query-response interaction guide users beyond the first page of result, and encourage exploration of unknown information? How does query-response improve users ability to search? [11]

Even perfect ‘search engines’ are limited by the interaction paradigm and the assumptions made about the searcher [13]. The SERP interface and the query-response interaction paradigm main focus is satisfying users with information (e.g. quick search, some topically relevant rank results and then moving on) instead of enticing, engaging and guiding users to discover and explore the unknown. For example, sometimes user’s relevance judgments, and search behaviour are the product of irrational decision making process where emotion (rather than cognition) and other contextual factors (such as search scenario, e.g. people searching for fun instead of performing a work task) might have more importance [16].

The reviewed IIR models highlights the complexity of information searching and the need of designing IIR-systems to support and guide users, who are often untrained and diverse, thought their searching tasks [11]. Understanding searching as a holistic process [10] (i.e. searching and browsing affected both by emotion and cognition) is vital to build better IIR-systems and increase the ability of user to search, by designing the search experience to encourage learning, sense making, exploration, user engagement and a fun information experience.

Although neither of the aforementioned IIR models include the curiosity concept nor they were made using leisure search scenarios (such as searching information for fun in social network) nor this research is based in any of them, this section had the purpose explaining why this research uses the IIR, an not the IR, viewpoint to investigate search; and why query-response interaction paradigm fails in many common search scenarios. However, there are several connections between IIR, leisure and curiosity research. For instance, Belkin's ASK model [64] is very similar Loewestein description: "information-gap theory views curiosity as arising when attention becomes focused on a gap in one's knowledge. Such information gaps produce the feeling of deprivation labelled curiosity. The curious individual is motivated to obtain the missing information to reduce or eliminate the feeling of deprivation gap in our knowledge" [36]; Saracevic's affective layer in the Stratified Model of Information Retrieval Interaction presumably might include curiosity and then be used as framework to study possible adaptations of search behaviour within the context of leisure search; Marchionini's model defines search a decision making process, implying that emotions and other human factors like curiosity could influence search; Bates' dynamic model states the possibility that an IIR-system response to a query could change the user's mental model, modify the initial information need, and subsequent interaction. Therefore, this model implies that a IIR-system response can be design to influence the user's affective state and evoked curiosity which could be the driven force of exploratory behaviour and 'berry-picking' in a learning scenario.

To summarize, the provided sample of IIR models show that search is always iterative conversation between the IIR-system and the user. Thus, the standard query-response paradigm is an example of how this conversation could be design, but it is not the only possible way of designing the information interaction, or the right one for every scenario. According to many researchers in the IIR field, there are many reasons to moved beyond this interaction paradigm and design a better search experience for exploratory and leisure search scenarios [11, 14].

2.4 IIR and Evaluation

In the real world, users search typically encompasses post-query browsing, detail assessment of the results, evolving information needs, multi-session search tasks and contextual factors that might influence the search process. Notwithstanding classical IR evaluation, based on the Cranfield paradigm, excludes all of those aspects and according to Armstrong et al. (2009) ‘there is little evidence of improvement in ad hoc retrieval technology over the past decade’[73].

On the contrary, the user needs and individual’s interactions with the IR-system are preponderant in the user-centric approach. Therefore, this section main focus is to relate important concepts to IIR.

Evaluation by definition is the systematic process of acquiring and assessing information to provide a useful feedback about some object. So, evaluation needs an object to be evaluated and some goal that ought to be reached. The following section shows that IIR evaluation has selected IR-system as the object of evaluation whilst the goal is typically the quality of the result (e.g. rank list of results). But what about the interaction itself, how does one measure not just the quality of the result, but the quality of the interactive experience, the search experience [14].

In 2009, Kelly [28] provided an extensive view into the interactive information retrieval evaluation landscape from typical IR evaluation (i.e. Cranfield paradigm) to a user-centred approach both in the laboratory and naturalistic scenarios.

This section has been included to provide a historical background for the methodological decisions made during this PhD work such as the adoption of the User Experience (UX) viewpoint to evaluate information searching and IIR-system, the creation of ‘simulated’ leisure search scenarios, etc.

2.4.1 Relevance

One of most important aspects of evaluation of IR-systems is the construct of relevance. Even measures like precision and recall are based upon relevance judgements made by topic experts acting as judges and following a set of criteria as those used by the Cranfield paradigm [62]. System-centered evaluation measures recognize what Saracevic labels algorithmic and topical relevance (i.e. documents topical overlap with the users information need) [74].

However if relevance is assumed as complex, subjective, and multidimensional, then the use of baseline assessments based on topicality is not particularly useful [75]. Therefore, performance measures for IIR scenarios need to acknowledge cognitive, situational, and mo-

tivational relevance. For example, based on empirical user studies, Shamber & Bateman identified ten groups of user's criteria to judge search results: depth/scope/specificity; accuracy/validity/credibility; clarity/readability; currency; tangibility; quality of sources; accessibility; availability of information/sources of information; novelty; and affectiveness [76].

IIR evaluation metrics focused on performance of IIR-Systems has hindered innovative research because they do not entirely represent the user's search experience [26]. Despite the mentioned facts, IIR has kept and maintained the standard paradigm of the IIR evaluation (e.g. Ingwersen & Järvelin extended the Cranfield method into the interactive laboratory setting with humans in the loop) because IIR has not been able to create a consensuated new standard of measures. Nonetheless, there are some proposals to create measures that acknowledge the iterative essence of search, and that focus in assessing the experience, rather than the effectiveness of the IR-system [26] and learning outcomes [11].

In 2007, Saracevic reviewed the history and evolution of relevance research pointing out that relevance, although difficult to define, will always be important to the IIR field because it is in the very nature and core of searching [77, 74].

2.4.2 User-centred approach

Human centered evaluation for IR is not new, according Ingwersen & Järvelin it dates back to mid-1970s. However, the fact of having 'real' users in evaluation has made very difficult the standardization and reproducibility of IIR experiments because there are many practical limitations in the empirical user studies (e.g. number of test persons, the data collection protocol). The main goal of this evaluation approach is not the narrow focus of finding the perfect IIR-system, but rather modeling a holistic IIR-system that helps people to achieve a broader goal and affects the user's well-being.

Therefore, when bringing research from information behaviour is evident that IR-system is just a tool to accomplish a more general task (e.g. Byström & Hansen work task [78]). In 2003, Borlund [29], using the concept of situational relevance, presented her influential framework for IIR evaluation using simulated task scenarios where the goal of the IR-system is to support users in a 'real' task. Using Borlund, and other authors, Kelly (2009) mentioned several methodological factors to guide new researchers in the set up of IIR experiments (e.g. variables, factorial design, baselines, learning effect, timing).

2.4.3 Classical Measures of IIR

From Kelly's (2009) journal paper, there are three basic types of evaluation measures: contextual (e.g. age, sex, experience as the accumulation of knowledge of skill, personality type, task type, location, time), interaction (e.g. number of queries issued, number of documents viewed, query length, dwelling time), performance (e.g. number of relevant documents saved, mean average precision, discounted cumulated gain) and usability (e.g. easy to use, require effort, preference) [28] .

This classification of measures offers an interesting hint into one big historical limitation of IIR evaluation metrics. For instance, usability measures are typically seen in IIR experiments as indicator of users' attitudes toward the IR-System or the search results. Despite the fact that in the IIR experiment users' may be requested, before and after the search, about their effort, satisfaction, skill's confidence and topical knowledge, most IIR researchers have ignored users "affective reactions to the experience, specifically their emotional responses to various stages of the interaction" (e.g. user's engagement) [24].

Therefore, even though usability is one of the main concepts in HCI to evaluate efficiency and user's satisfaction, in IIR research it has been separated from performance and efficiency concepts because IIR is biased by core IR system-centric metrics. Thus, limiting the possible range of concepts that could be studied in empirical user studies of IIR and failing to represent the complexity of the search experience.

2.4.4 Search Experience

In traditional IIR literature, the expression 'search experience' relates to contextual measures that reflect the gain of knowledge or skill after the interaction with IR-system in a given task. But in 2002, Norman [79] introduced the concept of user experience (UX) existing in e-commerce to HCI. UX explores how users feel about using products (i.e. the affective aspects of product use).

One comprehensive definition of user experience is provided by Hassenzahl, he states: "An experience is an episode, a chunk of time that one went through with sights and sounds, feelings and thoughts, motives and actions; they are closely knitted together, stored in memory, labeled, relived and communicated to others" [33]. UX has dynamic essence because users' internal emotional and cognitive state, and contextual circumstances change during and after an interaction with a product [80].

In IIR, the product is the IR-system. Therefore, researchers and practitioners of IIR field have increasingly sought to bring UX to IIR, and then assess affective and unexplored aspects

of search (e.g. engagement, absorption, serendipity) in order to design commercial applications to help users in ‘real world’ problems [14, 24, 81]. The search experience evaluates the quality of the search dialog (i.e. information interaction) from the perspective of a user within a given context.

From 2008, O’Brien & Toms have developed the User Engagement Scale (UES) in exploratory search environments to evaluate the users’ perception of multi-dimensional aspects of the search experience because IIR can not keep imagine that a single Likert-scaled question about user satisfaction is a precise and clear measure of a multi-dimensional complex experience like searching [26]. Also they suggest that IIR should start looking “beyond standard metrics” and bring “measures of fulfilment, play, and engagement” [24].

2.4.5 User Engagement

O’Brien & Toms defined user engagement as “a quality of user experience that describes a positive human-computer interaction” going beyond user satisfaction [24]. User engagement while searching encompasses users’ perception of the search experience. UES models six factors of the interactive information experience [24]:

- Aesthetic Appeal (AE): The users perception of the visual appearance of a graphical user interface.
- Endurability (EN): Users overall evaluation of the experience, its perceived success and whether users would recommend the IR-system site to others.
- Felt Involvement (FI): Users feelings of being drawn in, interested, and having fun during the interaction.
- Focused Attention (FA): Users’ mental concentration, absorption, and temporal dissociation or telepresence (i.e. “being in flow”).
- Novelty (NO): Users level of interest in the task and curiosity evoked by the system and its contents.
- Perceived Usability (PUs): Users affective (e.g., frustration) and cognitive (e.g., effort) responses to the system.

2.4.6 Serendipity

This property “may indeed be the holy grail of the search experience” [14] and plenty of information seeking research has seen its value, but serendipity is a difficult property to study

formally from an engineering perspective because it is hard to define and capture. For this research serendipity goes beyond accidental discovery and explores “what it means to have a prepared mind and an infrastructure to support discovery” [82].

Common design search patterns such as “See Also” panels, recent searches and viewed items are used to support serendipitous exploration on commercial sites like Amazon and eBay. In the academic field, recent systems have been designed to promote serendipitous discoveries over web search [83]. These have used standard information retrieval techniques to model interest based on previous digital experience such as email and chat archives, they then use this information to highlight terms or pages [84, 85]. The aid provided by these systems enhances the mind capability to recall previous experience in the moment of searching and assumes that increasing the minds ability of recall might fill the knowledge gap to explore and discover.

In this proposal, serendipity would be assumed as not just a personal act, but rather a social process where unexpected insights from other peoples behaviours may help the individual make valuable decisions they would otherwise have been unable to make. Every invention or discovery has been understood in a social context even when chance has played a vital role. The opposite is also true and is well documented, there have been cases where due to some social context, individuals have lacked the capacity to understand the value of other people or their own discoveries even when they are not “accidental” [86]. To exemplify the importance of social context, think of a traveller who is looking for new places of interest, and then instead of following the plans of written tourist brochures or guides, chooses to unexpectedly change their destination based on key recommendations given by another traveller or local person [87]. Certainly, social media as a ubiquitous communication channel could improve this kind of interaction.

2.5 Context and IR

The idea of context may seem obvious when we reflect on the way we understand conversations or information, but it could be quite a challenge to apply it practically. Besides context as term has been misused. It has become for information retrieval and computer science domains almost a buzz word to pique the interest of an incautious reader.

Most information seeking models represent mental constructs, such as goals or plans, and then present the possible procedures the seeker may take to carry them out.

Context is a powerful variable to understand and modify human behaviour. The situated actions theory comprehends actions which come as a result of an immediate situation [88].

The situation is understood as the asset that makes knowledge possible and directs action. A particular example of how contextual factors influence information behaviour, based on traditional libraries and without an IR-system, is the design layout of them and the experience created based on the layout. For instance, the fact that people had to walk past other stacks of books in order to reach their desired volume made people discover serendipitous information to their current situation. Thus, spatial context was important in the design of the library user experience.

In most of the previous sections from information behaviour to IIR, context and situation has been commonly identified as an important construct to evaluate and understand information behaviour and use. For example, in IIR Kelly mentions two contextual measures such as user task, and topic as to study information seeking behaviours in a job-related situation [28].

There are several definitions of context as both internal and external construct. For Schilit [89] context-aware computing defines context as “where you are, who you are with, and what resources are nearby”. On the other hand, Morse [90] context is explained as “implicit situational information”; Schmidt [91] goes further and interprets context as “interrelated conditions in which something occurs” pointing out possible relations between context features. A comprehensive definition for context in computer science is provided by Dey [92] when he defined it as “any information that can be used to characterise the situation of an entity”.

Moreover, this definition of context seems too abstract. Therefore, we reviewed the literature on context frameworks and adopted one due to its extensiveness and logical division of the context. It consists of: Environment, Spatio-temporal, Personal, Social, and Task categories [6] (Figure 2.15).

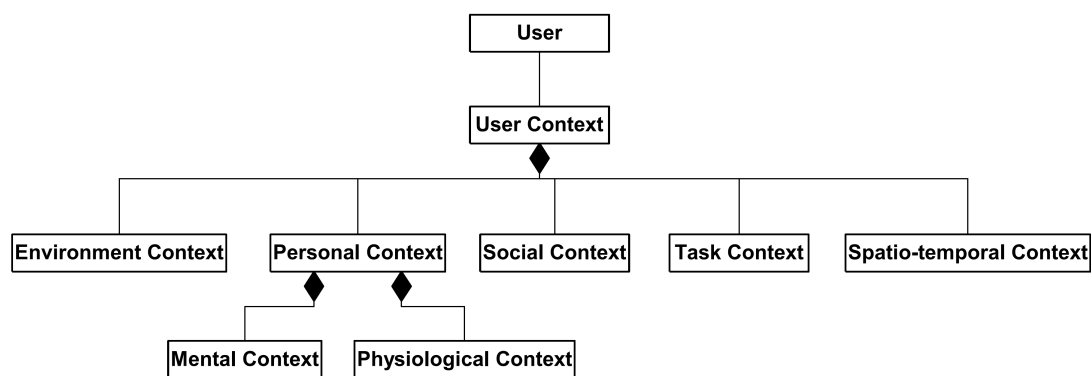


Figure 2.15: User Context Model [6]

In order to improve the search results, researchers in IR have used context to understand

and predict search moves. However, most of them have used the context concept within the IR system, but not when designing and assessing the search experience.

2.6 IIR and Search Interfaces

From Google to Netflix, all search interfaces share some design principles. In a very influential book, M. Hearst(2009) highlighted the importance of the search interface component in the IIR-system and summarised the development of this field from viewpoints of academic research and commercial systems [35]. Hearst's book was the first reference handbook of search interface design.

In 2011, M. Wilson grouped the main features of a search interface as input (e.g. search box), informational (e.g. number of results found, query reformulation, individual result, related searches), control (e.g. facet, pagination, Advance search features) or personalization (e.g. profile recommendation based on previous searches) [93]. His work offered an excellent general framework to study and evaluate search interface design. In 2013, Russel-Rose and Tate elaborated a comprehensive guide for the development of search interfaces from the User Experience (UX) perspective [14]. These publications are some of the most important publications in the field of search interfaces design.

The following section describes the modern components of search interfaces inspired by the cited literature. The section draws relations between the user interface components and the stages of the search process using a simplified version of the presented IIR models [5] (Figure 2.8). The section also uses Amazon.com and Booking.com search user interfaces to illustrate the state of the art of web search user interfaces.

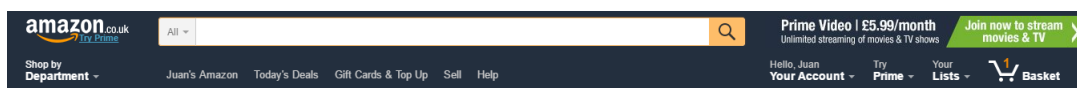


Figure 2.16: Amazon Search Box

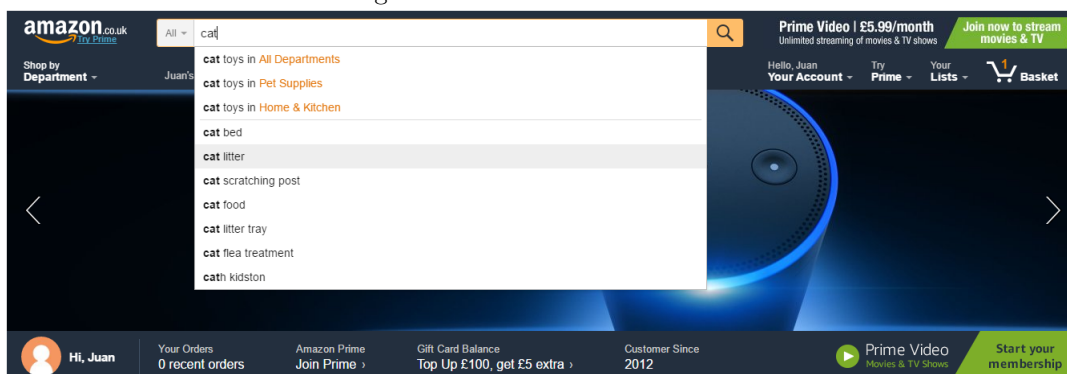


Figure 2.17: Amazon Search Suggestion for query formulation

2.6.1 Formulating the query

The search box is the standard way of communication between humans and the IIR system. There, in the search box, users are supposed to enter a sequence of terms that they perceive as a representation of their information needs (the query).

However, a search box needs to be more than an affordable space to write a query. When users attempt to formulate a query there are cognitive and emotional aspects. For example, most web search systems provide an interactive drop down list where the IIR system recommends queries or autocompletes partial words because an important principle in HCI is “recognition over recall”: the idea that people are better at recognizing things they have previously experienced than they are at recalling them from memory [14]. Then the user could select one of the suggestions or keep typing then entire query. Most search engines and search applications derive their recommendations from user’s individual search history and the aggregate search behaviour of many users.

Figures 2.16 and 2.17 show Amazon.com search box and Amazon’s autosuggest feature providing suggestions for the query “cat”. Figure 2.17 depicts that the first three recommendations provided a scope for the search using some product categories or taxonomies. This functionality is often use to support query disambiguation. These suggestions facilitate both known-item search and exploratory search behaviour within different search domains. Amazon’s autosuggest is a very powerful interactive feature because produces immediate feedback on keystroke and helps users in their dialog with the search application.

2.6.2 Examining Results

SERP design should accomplished two objectives: a) communicate the overall value and diversity of the result set b) convey the detail of each individual item [14].

Most IIR-systems present the search results as rank ordered list of documents based on a given relevance criteria. Although an order list of results does not represent any indication of relationship between result documents besides that the above document is estimated to be more relevant for a given query, the list of results offers the advantage of being “lean, ubiquitous and scalable; consistent, simple and intrinsic; and user and task inclusive”[94].

Based on Information Foraging Theory and the cognitive aspects of search, many practitioners follow the key principle of designing SERP to maximise information scent [49]. Back in 1998, Tombros & Sanderson showed that query-oriented summaries are strong indicators of relevance because they communicate how closely the query terms appear to one another on an individual document [95]. After their research a big field within the IR has been created to

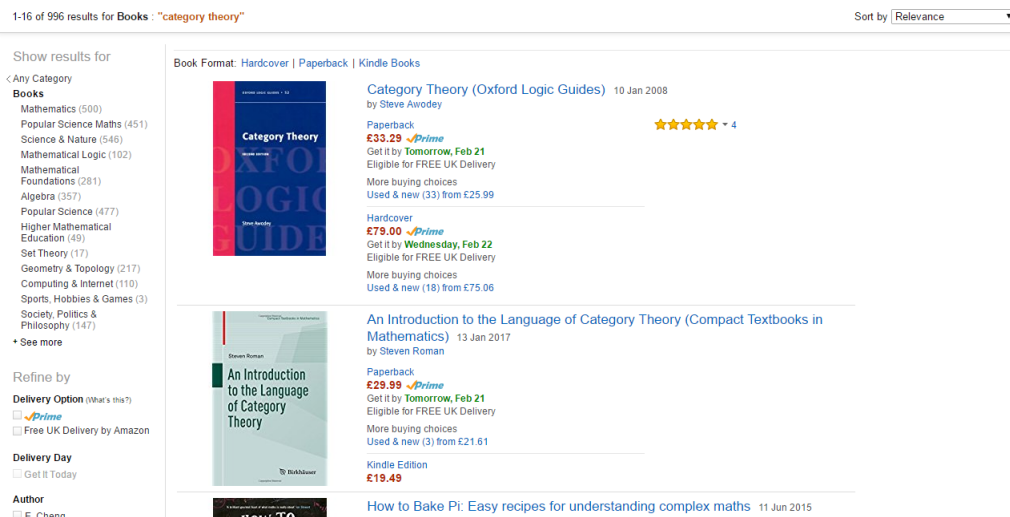


Figure 2.18: Amazon's SERP

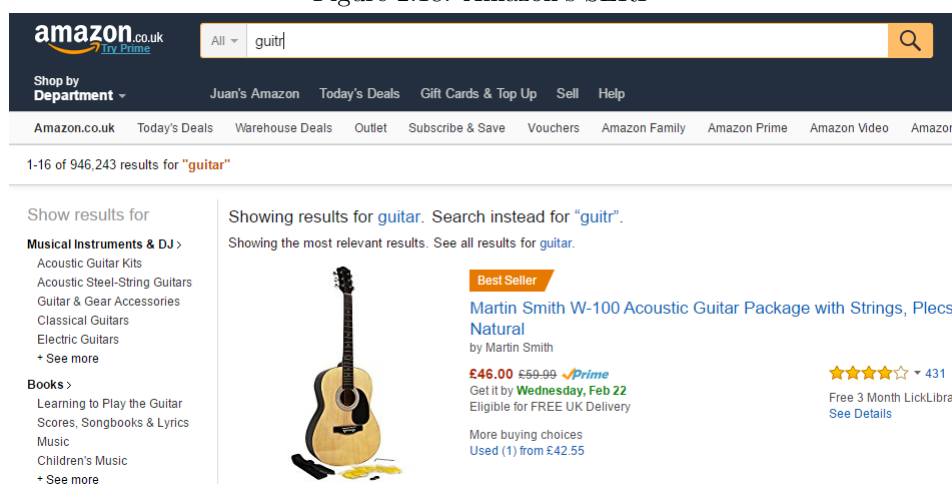


Figure 2.19: Amazon's misspelling auto-correction

support better text summarisation.

Drori & Alon (2003) demonstrated also that by presenting more detail information about individual items, users can browse individual results and verify the suitability without leaving the search interface [96]. Figure 2.18 shows how Amazon presents detail information of an item such as title, price, publication date, review ranked, available formats and thumbnail. Thus, users are help to make a better relevance assessment of the documents in the result page.

However, this idea must be balance with the important fact that if each item occupies more space, fewer results can be shown in the visible area without scrolling. As Russell-Rose & Tyler mentioned “there is no escaping between these opposing forces” [14].

Figure 2.19 highlights another important interactive scenario that all SERP must handle: searches with zero results. Figure 2.19 depicts a misspelled query term “guitr” and the Amazon’s ability to auto-correct the query and present the search results for “guitar”. SERP should not

just communicate the context of the search (e.g. zero results), but provide help to rectify the query [35].

Another important principle in the search interface design is “to maintain the context of the current search” [14]. This is usually done by displaying the number of matching results. Figures 2.18 and 2.19 illustrates how providing the total number of match documents in a SERP, it is implicitly telling the users the dimensions of the information spaces being search. The search performed in Figure 2.18 had less match documents than the “guitar”.

Depending on the query, SERP need to display heterogeneous collections of documents specially either to encourage exploratory search and discovery, or to help with query disambiguation by producing a more diverse set of results. One common approach to support is to return blended results or a composite search result.

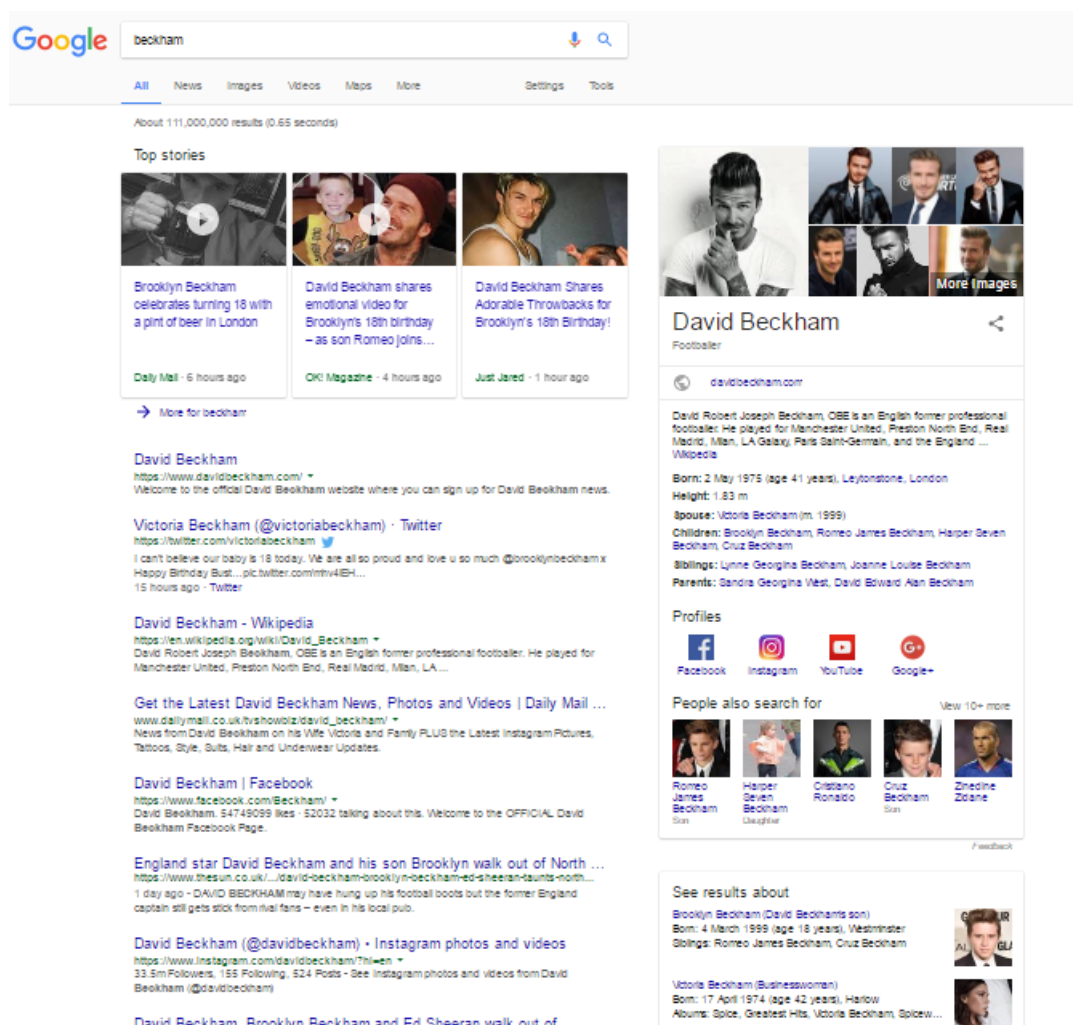


Figure 2.20: Amazon's SERP

Figure 2.20 depicts Google's search result page for the query "Beckham". The SERP provides recent news about the ex-footballer, his biography, pictures of him and more social media

content about him. There is also a section of related searches of other people who might have some relation with the retired player. These blended search layout encourage more than finding or locating a known document, they inspire searchers to explore what is new about the player, how other people are linked to the ex-footballer and how to find more information about the retired player.

Moreover, injecting diversity into the search result page might not be a sufficient condition to support exploratory search scenarios where searches cannot easily formulate effective queries, and decide when to stop searching. Stopping a learning task too quickly without in depth exploration might prevent searchers from knowing essential information about an specific subject [15] .

Most search interfaces lack the ability to communicate a) how much important information is available in the collection of documents b) how much of the information remains unexplored c) what different aspects (or dimensions) a topic has and d) the interrelation between those aspects (e.g. which one is more important). But there has been recently some research to overcome these limitations [11].

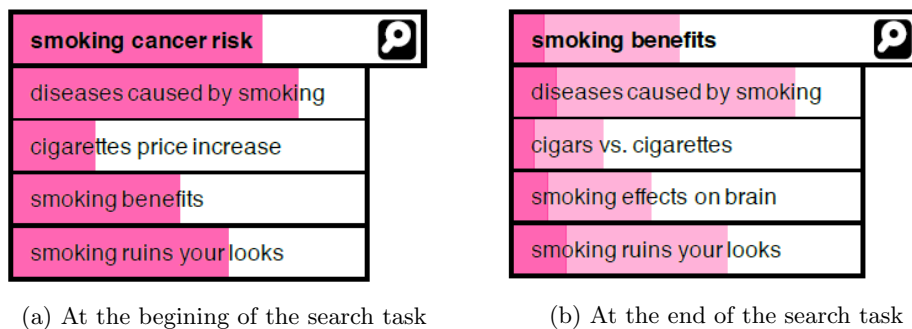


Figure 2.21: Scentbar suggestion query component. Screenshots are taken from [7]

For instance, Qvarfordt et al.(2013) designed a query preview interface component aiming to solve this problem [97]. Their SERP design show three kinds of information to represent the “current search context” in the form of a stacked bar chart: 1) the total number of newly retrieve documents, 2) the number of re-retrieve documents but not clicked ones, and 3) the number of re-retrieve documents that have already been looked by the searcher. Umemoto et al. (2016) created SERP and query suggestion component that visualize the guessed amount of important information that an user misses collecting from the search results after an individual query or during the span of a search session [7]. The idea is to communicate effectively the amount of information that users are missing in order that users can see their search progress visually. Figure 2.21 illustrates the searcher’s information gained and information missed during an exploratory search task with the goal of discovering the effect of smoking on human health

. However, none of these designs address the emotional aspect of searching, their main focus is the cognitive part of the search process.

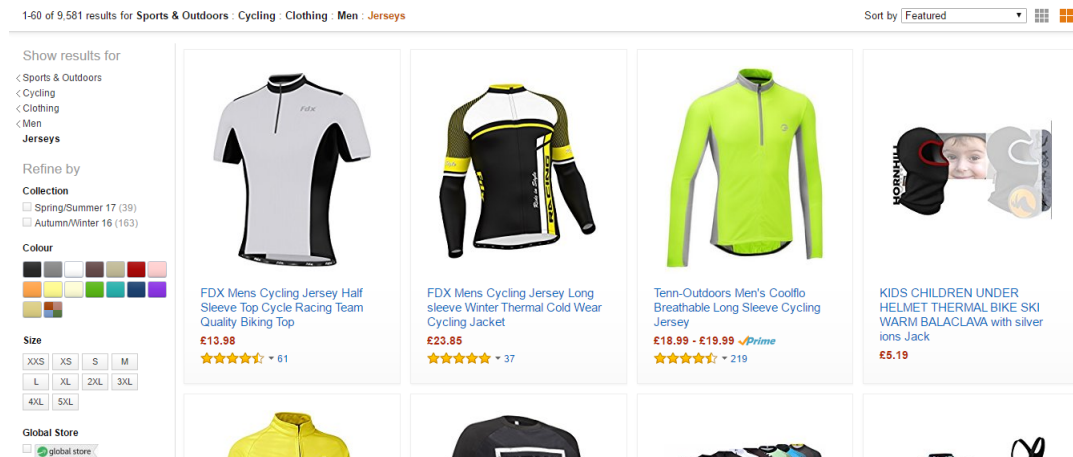


Figure 2.22: Amazon's facets for cycling clothes



Figure 2.23: Timeline Slider Facet in ACM's Digital Library for the query "IR"

2.6.3 Controls and Facets

In 2006, Hearst [98] presented evidence that faceted term search provides more effective information seeking support to users than typical query term search. The Flamenco project led by her was the most visible research project focusing on user interface aspects of faceted search.

Facets are independent dimensions by which a document could be classified. The whole idea behind facet search is to minimise the probability of reaching a zero result page by guiding users to more productive paths of information [14]. Facets are created by showing only available facet values given the current search context.

Figures 2.18 and 2.22 illustrate how Amazon uses search facets both to avoid dead ends

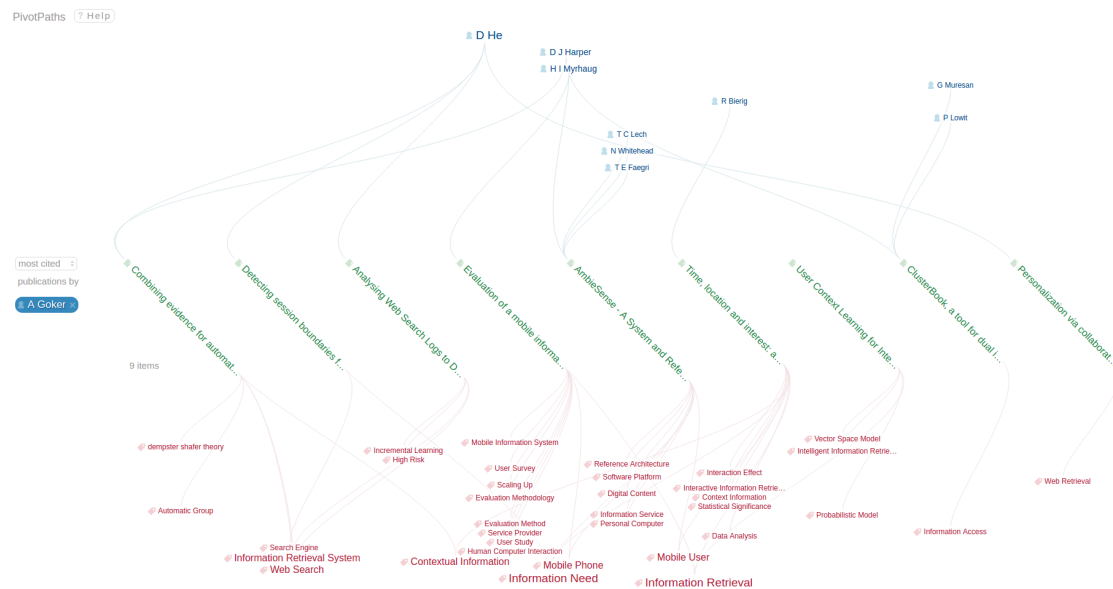


Figure 2.24: Pivotal Path Screenshot for the query “Ayse Goker” (see <http://mariandoerk.de/pivotpaths/demo/>)

in the search process and communicate contextual search information about the information space being explore. Figure 2.22 also shows that there are different types facets and ways to interact with them. For example, Amazon uses color buttons component visualization instead of normal category facet for searching cycling clothes. Figure 2.23 depicts the use of other interaction components such sliders to construct a timeline facet for searching publications in ACM’s digital library.

A number of applications are interested with understanding higher level patterns in a collection of documents. They used facets to provide an aggregate view of the collection and help in the discovery experience. This means a goal shifting for the search interface from locating a document to a more exploratory task. For example, Figure 2.23 provides an interesting insight into the development of the Information Retrieval (IR) research field. The aggregated timeline facet shows that the field had a lot of attention in the 1950, but in the 1970 the interest drop, and in the early 90s the field got traction again.

Aggregated data visualization could support better exploratory task scenarios by highlighting patterns in the information space. For example, PivotPaths [99] empowers an interesting combination of three ideas. Research publications are horizontally laid at the center of the screen, ordered by user defined relevance (either most cited, most recent or random), with authors given as a top layer and topics as a bottom layer. Links between articles and other objects are visible, with the location of authors and topics dictated by their relationships. As a result, proximity between authors (or topics) suggests contextual similarity. Figure 2.24 illustrates the

search results after “Ayse Goker” author query.

2.7 Casual-Leisure Search

From Taylor’s [12] “visceral needs” to “compromised need”, Belkin’s [64] “anomalous state of knowledge” and Bates’ [71] “cherrypicking or dynamic model”, IIR researchers have acknowledged the vital role of the user in the searching process and the intrinsic iterative nature of searching. They understand all IIR as a consequence of an information need that demands to be satisfied. Searching is always a rational iterative non-linear decision making process with the goal of maximize information utility and effectiveness (e.g. precision, recall, etc.). But is it really true for all iterative information retrieval behaviour? Given the reviewed literature from information behaviour to information seeking, this section finally introduces casual-leisure search, the motivation for investigating this field of IIR, the identified knowledge-gaps and connections with curiosity research.

2.7.1 Definition

Given the emergence of ubiquitous interconnected technologies and ‘Everyday Life’ [53] information seeking scenarios on them as never before, there has been an expanding research interest within the IIR field to study and understand Casual-Leisure Search or Fun Information Interaction (FII). Although fun information interaction could be a more general term because even in job-related or serious-leisure scenarios, search having fun could be important to enable creativity and well-being.

Fun information interaction could have the double meaning of: “searching for something that is fun” or “having fun while searching” [32]. Casual-leisure search might involve different scenarios and motivations. For example, searching funny videos to to pass time (kill time), browsing pictures of cats to relax, monitoring social media streams or news out of curiosity, on-line-window shopping to be entertained, reviewing your social network to escape from your current situation, etc.

Wilson & Elswailer [31] explain that casual-leisure search is often exploratory and motivated by hedonistic factors (e.g. boredom, curiosity, escapism, etc.). During casual-leisure search sessions, these hedonistic needs might be fulfilled interacting with the IR-System without eventually finding any useful information, and any other informational need could just be optional and transient. Therefore, even the search process could be “waste of time, pique interest, be fun” (Figure 2.25).

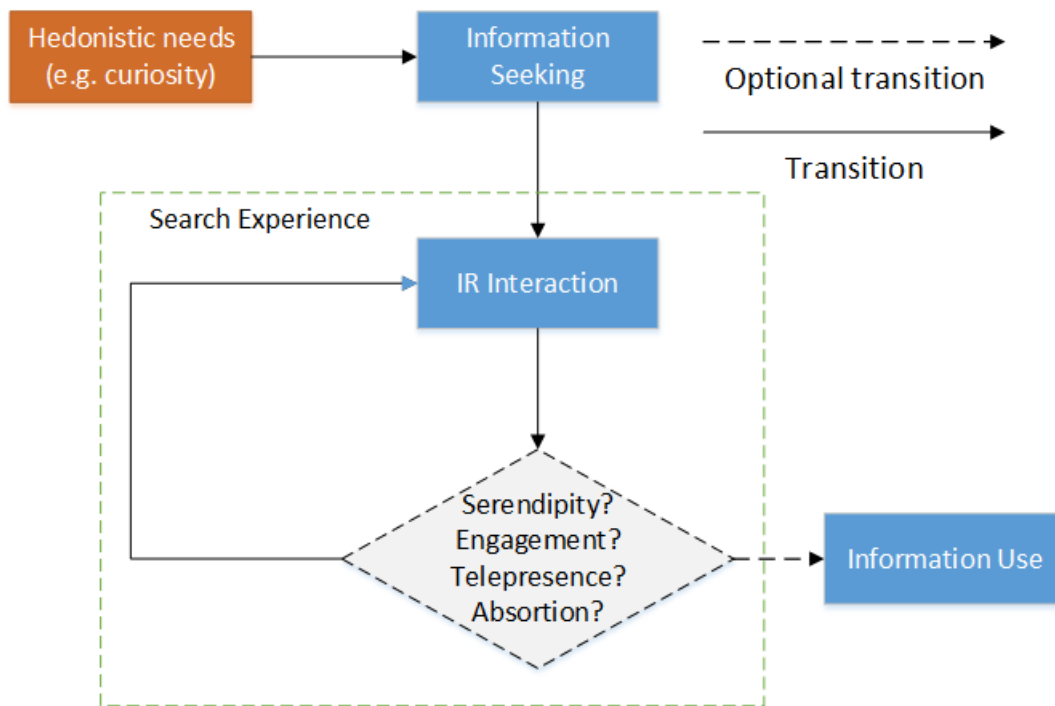


Figure 2.25: Casual-Leisure Search Model

2.7.2 Motivation

Although search is commonly understood just as a findability problem and restricted to a query-centric paradigm, IIR has highlighted the need to address exploratory search work-task scenarios [11, 14]. For example, Teevan et al. [13] showed from an observational diary study that even “perfect search engines” based on the keyword paradigm do not perform well when users know exactly what they are seeking. They discovered that people searching their own personal information displayed an orienteering behaviour using their local context instead of a fast keyword search approach because this search tactic did not require users to clearly specify their information need. Indeed, the user “is an integral part of the information system and not a stranger knocking at the door for directions” [12].

Social media (e.g Facebook, Twitter, Instagram, etc.), mobile devices and other pervasive technologies have made information accessible to people in leisure scenarios and revealed information searching behaviours motivated by hedonistic rather than informational need. In 2010, as already mentioned, Wilson and Elsweller [31] demonstrated how classical IIR models focused on job-related scenarios fall short in explaining common casual-leisure search behaviours because they were created in LIS. From empirical evidence, they suggest that 1) casual-leisure

search is an exploratory scenario and 2) IR-systems for leisure search scenarios should be designed to address hedonistic needs.

In recent years, casual-leisure search field has brought a number of workshops and events to foster research and unit efforts in the IR community (e.g. “Entertain Me” [100], “Searching4Fun” [101], “Evaluation of IR Dagstuhl” [32]). This IIR field has many challenges yet to be address because when searching for fun users moves and tactics, their motives, their standards for assessing success, and their interaction behaviour all contrast with typical search. However, after having so much access to ubiquitous technologies, these behaviours of ‘everyday life’ are probably representing a large amount of search queries in the web and social media [101].

2.7.3 Limitations and Challenges of IIR for Casual-Leisure Search

Casual-leisure Search changes most of the assumptions about the holistic experience of searching [32]. Therefore, this field needs new evaluation methods and measures. For instance, when considering casual-leisure search, stopping behaviour and dwelling time might be interpreted in a different way as in classical IIR (i.e. stopping could mean the user is not really having fun or entertained by the IR-system; a bad result could so discouraging that the person may stop the search session whilst discovering a new surprising information could trigger more exploration).

Following the existing literature in casual-leisure search, there are two possible ways to evaluate IIR-systems design to satisfied users’ hedonistic needs: engagement and state change (e.g. bored to not bored, stressed to relax, sad to happy) [32, 31, 102]. Nonetheless, it difficult to generate casual hedonistic needs on users in a laboratory setting because they are intrinsic to the user. How does one make an individual sad or bored in a way that they will naturally seek to entertain themselves?, how to reproduce casual-leisure behaviour in a controlled setting?, What are the best methodologies to assess casual-leisure search? [32]. This raises questions beyond technical developments and IIR. For example, is it ethical to make someone feel sad, and anxious for the sake of a casual-leisure research?.

Although there are many differences with job-related IIR scenarios, there is some common ground too. For instance, what is the purpose of IIR-system? In the past Artificial Intelligence (AI) claim that retrieving information really meant getting an answer to a question. But there are many search scenarios, even in the work task context, where users interacting with an IR system do not seek an answer, but to access and explore possible relevant sources of information as part of an information seeking task (e.g. A trivial question, has X anything to do with Y? Rather than to get an answer, users would want to see what has been said about X and Y) [15]. The purpose of search, as Bates’ dynamic model implies, is not to guide users to the ‘right’

answer, but the ‘right’ question. Thus, the purpose continues to be the same in both leisure and job related IIR-scenarios, what it changes from scenario to scenario are the aforementioned assumptions because ideally, at least in principle, an IIR-system should be both fun to use and an efficient tool to solve some kind of informational need.

Many IIR-systems are built, at least partly, with the idea of inducing the user to interact with the system without a predefined purpose, and to retain this user in interaction; some encourage serendipity and learning (e.g. students might be encouraged to go beyond one or two articles in Wikipedia); others support a certain kind of e-commerce behavior and advertising, etc. For example, Pearce et al. [27] have created an IIR-system to explore information, as opposed to search for it, by designing a playful system to support and motivate ongoing exploration.

Also another important common factor for IIR research, both in job and leisure scenarios, is that searching, browsing and exploring are not always conscious rational actions, and even when they are rational, our feelings and emotions could determine and influence the search experience (e.g. the role of anxiety on ISP [3]). For example, Wilson and Elswiler reported that some users search and browse for more time than they had planned because they were emotionally engaged in the activity [31]. Thus, modeling users’ affective state and supporting some emotions through the interaction might be a big innovation for IIR-system (e.g. e-learning information system that encourages students to explore with curiosity).

Users’ search behaviour in IIR-system is limited by the standard query-response interaction paradigm which main focus is satisfying users with information (e.g. quick search, some topically relevant ranked results and then moving on) instead of enticing, engaging and guiding users to discover the unknown. According to the reviewed and small literature of casual-leisure search, users who have a strong hedonistic motivation (such as curiosity) could engage interacting with an IIR for more time than what they have thought. During this leisure seeking scenario, users get pleasure from exploring and interacting with the IIR system. Therefore, drawing from Wilson and Elswiler’s reported studies [31], it might be possible that IIR researchers and practitioners develop better information systems and create a better search experience by mapping the psychological construct of curiosity [18] to the design of IIR-systems.

After reviewing the literature of curiosity theory, information behaviour models, and interactive information retrieval, it is clear that IIR-systems must go beyond helping people to find what they already know (query-response paradigm), they must stimulate or motivate exploration (e.g. using the dialog they establish with the user) and improve users’ ability to search.

This PhD project proposes the construct of curiosity as an hedonistic need that could be generated and stimulated within the holistic searching experience by a given IIR-system.

According to the presented psychology literature, curiosity is the most important driving force of cognitive development and exploratory behaviour. The fundamental goal of curiosity is not to acquire knowledge but to bring pleasure. Therefore, this research main hypothesis is that one possible way of achieving a more enticing and engaging search experience is to design IIR systems for curiosity [103]. The next chapter discusses the first experiment done to study casual-leisure search behaviours in action using social media content from Twitter.

Chapter 3

Exploring the effects of curiosity driven design in the search experience

This section contains the first experiment done in this research with two main novel contributions. First, the design of simulated “casual-leisure” search scenarios such as “wasting or killing time” or “exploring for the experience” in a experimental setting using the user’s context to evoke curiosity from social media information. Second, we found that for participants interacting with a novel curiosity driven application in a simulated casual-leisure scenario, their relevance judgments, cognitive absorption and the overall UX were strongly affected by the feeling of curiosity (e.g. users find irrelevant information but are happy with their search experience; users are deeply absorbed by the interaction, then they explored for more time). The experimental setting compares information searching behaviour between curiosity-driven and classical query-response search experience. The information searching behaviour was evaluated by using psychometric questionnaires and recording the user’s simulated search session length.

3.1 Social Media Search

Media content creation, edition and distribution has been transformed by Social Network Sites (SNS), smart mobile devices like mobile phones, tablets and other ubiquitous sensors. World events are documented and captured in real-time by individuals and organizations (e.g. sport events, earthquakes, hurricanes, stock markets, etc.) [104]. Media content is generated at astonishing rates by heterogeneous multilingual social mass of people who just want to communicate their thoughts, opinions and values.

Streams of media content are extremely important as sources of information. But most of their true value is still undiscovered due to complex technical data mining challenges such as

inconsistent quality (e.g. photomontage, stemming text with misspelling or social slang), lack of data (e.g. few occurrence of the same information) and the dynamic nature (e.g. newer information makes older sources irrelevant, millions of messages every minute force index scalability).

More than ever before media content posted online has spatio-temporal and social context metadata. For example every photo and message upload in the SNS has a creation date and it is created by someone connected with a social graph of friends or followers. In addition, plenty of media content is loaded with geo-information due to inclusion of GPS and other location systems. This research endeavours to use the current or desired “surrounding context” concept as a lens to filter information from social media and generate curiosity-driven search experience [6].

Recently there has been a lot of interest around using contextual features around a social media content (e.g location and time). For example, Whooly [105] is a web application that connects people with their hyper-local communities using event detection algorithms over Twitter data. Their search experience is not driven by contextual features following a similar layout of typical curated news media agregators.

Social media was chosen as an interesting information and communication environment for “casual” leisure search research because of the kinds of interactions this research was keen to observe and study. According to the literature social media information environment was relevant forum to investigate behaviours such as ‘Needless browsing’ or “Exploring for the experience” [16, 31].

Another key reason that influenced the use of Twitter social media data for this research, it was my employment with Ambiesense Ltd ¹ on a FP7 research project named WebinOS ² where our responsibility was to create a prototype application for smart-cities using the WebinOS platform and social media data with geo-location.

3.2 Motivation

In the current study, user’s engagement and behavioral measures were employed to examine and compare the social media search experience between two search application one based on query-response paradigm (i.e. Twitter search service) and the other based on curiosity arousal principles (i.e. Ambiecities). The purpose of the study was to gather empirical evidence of the usefulness of incorporating curiosity in the design of IIR systems, to learn about the effect

¹<http://www.ambiesense.com/>

²<https://developer.webinos.org/>

in user's engagement, possible effects in user's relevance assessment, how users exploration is increase when motivated by a hedonistic need, and then elaborate a set of design recommendations to enhance the search experience in leisure scenarios.

This experiment adopted user experience (UX) as a lens for exploring IIR in social media casual-leisure scenarios. The UES psychometric scale was used to capture both utilitarian and hedonic aspects of the search experience [26]. As already mentioned, UES is a multidimensional scale that measure pragmatic and hedonic aspects of the search experience. UES contains items related to novelty, aesthetically appeal, felt involvement, focused attention, durability and perceived usability. These items are important aspects of UX and they address both user's perceptions of the IIR-system and of themselves using the IIR-system [33].

Although there are several behavioural metrics in IIR (e.g. number of results retrieve, query changes, click behaviour, etc), this study employed time on the task (i.e. dwelling time). In classical IIR, more time on a task is often associated with poor performance (i.e. the IIR system is inefficient). However, O'Brien & Toms showed that this assumption could not be true when analysed from the lens of UX [26]. Based empirical evidence, they mentioned that time spent reading, browsing and in total may indicate a more engaging UX. Consequently, this research postulated that more total time spent will be associated with high scores on the self-reported user's engagement scale questionnaire, specially when user's interacted with curiosity-driven search application.

This study did not used other IIR behavioural metrics because it was difficult to record users' search behaviour when they were interacting with the Twitter commercial search service. Other reason for measuring only time on the task was that the main focus of this experiment is to compare user's engagement self-report (e.g. Focus Attention) and their subjective overall evaluation (e.g. relevance criteria for evaluating, comments, etc) with the perceived time and the recorded time on the task. Although future studies during this PhD will include other behavioural metrics.

To wrap up, the aim of this study was to research user's social media search behaviour in a casual-leisure scenario; and the effect of designing for curiosity in social media search applications. Combining self-report questionnaires (e.g. UES sub-scale) and behavioral data (e.g dwelling time), this research investigated how these different metrics related to each other in order to provide a 'true' picture of the holistic search experience.

3.3 Curiosity driven search application (Ambiecities)

The following section discusses the curiosity principles employed during the design and creation of Ambiecities³, a curiosity driven search application. This section also presents that some information systems haven been built for curiosity and user studies where this human drive of behaviour has been investigated.

In order to increase the “casual search behaviour” specially for the study, it was assumed that the search experience should be “session focused rather than result focused” as mentioned by Elswailer et al. [16]. To create such a UX, the design of Ambiecities had to make the search journey itself (i.e. the interaction and exploration) as important as the destination.

Ambiecities evoked curiosity by creating an interactive UX around the user’s context. The web application was built around the “transient information need” [16]: what is happening around a location according to Twitter?. The application uses Web Sockets, and Geo-location. There were two views: Map and List as shown in Figure 3.1.

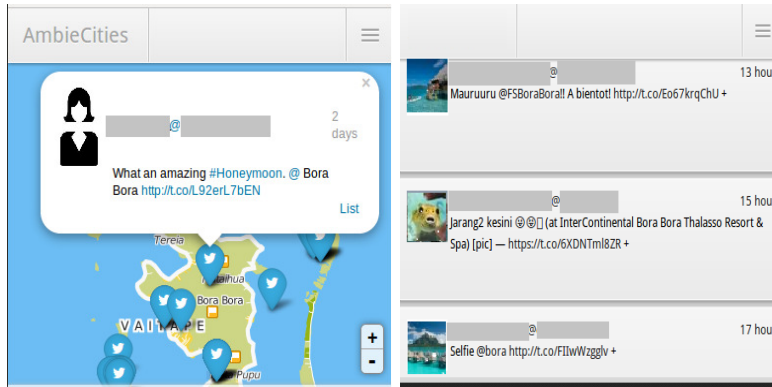


Figure 3.1: Ambiecities Mobile Map and List View

3.3.1 Curiosity in Information Systems and IIR

Although the curiosity construct is largely absent from the literature of IIR, LIS, HCI and CS, there are some few exceptions. For instance, serendipitous experiences have been associated with curiosity [17]. Hundt et al. (2012) built a IIR-system called the “*Bohemian Bookshelf*” using the visual aesthetics and animation factors to promote curiosity and provide a starting point for a more in depth interaction [106]. They designed the UX considering specific features like a “visually distinct interface, visual metaphors, the representation of unusual data facets and incorporation of visual cues to facilitate the interpretation of data“. Participants evaluating the Bohemian Bookshelf mentioned that the visually attractiveness motivated them to take a

³<http://www.ambiecities.com/main/>

closer look and catch their attention; then when they stumble upon some interesting result, they might want to read it and potentially lead to serendipitous discoveries.

Another example of curiosity and serendipity in LIS is provided by the Björneborn's (2008) ten dimensions of the physical library space that facilitate serendipity [107]. He mentioned that two key dimensions of the library space design were a curiosity-invoking display of library resources, and explorability, which is the encouragement of the individual to move freely around the library, to explore and follow their curiosity. Björneborn argues that in a library people must be curious about what is being displayed and become actively engaged rather than to be passive observers [107].

There are interesting studies relating curiosity to users' behaviour in e-commerce and marketing. Lee & Kozar (2009) reported a positive relationship between on-line store websites' ability to invoke curiosity on customers' affective appraisals and purchase intention [108]. They demonstrated that the ability of a website to invoke curiosity, interest and stimulate further viewing (i.e. 'mystery' in virtual spaces in their own words) could have a positive effect in UX. For instance, 'bonus', 'hot buy', 'learn more', 'compare price', or 'try it' cues may make on-line customers curious, and this may positively affect their appraisal and the purchase intention. Koo & Ju studied the effect of online atmospherics (i.e. the designing of buying environments to produce specific emotional effects in the buyer that enhance purchase probability) and perceptual curiosity on emotions and on-line shopping intentions. They recommended that the UX of e-commerce websites should concentrate on evoking emotions of their customers such as pleasure and arousal (i.e. 'the extent to which an individual who engages in an on-line shopping feels stimulated, active, or excited when he navigates web pages of the on-line store').

Koo and Choi (2010) investigated the moderating role of epistemic curiosity (EC) on link between knowledge search services user motives and intentions [109]. Respecting I-type EC (i.e. autotelic personality, people who enjoy acquiring and discovering new knowledge for the intrinsic joy), they concluded that people with high I-type EC are attracted to knowledge systems with features such as ease of use and enjoyment. Besides, they informed drawing from empirical data analysis that the selected knowledge search services companies failed to deliver satisfied level of services for interactivity and responsiveness enough to have an impact on user motives. Therefore, they recommended that system designers should improve the interactivity of the service by "enhancing two-way and concurrent communications among users and initiate services facilitating multi-directional discussions and timely exchanges of feedbacks among users". These literatures show a potential of path IIR research using curiosity constructs from Psychology, because searching like buying or purchasing is a decision making process that

could be influence by curiosity.

From the field of HCI, Tieben et al. (2011) developed a curiosity model around these principles to design playful interactive systems [110]. Instead of focusing on the content, they analyzed how the interaction by itself could affect the users curiosity and change people's behaviour. They developed five interaction scenarios by mapping the actions of passers-by into a sound output from speakers. They expected each of those scenarios to influence the way students walked through the corridor. To evoked this behaviour, they employed the five curiosity principles, with each scenario focusing on a specific principle.

After their 'everyday life' experiment in a school context (passers-by did not know that they were participants in a field study), they concluded that all principles have a compelling effect on people's behaviour, depending on the specific application and context. However, it seems that combining these principles for a specific target group, can lead to even more significant results. They specially argued that the combination of novelty, complexity and uncertainty creates a "promising, repeating stimulant for exploration"[110]. Tieben et al. [110] investigation has influenced and inspired the idea of generating exploration and enticing search experience by using curiosity principles, as outlined in chapter 2.

From CS, Wu et al. (2014) define a computational model of curiosity based on the psychological theory by Berlyne for virtual learning enviroments (VLE) [111]. Their computational model used three curiosity principles in virtual learning environments, i.e novelty, surprise, conflict and uncertainty. They found that curious learning companions can help the learners to explore more knowledge in the VLE and avoid confusion.

3.4 Designing for Curiosity

While curiosity may well be considered a personality trait, curiosity can be evoked through the following four principles, as described in the previous section: novelty, complexity, uncertainty and conflict [18].

The aim of this research was to explore if it was possible to evoke curiosity and exploratory behaviours through the interaction with an IIR-system for social media, using the four principles o collative variables as guides to design the search experience.

Therefore, Ambicities main purpose was to engage people during the search session, entice them to explore for a longer period and to experience 'flow' [23] (or at least cognitive absorption) rather than an IIR-system which goal is to retrieve topically relevant documents as quickly as possible, and then move on. Also, in contrast with a classical IIR-system design where high

topically relevant content is a key factor of the UX, Ambiecities did not filter, or rank Twitter content based on the quality of the Tweet, or other social signals. The content was not controlled because one of the objectives of the study was to investigate how the interaction alone could evoke curiosity, how the curiosity feeling produce by the interaction might trigger a longer search session, more cognitive absorption, and affect the overall evaluation of the UX. All these having in mind that the search scenario is about solving an user’s simulated hedonistic need (or experimental stimulated by evoking curiosity), and not a informational one.

Previous research found a strong relationship between context and the users’ motivation in casual leisure scenarios [16]. Spatio-temporal context features like Now, Recent, Near Me, Near a particular Location (e.g. clicking a location button, dragging the map to a particular location or typing the name of a particular location) were used as filters to retrieve information and enabled users to choose their “desired context” or “current context” [6]. Figure 3.2 shows the software architecture of Ambiecities and the interaction design using the “surrounding context” concept as lens to present and filter information.

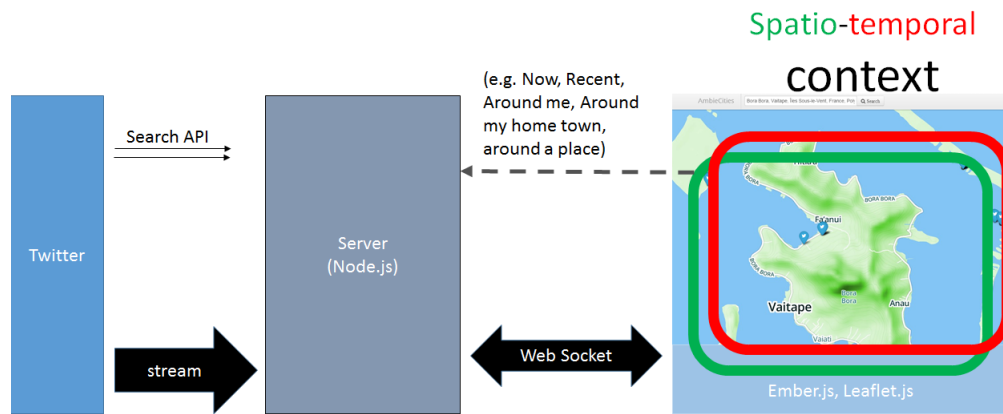


Figure 3.2: Ambiecities Software Architecture and Interaction Design.

Ambiecities was initially conceived as a simple extension of smart travel assistance application created for the WebinOS Project ⁴ using data with geo-location (e.g. images from Flickr ⁵, and microblogging from Twitter). The application was designed to create a full-duplex communication channel between users and the data coming from the third parties service using map visualization. In fact, the original name of the application was “Travel Companion” ⁶.

However, after showcasing the application in several forums (e.g. project meetings, other colleges in the university), it became evident to this research that users’ curiosity was aroused by the interaction defined in the application and the thought of being able to explore what it is happening in a remote location through the lenses of map based visualization. For example, in

⁴<https://developer.webinos.org/>

⁵<https://www.flickr.com/>

⁶<https://github.com/juanmillan85/app-travel-companion>

one presentation an audience interacting with Ambiecities kept exploring different locations of their interest (e.g. where they were from, where they lived, where they were going for holidays) and laughing at the content of some of the tweets. Then, this research decided to migrate this application outside the WebinOS platform to be used as a standard web application.

Instead of following the “cold” query-response IIR approach (i.e. submitting a query, then getting a static rank list of results, and moving on; also assuming that digital search is always a rational decision making process with only pragmatic goals), Ambiecities created a continuous and highly interactive search experience where a user would see markers pop up around the map, they were exploring at the time. The most recent tweet is represented with a red marker, and all other other tweets are represented with a blue marker. This interaction continues during all the search session. Therefore, the red marker changes location every time a new Tweet comes within the spatial-temporal context expressed by the map GUI.

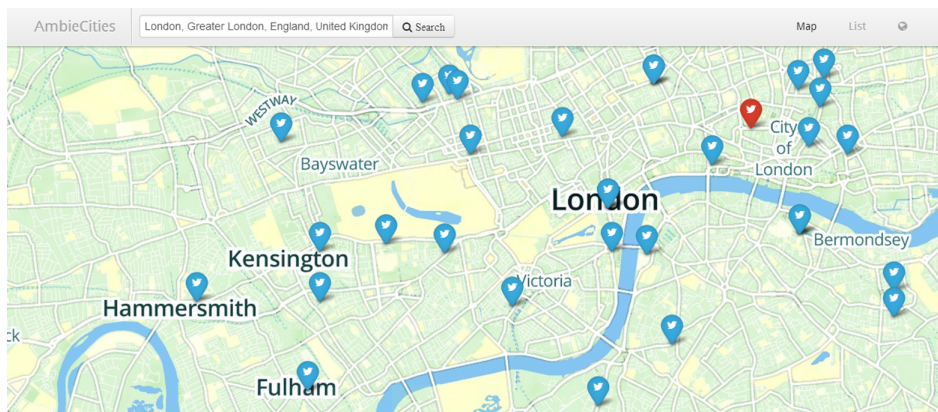


Figure 3.3: GUI Screenshots. London, UK Map and real-time Twitter markers. Each marker contains a Tweet with links and photos

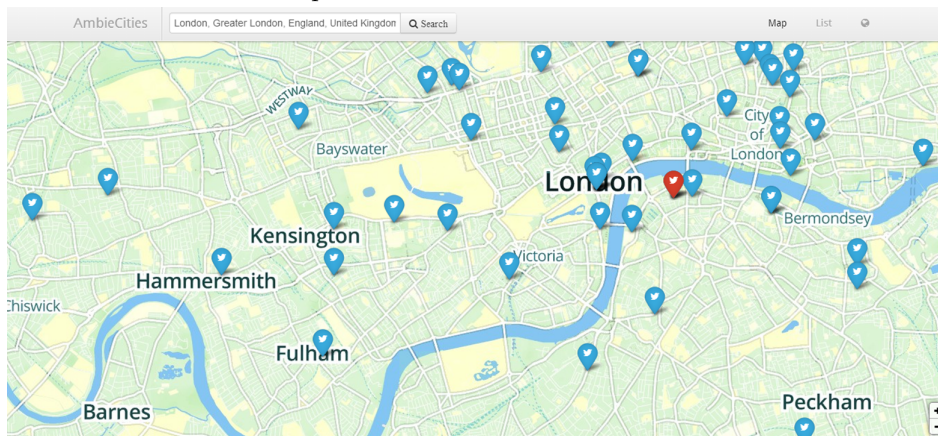


Figure 3.4: GUI Screenshots. London, UK Map after 3 seconds, and new Tweet appears. The red marker have changed location

Figures 3.3, 3.4, 3.5 and 3.6 illustrate the search experience by showing a set of screenshots of the user interface of Ambiecities. For example, Figures 3.3 and 3.4 show the large amount

of Tweets with geo-location produce during hypothetical searching and browsing session in London, UK. Similar to Tokyo, New York and others ‘big’ cities, London seems to have a lot Twitter users providing Tweets with geo-location. In contrast, Figure 3.6 shows that not many people in the city of Regensburg used Twitter with geo-location. Therefore, depending in the area that users would wanted to search (using the map or the top search box to retrieve a particular place, city or country), Ambiecities will retrieve a constant stream, or just a few Tweets.

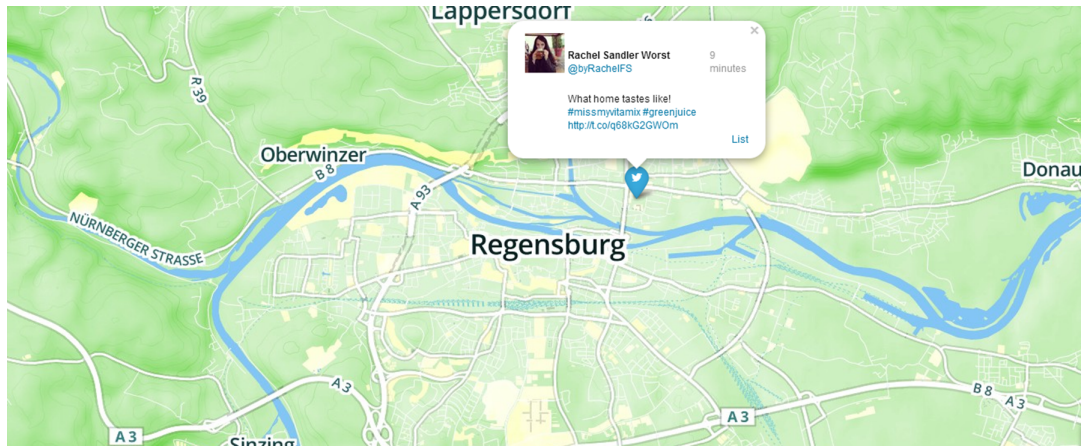


Figure 3.5: GUI Screenshots. Regensburg, Germany Map just one Tweet after 2 minutes



Figure 3.6: GUI Screenshots. Japan Map after 4 seconds

Also to encouraged the exploration of other places and to get people inspired, each time the application is refreshed, the user would start from a new location chosen at random from a set of pre-recommended locations (e.g. Paris, Caracas, Bora Bora, Cancun, etc).

Searching and browsing are decision-making process that can be highly emotional. The

process of choosing and the subsequent decision outcome may make us happy or sad, excited or bored, curious or feel regret. Following Kuhlthau's "uncertainty principle" [3], "cold" IIR-systems designed a positive search experience by providing users with high topically relevant and useful information. Thus, avoiding feelings of regret when the information seeking task is finished.

However, there are other powerful emotions and paths to evoked them that IIR-systems can used to create a positive UX and guide users exploration to the 'unknown' discoveries. For instance, drawing from two experimental studies presented by Van Dijk & Zeelenberg [112] where the role of curiosity and regret aversion were investigated in how decision-making occurs under uncertainty. They demonstrated that curiosity may overcome regret aversion in a decision making process (in other words, the feeling of curiosity could kill regret). Therefore, this research hypothesized that Ambicities interactivity and aesthetic appeal factors could be design to evoked curiosity and create a positive search experience, even without filtering and presenting high quality social media information.

Based on Berlyne's arousal mechanism to evoke curiosity [18], this research designed information interaction scenarios around novelty, uncertainty and their combination (Table 3.1). For example, the novelty and surprise factor is highlighted by the continuous stream of markers within the spatio-temporal context expressed using the map GUI. The continuous interactivity was created using HTML5 WebSockets technology to enable both push and pull interaction. Some possible curiosity driven scenarios enabled by Ambicities are: "I was looking the red tweet marker but when I was going to click, the red marker changed to blue and a new red marker appear. What is happening in this new location?, I am missing something", "I was reading a Tweet after clicking a marker, but in the background I could see that more Tweets were appearing a given location, what is happening there?", "every time I refresh I get a different place, so what is happening in xxx location?", "I dont need to think just pan around, and monitor what is happening right now in my hometown", "This new tweet was generated very closed where I was born/ I am living / I want to go for holidays/ the beach I want to go with by wife", etc.

On the other hand, uncertainty was created because all the markers were drew with the same colors and shape, they did not give a sign or cue of marker's content (i.e. the content of the tweet). This probably created curious questions in the mind of the users while searching. For instance, "What is this marker about?", "Is about the place?", "Is going to be interesting?", "Is it similar to the other near by Tweet?", "Who is sending this Tweet?", "Wow! a tweet inside of Buckingham Palace, is it the Queen?", "Where is the red marker? I need to browse more in

order to see the latest Tweet in the map”, etc.

Novelty: continuous search interaction using a map layout query the desire spatio-temporal context. the interaction is created by web sockets HTML5 technology	Uncertainty: there was not a sign or cue about the content of the marker. The tweet on the map was represented by a moving red marker
Scenarios: I was looking the red tweet marker but when I was going to click, the red marker changed to blue and a new red marker appear What is happening in this new location?, I am missing something I was reading a Tweet after clicking a marker, but in the background I could see that more Tweets were appearing a given location, what is happening there? Every time I refresh I get a different place, so what is happening in xxx location? I dont need to think just pan around, and look what is happening right now in my hometown This new tweet was generated very closed where I was born/ I am living / I want to go for holidays/ the beach I want to go with by wife	Scenarios: What is this marker about? Is about the place? Is going to be interesting? Is it similar to the other near by Tweet? Who is sending this Tweet? Wow! a tweet inside of Buckingham Palace, is it the Queen? Where is the red marker? I need to browse more in order to see the latest Tweet in the map

Table 3.1: Curiosity driven scenarios evoked by novelty and uncertainty

All of the curiosity principles are interrelated[18]. Complexity and novelty could be seen together. A stimulus with short-term novelty might have higher degree of temporal complexity than purely repetitive patterns [111]. For example, a the red marker moving around the map to unexpected places rather than in a pre-allocated position.

Conflict could also be observed in Ambiecities after a repeated exposure to the application. For example, exploring a similar or the same location in different moments (e.g. morning, night) and get different responses that are incompatible with each other. This principle could be used to model user’s curiosity as in virtual worlds where learning agent models conflict as the incompatibility between the human learners understanding and the expert knowledge embedded in the virtual world [111].

However, complexity and conflict curiosity arousal principles were not used explicitly in the design of Ambiecities because the search experience of application was created for leisure search scenarios where the exploration should be trigger by the “intrinsic enjoyment of exploring’, and not the desired of accumulate knowledge or information need. The goal of Ambiecities was to help users reaching an emotional-motivational state of curiosity in their exploration without any explicit and challenging “information gap”. In other words, this research intended to generate CFI (feeling-of-interest) curiosity instead of CFD (feeling-of-deprivation) [19].

3.5 Methodology

In order to study casual-leisure search behaviour and the effects of designing the searching experience based principles of curiosity theory, the experiment followed a user-centred approach both in the laboratory and naturalistic scenarios. [28]. As explained in chapter 2, the complexity of IIR necessities a mixed-methods perspective to UX measurement where behavioural IIR metrics provide the “what” and subjective, self-reported metrics the “why”. Although users could under- or over-report their search experience, the recorded behavioural data may provide evidence of the reliability of their subjective responses. For most part of the data analysis and data collection, this research followed the methodology describe by O’Brien and Toms to evaluate and measure different facets of UX [26].

3.5.1 Data Collection

The experiment was conducted both in the laboratory and in a more naturalistic scenario. The participants sample comprised 5 in the laboratory and 23 who joined the on-line study. In the laboratory, a Dell desktop computer with a webcam was used. Morae software captured the whole browsing session of the participants, and recorded their comments, heads and upper body reactions as they browsed and searched. Both in the laboratory and on-line setting, SurveyGizmo was employed to guide users through the experiment and store their responses to the questionnaires. The whole protocol was made available in English and Spanish.

There were several reasons to have both experiment settings (i.e. the laboratory and on-line). First, having a sample of the participants in a controlled usability laboratory helped this research to get more subjective feedback about the experiment and the UX because all the participants were interviewed after they finished. Second, having two setups for the experiment allow this research to compare users’ search behaviour in different circumstances and analyse the possible influence of an observer, and formal laboratory. Third, due to some logistical constraints this research was not able to get more participants even-though the voluntary invitations for the laboratory were shared by different channels of communication e.g. announcements via social media, email. The experiment was opened during 2 months in spring 2014 .

The SurveyGizmo guide or protocol was divided in the following sections:

1. An ethically consent form that described the experiment and the protocol of the study.
2. A demographic questionnaire that enquired participants about their gender, age, level of education, familiarity with social media sites, and some data about where do they look for information before choosing a leisure activity.

3. An introduction to a given interface and its functionality. Then the open-ended leisure search scenario with the link of the IIR-system. When they clicked a new browser tab was created. Below the link they were presented with long area section where they could write or paste the results and comments about the experience.
4. After they finished their browsing session in the other tab and they clicked a done button, they were requested to estimate how long they had searched or browsed during the simulated scenario.
5. An user's engagement perception questionnaire (UES; 26 items, E-shopping, O'Brien & Toms, 2010 [113]; 28 items, wikiSearch, O'Brien & Toms, 2013 [24]). The wording of the UES was replaced in order to be used for social media search. The items in the questionnaire were presented in a random order for each participant.
6. An questionnaire taken from Moshfeghi & Jose [102] was administered in order to compare task perception between the two applications. Four participants on the laboratory did not answered this section because this questionnaire was included after the second week of the study.
7. There were other subjective questions about the search experience (e.g. What aspects of the search experience did you like?, While using the App, did you enjoy your exploration?), the content and relevance assessment (e.g. While using the App, did you feel the information presented was relevant?, While performing any assigned task, did you find other information not related to the tasks interesting and valuable to you?); finally if they have learn something new or unexpected and suggestions for the UX.

In the laboratory, the subset of interviewed participants were asked to described their search experience, what they found the most and least engaging of their experience.

3.5.2 Procedure

Participants were recruited through various means such as email (e.g. the internal list of emails of City London University for students, and staff members), social media sites (e.g. Twitter, Facebook, LinkedIn). Volunteers for the laboratory responded by choosing a slot in a Doodle poll and sending an e-mail to the researchers if they had further questions about the location of HCI laboratory. The other participants just went directly to SurveyGizmo and accessed the experiment on-line. The duration of the experiment was expected to be between 45 minutes and 1 hour for each participant.

At the start of the session the participant was welcomed and led through an ethically-approved consent procedure. Following this, the experiment commenced, and they started to complete a demographic questionnaire. Next, they were introduced to either of the systems whether they had experienced Twitter or not. Then they were presented with the search task scenario based on Borlund’s simulated scenario [29] to situate participants in a potential casual-leisure scenario⁷.

After reading, participants were randomly assigned Twitter or Ambicities. this assignment was similar to Hu et al. [105]. During the simulated leisure search session, there was no minimum or maximum time for the task. Most users were recommended to use a browser such as Google Chrome, Internet Explorer or Mozilla Firefox. Figure 3.7 depicts both IIR-systems.

Twitter vs Ambicities

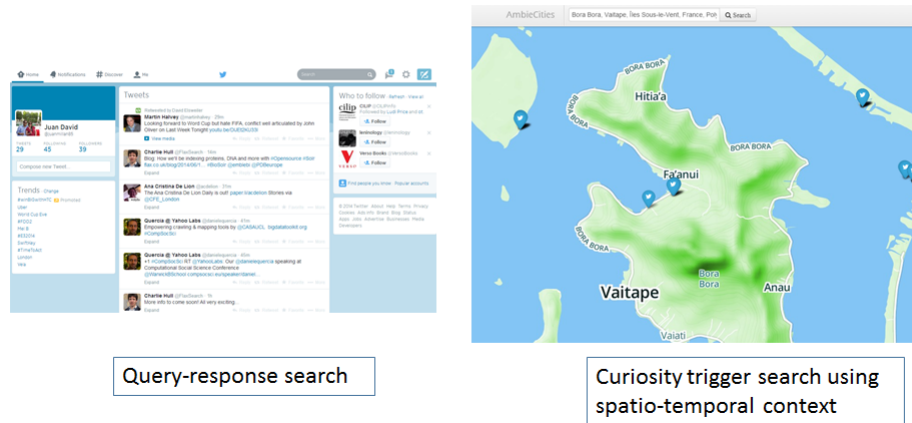


Figure 3.7: Ambicities vs. Twitter Experiment Setup

Twitter information seeking and search experience designed have been described by Russell-Rose, Lamantia & Makri [114]. They highlighted that the web interface was built around the four more important facets of Twitter.com: “the content and activities of people in the users personal network (Home); interactions with other users (Interactions); the users profile (Me); and a digest of content from all users in the twitter.com network (Discover)”. During the experiment, this research expected most users would choose to interact with the search service and ‘Browse categories’ to discover information and have fun similar to what it was reported by Elswailer et al. [16]. Twitter search service follows a query-response design and retrieves highly topically relevant content ranked using various social media variables such as likes, and shares; after performing a search, twitter allows the monitor of newer relevant tweet in the top of the web interface; users browse their results throughout an infinite rank list of results from top to bottom.

⁷In the following sub-section the search scenario will be discussed in length

On the other hand, Ambiecites' search experience design has a curiosity driven perspective. UX was created around the interactivity and the aesthetic appeal of the map GUI instead of retrieving high topically relevant content. Ambiecities did not apply any relevance function or quality filter on the incoming tweets. Although before starting the experiment, this research estimated that bad content could have a negative effect in the perception of the search experience and their engagement. For example, content such as irrelevant tweets (e.g. "Good morning London", "Buenas noches Cali"), racist, xenophobic, rude comments (e.g. "I hate 'XXXXX' people"), and even pornographic (e.g. tweets with links to other websites or pornographic pictures). This investigation wanted to observe how important the interaction design by itself and the motivation were to the overall evaluation of UX in a leisure scenario.

3.5.3 Search Scenario

After reviewing IIR literature and evaluation methodologies, this research choose to create simulated casual-leisure search scenario based on Borlund's simulated task scenarios. This scenario was not about just about finding or locating specific news, events or information, but about the "experience." The same scenario for all participants described a non-intentional or loosely-define interaction (i.e. no predesigned information need) with IIR-system. The scenario should encourage participants to do what comes natural in order to have fun and joy while interacting IIR-system.

The search scenario was also constructed drawing from Wilson & Elsweler comments regarding a casual-leisure search evaluation [31]. The mentioned that future research in casual-leisure search "must consider how [...] with high ecological validity [...] studies *could be created* where users are provided with hedonistically motivated tasks. Studies could be designed, for example, where users are told that there is a unforeseen delay and told they may use a computer while they wait. Then, when they appear to be bored, or after a reasonable amount of time, the faux-study continues". However, this research have never indented to make participants sad, bored or wait, the simulated scenario was built helping participant to imagine such a 'real' situation of being bored whilst waiting in particular place after an unforeseen event.

Also this research instead of provide users with 'real' hedonistically motivated task (having in mind the dubious ethical challenges involve in making a participant angry, bored, sad, etc), this research let participants interact with a social media IIR-system and observe how naturally they were able to enjoy their holistic search experience and how much they were distracted from their current simulated situation. Specially, when they were interacting with a hedonistically based an IIR-system (i.e. using the hedonic value of curiosity). This methodological idea

of moving the hedonistic perceived value from the search scenario description and protocol planning to the interaction with IIR-system built around hedonistic need (i.e. curiosity) has not been presented in IIR to study casual-leisure search in action so far according to the literature reviewed. In light of the aforementioned, this change in the methodology could be a novel contribution to IIR field if this research observes and records evidence that casual-leisure search behaviours such “Exploring for the experience”, “Needles Browsing” happen as a result of participants interacting Ambicities and their feeling of curiosity evoked by the application [16]. This finding would open another methodological way to study the search and browsing practice of what comes naturally in order to satisfy a hedonistic need (i.e. casual-leisure search) in action.

The defined simulated casual-leisure search scenario in order to generate an information environment for participants was:

‘You went to the best attraction in your city using public transport. Some good friends are visiting the city for their holidays and you are going to meet them at the station. You have been waiting them for almost half an hour. They send you a message saying they have got lost, but they will be coming as fast as they can. Feeling bored, you happen to see a big public touch screen with an app displaying what people are talking about. In order to change your mood, you would like to explore what is happening in your city or other parts of world while you wait for your friends’

To summaries, participants were asked to explore “what things are happening in their city or other parts of world” as proxy to seek enjoyment and entertainment while they wait for their friends in the simulated scenario.

This scenario could be related with an entertainment search intent described by Moshfeghi & Jose [102] as well. They categorized this search intent as entertainment by adjusting mood (ENM). Therefore, the main challenge for the IIR-system in this simulated search scenario was to provide an UX that encourage participants to have fun exploring social media information or at least to forget their boring simulated situation while distracted and engaged with the IIR-system.

As already mentioned, Twitter was chosen because previous researchers have highlighted microblogging as an important scenario in casual leisure search behaviour [16, 31].

3.5.4 UES for leisure

The wording of the UES items changed from WikiSearch to the App for both Ambicities and Twitter. No more adaptations were done for the purposes of maintaining the validity of the questionnaire to measure the UX.

Table 3.2 shows the total of 26 UES items grouped by six factors, Focused Attention (FA), Perceived Usability (PU), Aesthetics (AE), Endurability (EN), Novelty(NO), and Felt Involvement (FI). Participants rated all items using a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5) with a sixth option for not applicable.

Questions	UES Sub-scale
The time I spent searching just slipped away.	FA
I felt in control of the searching experience.	PU
I was really drawn into my searching tasks.	FI
I felt involved in the searching tasks	FI
Searching using the App was worthwhile	EN
I felt frustrated while using the App	PU
I was absorbed in my searching task.	FA
Using the App was mentally taxing.	PU
My search experience was fun	FI
I could not do some of the things I needed to do using the App.	PU
I consider my search experience a success.	EN
This search experience did not work out the way I had planned.	EN
I would recommend the App to my friends and family.	EN
I was so involved in my searching task that I lost track of time.	FA
I felt annoyed with using the App.	PU
This search experience was demanding.	PU
The App interface is attractive.	AE
My search experience was rewarding.	EN
I felt discouraged while using the App.	PU
The App interface is aesthetically appealing.	AE
I found the App confusing to use.	PU
I felt interested in my searching tasks.	NO
I continued to used the application out of curiosity.	NO
The screen layout of the App appealed to my visual senses.	AE
I blocked out things around me when I was using the App.	FA
I lost myself in this searching experience.	FA

Table 3.2: User Engagement Scale Questionnaire for the casual-leisure scenario

Drawing from observation made by O’Brien & Toms [113, 24], PU is an important variable in predicting searchers lasting impressions of the UX as worthwhile, rewarding, etc and mediated the relationships between AE, NO, FA, FI and EN. AE and NO “were predictors in the model, indicating whether [...] *participants* would chose to invest their attention and become involved in the searching encounter” [24]. FA was a very important specially in the e-shopping scenario where the ability of consumers to reach a Flow [23] state contributed to compelling on-line shopping experiences [113]. Overall, users level of FI foretold PU and EN (i.e. the overall UX

evaluation).

Another interesting observation from O'Brien & Toms [24], when testing the generalizability of UES for exploratory search, is that the average PUs rating was significantly higher than FAs. From their discussion, it seems that the nature of laboratory setting and the exploratory search task may have made a state of Flow and cognitive absorption difficult to achieve [23]. Therefore, there is a greater probability that in a naturalistic environment or scenario, participants FA and PU may complement each other because they might not be constrained by time or other externally motivated task factor (e.g. a formal HCI laboratory, the presence of a researcher, etc).

3.5.5 Participants

There were 28 participants, 19 males (67.9%) and 9 females (32.1%). Half of the sample was between the ages of 25 and 34 (n=14, 50.0%), approximately one third was between the ages of 18 and 24 (n=8, 28.6%). The remaining participants were between 35 and 54 (n=4, 14.3%), or over 55 (n=2, 7.1%).

The participants in the study were considerably well educated: 11 people (39.3%) held an undergraduate degree, 11 people pursued a masters degree (39.3%), 4 people had a trade or other technical school degree (14.3%), 1 held a medical degree (3.6%), and 1 person had a doctorate degree (3.6%). Sixteen reported that they spoke more than 2 languages.

Most reported daily use of social media information and familiarity with popular social networks sites through desktop computers or mobile phones. Participants were members of different social media sites: Facebook 89.3% (n=25), Twitter 82.1% (n=23), Linkedin 82.1% (n=23), Google+ 67.9% (n=19). The participants answered that when they choose a leisure activity 96% use Internet and 77% follow Word of Mouth.

3.5.6 Data Preparation

Data were analyzed using Excel statistical software package and Real Statistics Resource Pack macro (i.e. an Excel add-in which extends Excel's standard statistics capabilities by providing advanced worksheet functions and data analysis tools). The negatively worded UES items were reverse coded. Cronbach's alpha was employed to investigate the internal consistency of each sub-UES based on DeVellis' guidelines [115]: below 0.60: unacceptable; between 0.60 and 0.65: undesirable; between 0.65 and 0.70: minimally acceptable; between 0.70 and 0.80: respectable; between 0.80 and 0.90: very good; much above 0.90: attempt to reduce the number of items. The sample means and standard deviations for each sub-scale, and the correlations amongst

the sub-scales were examined.

Besides the UES questionnaire, in a following post-task questionnaire participants perception of their completed task in terms of the difficulty of the task, the familiarity of the participant with the task, the extent to which they found the task stressful, interesting and clear were measured by asking the following question ‘*The task we asked you to perform was [easy/stressful/interesting/clear/familiar] (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”)*’. The perception post-task questionnaire was taken from Moshfeghi & Jose [102], and the sample mean and standard deviation were investigated in relation to their previous work in entertainment based search and with previously mentioned UES.

The total time of the task was recorded in seconds. Thus, the values were converted into minutes and seconds because this scale was more appropriate to understand and analyse the data; besides the total time estimated after the search task was given by participants in minutes.

The complete protocol and questionnaires for the experiment were provided both in English and Spanish. Therefore, all the answers for open questions and comments were translated to English by a native Spanish speaker.

3.6 Results

User’s engagement scale scores were analyzed in comparison with participants’ total time on the task, total time on task estimation, perception of the task, relevance assessment and other subjective questions about the search experience.

3.6.1 Reliability analysis of UES

As shown in Table 3.3 and Table 3.5, the FA, PU, and NO were highly reliable for both Ambicities and Twitter participants. For the Ambicities group of participants, AE and FI sub-scales were in a respectable range of internal consistency, and EN in undesirable. If the item (“This search experience did not work out the way I had planned”) was eliminated from the EN sub-scale, the value of Cronbach’s alpha is improved (from 0.62 - 0.93) as shown in Table 3.4. Perhaps, one explanation could be that participants answered this question after experiencing Flow and cognitive absorption during the search scenario. Thus they might have expended more time, and explore more information than they had planned or imagined and then evaluated this UX factor as positive. Although DeVellis recommends that sub-scales over 0.90 should reduce the number of items, eliminating more of them in the FA, and EN sub-scales

would not have reduced the Cronbach's alpha significantly.

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	5	3.30	1.08	0.92
PU	8	3.65	0.75	0.85
FI	3	3.40	0.96	0.71
EN	5	3.14	1.18	0.62
AE	3	3.69	0.70	0.78
NO	2	3.79	0.96	0.75
Total	26	3.50	1.08	0.77

Table 3.3: Ambicities UES

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	5	3.30	1.08	0.92
PU	8	3.65	1.15	0.85
FI	3	3.40	0.96	0.71
EN	4	3.21	1.17	0.93
AE	3	3.69	0.84	0.78
NO	2	3.79	0.96	0.75
Total	25	3.51	1.07	0.82

Table 3.4: Ambicities UES with 25 questions

Participants using Twitter had a respectable internal consistency answering EN sub-scale, unacceptable for FI sub-scale, and undesirable for AE sub-scale. Table 3.5 presents the detail values of Cronbach's alpha for the Twitter group of participants. If the item ("I felt involved in the searching tasks") for FI sub-scale and the item ("The screen layout of the App appealed to my visual senses") for AE sub-scale were eliminated, the value of Cronbach's alpha grew for FI (from 0.56 - 0.82) and AE (from 0.61 - 0.78). The eliminated items were chosen after analysing the internal correlation between AE and FI sub-scales. Table 3.6 exposed the values of each sub-scale when the number of items is reduced to 24 for Twitter group of participants .

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	5	2.87	1.141	0.84
PU	8	3.84	0.98	0.87
FI	3	3.43	0.91	0.56
EN	5	3.46	1.18	0.74
AE	3	3.52	1.04	0.61
NO	2	3.29	0.90	0.87
Total	26	3.40	1.03	0.75

Table 3.5: Twitter UES

Means were determined by summing participants ratings of items within each sub-scale and dividing by the total number of items for that sub-scale; these individual scores were then computed to obtain means and standard deviations for each sub-scale. The UES sub-scales evaluated most highly on average by participants were PUs (Ambicities, mean 3.64;

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	5	2.87	1.14	0.84
PU	8	3.84	0.98	0.87
FI	2	3.54	1.00	0.82
EN	5	3.46	1.18	0.74
AE	2	3.54	1.07	0.86
NO	2	3.29	0.90	0.87
Total	24	3.42	1.04	0.83

Table 3.6: Twitter UES with 24 questions

Twitter mean 3.84) and AEs (Ambicities, 3.69; Twitter, 3.52). On five-point Likert scale FI (Ambicities, 3.40; Twitter, 3.43), NO (Ambicities, 3.79; Twitter, 3.29), EN (Ambicities, 3.14; Twitter, 3.46) were rated average between 3.30 - 3.42. The lowest sub-scale value on average was FA 3.09 (Ambicities, 3.33; Twitter, 2.87). FA and NO sub-scales seemed to be the most representative difference when evaluating the UX between the two groups of participants.

The ShapiroWilk test of normality refuted the assumption of normally distributed data for FA sub-scale. This research used nonparametric test of Mann Whitney U to compare groups according to the rating of FA sub-scale. Responses to FA sub-scale showed that median latencies in the group of Ambicities and Twitter were 3.3 and 2.87; the distributions in the two groups differed significantly (MannWhitney U = 2001, $n_1 = n_2 = 14 \times 5 = 70$, $P < 0.027$ one-tailed, $r=.16$). Therefore, there is non-parametric 0.973 confidence interval that a randomly chosen subject from the Ambicities group has a FA higher score than a randomly chosen subject from the Twitter group. This no-parametric test rejects one of the formulated null hypothesis at the beginning of this study regarding the effect of curiosity driven search (Ambicities) in FA during casual-leisure search on social media. However in order to provide a more evidence, this research compares this findings against the recorded behavioural metrics and subjective responses.

The assumption of normality on the NO sub-scale was tested using ShapiroWilk as well. But again the distribution of rates rejected the null hypothesis of normally distributed data. Thus, this research used the nonparametric test of Mann Whitney to compare whether or not the two sample distributions of NO ratings differ significantly. The test of NO sub-scales demonstrated that median latencies in the group of Ambicities and Twitter were 4 and 3.5; the distribution in both groups was statistically significant (MannWhitney U = 276, $n_1 = n_2 = 14 \times 2 = 28$, $P < 0.044$ two-tailed, $r=.27$) Therefore, there is non-parametric 0.95 confidence interval that a randomly chosen subject from the Ambicities group has a NO higher score than a randomly chosen subject from the Twitter group. This second no-parametric test rejects another formulated null hypothesis at the beginning of this study regarding the effect of curiosity

driven search (Ambiecities) in NO during casual-leisure search on social media.

3.6.2 Relationship between the UES sub-scales

Pearson's correlation amongst the UES sub-scales were analysed for both groups of participants Ambiecities and Twitter. Table 3.7 shows all the results. High correlations were observed between most UES sub-scales.

Measure	FA	PU	FI	EN	AE	NO
Ambiecities						
FA	1					
PU	0.17	1				
FI	0.64	-0.01	1			
EN	0.81	0.32	0.79	1		
AE	0.51	0.01	0.18	0.25	1	
NO	0.66	0.48	0.52	0.63	0.44	1
Twitter						
FA	1					
PU	-0.04	1				
FI	0.64	0.28	1			
EN	0.64	0.31	0.51	1		
AE	0.20	-0.26	0.03	0.26	1	
NO	0.61	0.22	0.77	0.72	0.26	1

Table 3.7: UES Factor's Correlation

For the Ambiecites group, correlations above 0.5, as observed between FA and EN, FA and FI, FA and NO, FA and AE, EN and FI, NO and FI and NO and EN, indicated that some of the items within these sub-scales may load on more than one factor. However, because the sample size limited the possibility of using statistical techniques such as factor analysis to look at the dimensions of the combined scales, they were investigated as separate sub-scales. There was negative correlation between PU and FI ($r=0.01$). This demonstrated that these sub-scales represented distinct dimensions of UX, but their relationship could require further exploration because according previous studies to O'Brien and Toms [113] mentioned that PU mediated and predicted FI. For example, in the e-shopping scenario O'Brien and Toms mentioned "if the user has experienced Felt Involvement, it is because the usability of the system did not interrupt or prevent them from enjoying themselves; in this situation, judgments of the Perceived Usability will be influenced by the level of involvement achieved". Therefore, these correlations are evidence that the Ambiecites group evaluated their UX beyond utilitarian and pragmatic dimension, and other UES sub-scales such as FA, and NO became better predictors FI, and EN than PU [113].

On the other hand, Twitter participants UES sub-scales indicated high correlation between

FA and FI, FA and EN, FA and NO, FI and EN, FI and NO, NO and EN. There was also a negative correlation among AE and PU ($r=0.26$). This established that these sub-scales represented different dimensions of UX, but their relationship could require further exploration because most of previous studies have shown a different association. This results could be related with the webcast study by O'Brien and Toms [116] where negative correlation were obtained and AE was not correlated with FA or PU. Contradicting the shopping scenario which found a predictive relationship between these factors [113]. Perhaps, participants familiarity with the system may have made the interaction efficient and as planned, but not graphically interesting, attractive or appealing.

3.6.3 Total scores of UES

Tables 3.8 and 3.9 provides details regarding the mean and standard deviation of each UES item for both groups of participants.

Questions	UES Sub-scale	$\bar{x}$	σ
The time I spent searching just slipped away.	FA	3.79	1.19
I felt in control of the searching experience.	PU	3.64	0.84
I was really drawn into my searching tasks.	FI	3.36	1.00
I felt involved in the searching tasks	FI	3.5	1.012
Searching using the App was worthwhile	EN	3.21	1.19
I felt frustrated while using the App	PU	3.57	1.02
I was absorbed in my searching task.	FA	3.14	1.17
Using the App was mentally taxing.	PU	3.79	0.97
My search experience was fun	FI	3.36	0.92
I could not do some of the things I needed to do using the App.	PU	3.07	1.21
I consider my search experience a success.	EN	3.28	1.26
This search experience did not work out the way I had planned.	EN	3.14	1.23
I would recommend the App to my friends and family.	EN	3.28	1.26
I was so involved in my searching task that I lost track of time.	FA	3	0.88
I felt annoyed with using the App.	PU	3.71	1.33
This search experience was demanding.	PU	3.93	1.07
The App interface is attractive.	AE	3.85	0.86
My search experience was rewarding.	EN	3.07	1.07
I felt discouraged while using the App.	PU	3.5	1.09
The App interface is aesthetically appealing.	AE	3.78	0.89
I found the App confusing to use.	PU	3.93	1.00
I felt interested in my searching tasks.	NO	3.57	1.02
I continued to used the application out of curiosity.	NO	4.00	0.87
The screen layout of the App appealed to my visual senses.	AE	3.42	0.75
I blocked out things around me when I was using the App.	FA	3.14	1.17
I lost myself in this searching experience.	FA	3.43	0.94

Table 3.8: Ambicities UES total scores and standard deviation for each question

For the Ambicities group and taking out PUs reverse coded, the three highest item scores were “I continue to used the application out of curiosity” (NO sub-scale, 4.0), “the App interface

Questions	UES Sub-scale	$\bar{x}$	σ
The time I spent searching just slipped away.	FA	3.29	1.32
I felt in control of the searching experience.	PU	3.93	0.62
I was really drawn into my searching tasks.	FI	3.21	0.89
I felt involved in the searching tasks	FI	3.21	0.70
Searching using the App was worthwhile	EN	3.57	1.02
I felt frustrated while using the App	PU	3.79	0.97
I was absorbed in my searching task.	FA	2.71	0.73
Using the App was mentally taxing.	PU	4.21	0.89
My search experience was fun	FI	3.86	1.03
I could not do some of the things I needed to do using the App.	PU	3.43	1.09
I consider my search experience a success.	EN	3.93	1.00
This search experience did not work out the way I had planned.	EN	3.5	1.09
I would recommend the App to my friends and family.	EN	3.93	1.07
I was so involved in my searching task that I lost track of time.	FA	2.64	1.39
I felt annoyed with using the App.	PU	3.93	1.00
This search experience was demanding.	PU	3.79	1.05
The App interface is attractive.	AE	3.71	1.20
My search experience was rewarding.	EN	3.36	1.22
I felt discouraged while using the App.	PU	3.86	1.10
The App interface is aesthetically appealing.	AE	3.36	0.93
I found the App confusing to use.	PU	3.79	1.12
I felt interested in my searching tasks.	NO	3.14	0.77
I continued to used the application out of curiosity.	NO	3.43	1.02
The screen layout of the App appealed to my visual senses.	AE	3.5	1.02
I blocked out things around me when I was using the App.	FA	3	1.04
I lost myself in this searching experience.	FA	2.71	1.14

Table 3.9: Twitter UES total scores and standard deviation for each question

is attractive” (AE sub-scale, mean 3.85) and “the time I spent searching just slipped away” FA item (FA sub-scale, mean 3.79). The lowest score was “I was so involved in my searching task that I lost track of time” FA item (mean 3.0). For the Twitter group and taking out reverse coded too, the two highest item scores were “I felt in control of the searching experience” (PU sub-scale, mean 3.93) and “I consider my search experience a success” (EN sub-scale mean 3.93). The lowest score was “I was so involved in my searching task that I lost track of time” FA item (mean 2.64).

The total UX score for each participant using UES were computed by adding the averages of each sub-scale together and dividing by the total number of sub-scales, because each sub-scale incorporated a different number of items. Each item was rated between 1 - 5. So, the maximum score for a given sub-scale was 5. Table 3.10 shows the scores for each of the 28 participants. The total score for both groups Ambiecities (mean 3.50) and Twitter (mean 3.40) is very similar. Thus, the overall UX is apparently similar for both applications according to the self-report UES.

Participant #	Ambicities	Twitter
1	3.63	2.62
2	2.97	3.75
3	4.59	3.39
4	2.88	4.08
5	4.39	4.21
6	2.25	3.25
7	2.73	3.92
8	3.56	3.48
9	3.59	3.71
10	3.54	2.84
11	2.91	2.45
12	3.55	3.40
13	4.14	2.94
14	3.54	3.50
Total	3.45	3.40

Table 3.10: Total Scores for all the 28 participants

Application	n	Σ Time	$\bar{x}$	σ
Ambicities	14	03:53:28	00:16:41	00:13:36
Twitter	14	01:44:14	00:07:27	00:05:57

Table 3.11: Time Spent Searching (Hours: Minutes: Seconds)

3.6.4 Browsing and Searching Behaviour

Although there are several interaction metrics for IIR, the only recorded behavioural metric was total time on the task. The total time was basically the time spent (in seconds) on the page where participants got the link of the search application whether Ambicities or Twitter, and they had a large text area to write or comment what they did during their search experience.

The simulated search sessions lasted an average of 12.06 minutes. Table 3.11 summarizes the time spent in the simulated scenario for all participants in both groups.

Participant's searching and exploring with Ambicities spent a total of 3 hours, 53 minutes and 28 seconds, almost the double of the Twitter group of participants who spent 1 hour, 44 minutes, and 14 seconds. Although there are more variability in the Ambicities sample, Ambicities' shortest search session was 4 minutes, and 17 seconds. In contrast to the Ambicities group, the Twitters' group shortest session was 53 seconds.

This research used nonparametric Mann Whitney U and Kruskal-Wallis test to contrast the measure of total time spent on task due to the given sample size and because the ShapiroWilk test of normality refuted the assumption of normally distributed data for this sample.

Total time on the search task showed that median latencies in the group of Ambicities and Twitter were 00:11:45 and 04:32; the distributions in the two groups differed significantly (MannWhitney U = 45, $n_1 = n_2 = 13$, $P < 0.05$ two-tailed, $r=.39$). Therefore, there is non-

parametric 0.95 confidence interval that a randomly chosen total time measurement from the Ambiecities group is higher than a randomly chosen subject from the Twitter group. This no-parametric test rejects one of the formulated null hypothesis at the beginning of this study regarding the effect of curiosity driven search (Ambiecities) in the time spent exploring during casual-leisure search on social media.

Pearson's correlation assuming a linear relation between time on the task and FA sub-scale for Ambiecities was not statistically strong ($r=0.30$), the measure show a positive relation. But, these two results add evidence to the hypothesis that curiosity driven search design encourage more FA and entice participants more exploration than the standard query-response paradigm that focuses in the pragmatic value of searching. In the discussion section, total time on task results are confronted with subjective responses of various parts in the questionnaires because these answers provide better a insight into "why" participants interacting with Ambiecities responded differently, both evaluating the FA sub-scale and searching, on the same search scenario.

However, if the sample is splited using the median of FA sub-scale in two groups. Those who did not experience focused attention ($FA < 3.0$), those who did ($FA > 3.0$); then with Kruskal-Wallis, this research tested the dependant measure of total time spent between the two groups. Participants who rated their FA above the median spent more total time on task (mean: 14 minutes, 08 seconds) than the ones who marked their FA below the median (mean: 7 minutes, 33 seconds, $\chi^2[2]=4$, $p=0.045$). So, grouping the participants give evidence regarding the strong relationship between FA sub-scale scores and the total time spent on the task.

Another important evidence of higher FA and cognitive absorption in the curiosity-driven group of participants, it was provided by the estimated total time. As mentioned in the procedure section, participants were asked to guess the total time they spent during the previous task immediately after they have finished the task. Table 3.12 shows all detail results regarding each participants total time recorded, estimated total time, and the absolute error.

Twitter's participant total task estimation summed up 1 hour, 55 minutes (estimated total task mean 00:08:13, absolute error mean 00:04:59). They overestimated their recorder time spent in 10 minutes, 46 seconds because their total time on task recorded was 1 hour, 44 minutes, 14 seconds (Relative error of 9.63%) . In contrast the Ambiecities group summed up a total of 3 hours, 15 minutes (estimated total task mean 00:13:56, absolute error mean 00:09:12). They underestimated their recorder total time by 38 minutes, 28 seconds because their total time on task recorded was 3 hours, 53 minutes, 28 seconds (Relative error of 19.73%). Therefore, this difference in absolute error and relative shows that participants' interacting with Ambiecities

Ambiecities			Twitter		
Recorded Total Time	Estimated Total Time	Absolute Error	Recorded Total Time	Estimated Total Time	Absolute Error
09:04	17:00	07:56	02:56	05:00	02:04
11:13	30:00	18:47	03:49	05:00	01:11
14:03	10:00	04:03	10:19	15:00	04:41
05:12	05:00	00:12	07:34	05:00	02:34
06:00	05:00	01:00	15:58	03:00	12:58
05:32	10:00	04:28	13:27	02:00	11:27
12:18	08:00	04:18	02:03	05:00	02:57
41:50	20:00	21:50	16:34	30:00	13:26
11:45	15:00	03:15	04:11	15:00	10:49
09:34	05:00	04:34	02:50	03:00	00:10
04:17	15:00	10:43	04:32	02:00	02:32
43:42	15:00	28:42	00:53	02:00	01:07
29:43	20:00	09:43	16:47	20:00	03:13
29:15	20:00	09:15	02:21	03:00	00:39

Table 3.12: Detail Participants Total Time on Task, Estimated Total Time on Task, and Absolute Error (Hours: Minutes: Seconds)

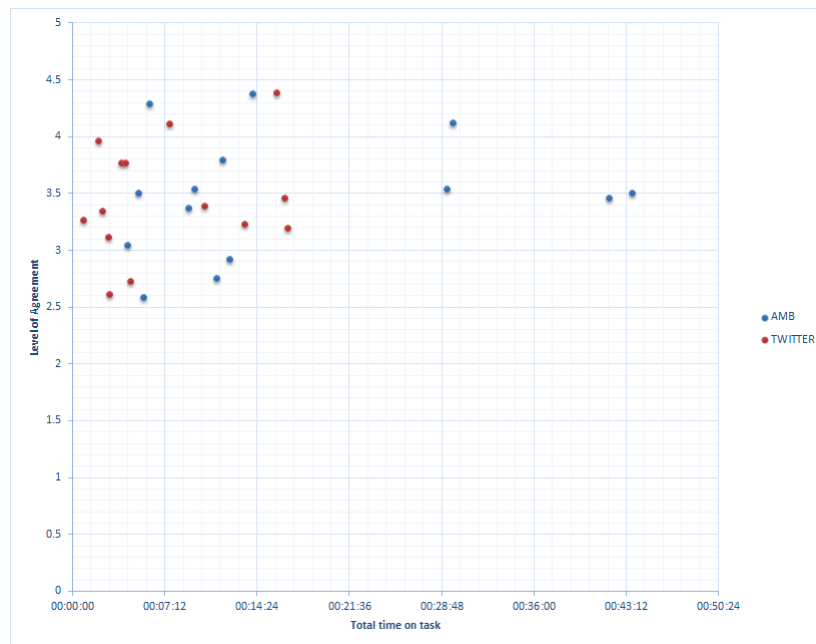
had less accurate guesses than those searching and exploring with Twitter. Supporting the idea that curiosity-driven design had a effect in the cognitive absorption (i.e. state of deep involvement with software) of participants during the searching task.

On the other hand, an Ambiecities participant in the laboratory overestimated the total time on task in more than 18 minutes, and this might be indication of a negative search experience instead of positive. According to Morae recordings, this participant had a very engaging initial experience (commenting and laughing while searching about some tweets), and finding interesting tweets when browsing their home country. It seems that towards the end of the interactive session, and when looking for tweets in New York, the participant found out some racist tweets with offensive content, and decided to stop. So, this bad disengagement could be reflected in the participant's estimation of the total time on the task, and the low focused attention score (FA sub-scale mean 2.6). Although the participant rated high AE sub-scale (mean 4.00) and agree to the statement "I continue to used this application out of curiosity", it seems the quality of the content affected the overall evaluation.

Therefore, Ambiecities interaction induced participants to state of deep cognitive absorption, but sometimes failed to provide an overall positive search experience, because there must be an equilibrium between the hedonic and utilitarian aspects of information searching (in this particular case a minimum of quality content). However, most of Ambiecities participants in the leisure search scenario assessed their search experience as positive based on the interaction itself as reflected by subjective comments as well as UES ratings.

Figures 3.8, 3.9, and 3.10 compare UES total scores, FA sub-scale scores, and NO sub-scale

scores with total time on task. Figure 3.10 shows an interesting relationship between higher NO sub-scales scores and the longer information search sessions.



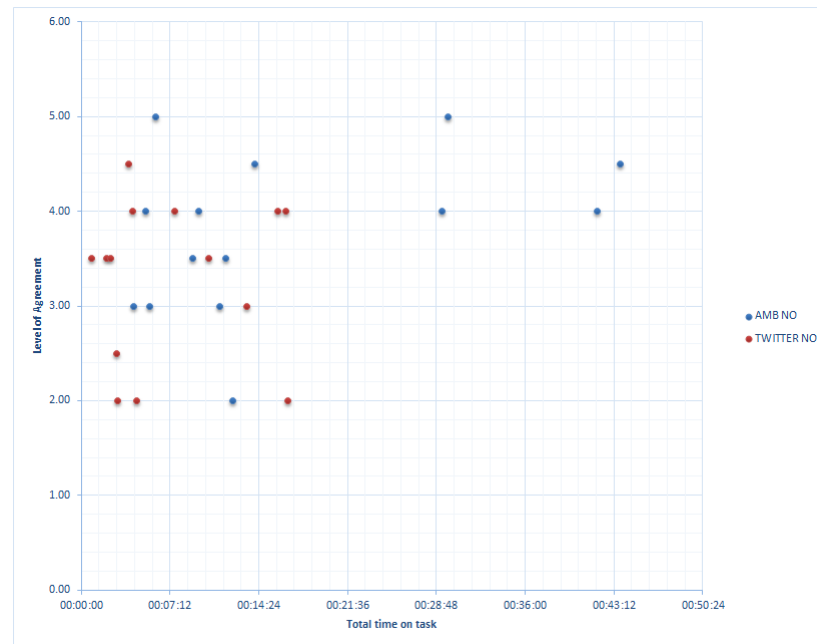


Figure 3.10: NO sub-scale Scores vs. Total Time On The Task

“I felt like I was clicking on tweets’ markers because I could not help it, not because I thought I would find something interesting. I click because I was feeling nosey. After 20 clicks I realized I did not care about them but I still wanted to read them. The tweet was there; I had to look at it”.

She also said about her experience: *“A good way to waste time [...]”* and when discussing the time spent she answered:

“I felt exploratory and just waiting for something to come up, clicking around and seeing new tweets pop up”.

Based on the laboratory observations and the self-reported questionnaires, the participants’ information searching behaviour using Ambicities usually followed an initial navigation to known places around their current city. Then they navigated to places where they had some emotional relationship or interest, like “my hometown”, “my home country”, other cities they have visited, where friends live, or where they wish to travel.

In contrast, most participants using Twitter followed look-up search with shorter search sessions. Others just used what people or organizations in their network were talking about. For example, Karen, who followed Londonist⁹, without triggering any search looked at her Twitter timeline when performing the simulated scenario. She found a tweet from Londonist and went to the official site. Later she said:

⁹<http://londonist.com/>

“I need to check Londonist more often, there are cool things”.

She described orienteering behaviour similar to Teevan et al. [13] by using her social context (e.g. friends, organizations she follow, etc.).

3.6.5 Relevance judgments in a Casual-leisure Scenario

As already stated, participants using Ambiecities spent more time on average exploring than those using Twitter even when they felt the content was not relevant, because they were driven by curiosity. This scenario illustrates the effect of eliciting curiosity using context triggered search for social media rather than query-response search. Designing the search experience with spatio-temporal context features (e.g Now, Recent, Near Me, Around a Location of Interest) encouraged hedonistic motivation instead of a pure informational by inviting people to keep exploring and waiting for more tweets whether relevant, non-relevant, interesting, etc.

In the post questionnaire, participants had similar scores to the EN and FI sub-scales. But when they were asked *“which of the following features made the information display by the App enjoyable for you? (Please tick all that apply) a)topically relevant; b)location; c)freshness; d) user interface; e) social relation”*: 21% participants using Ambiecities selected Topically Relevance in comparison with 71% using Twitter. This interesting finding highlights how the experience (e.g. the interaction using spatio-temporal context) was more important than retrieving relevant results. Figure 3.11 depicts the answers to this question. In a way, the search journey for some became more important than the outcome of the process in terms of relevance and usefulness (e.g finding a relevant tweet) as reported in [16].

When using Ambiecities, Caroline and other users’ behaviour were very similar to real casual search “wasting time scenarios” where behaviours like “exploring for the experience” and “needless browsing” were identified [16]. As already mentioned four participants used Ambiecities in the on-line survey for more than 29 minutes whilst the longest Twitter search session was less than 17 minutes.

In the sample, there is also evidence of participants being likely biased towards their ratings of the Twitter search experience. For example Charles rated his search experience as fun, and commented:

“It is a good distraction while I wait for my friends”.

Nonetheless the dwelling time recorded for him was 123 seconds (2 minutes and 3 seconds). Thus, given the simulated scenario (i.e. non-goal waiting scenario), this dubious behaviour raises the question if it was really fun. While Alexander, who searched 29 minutes and 15

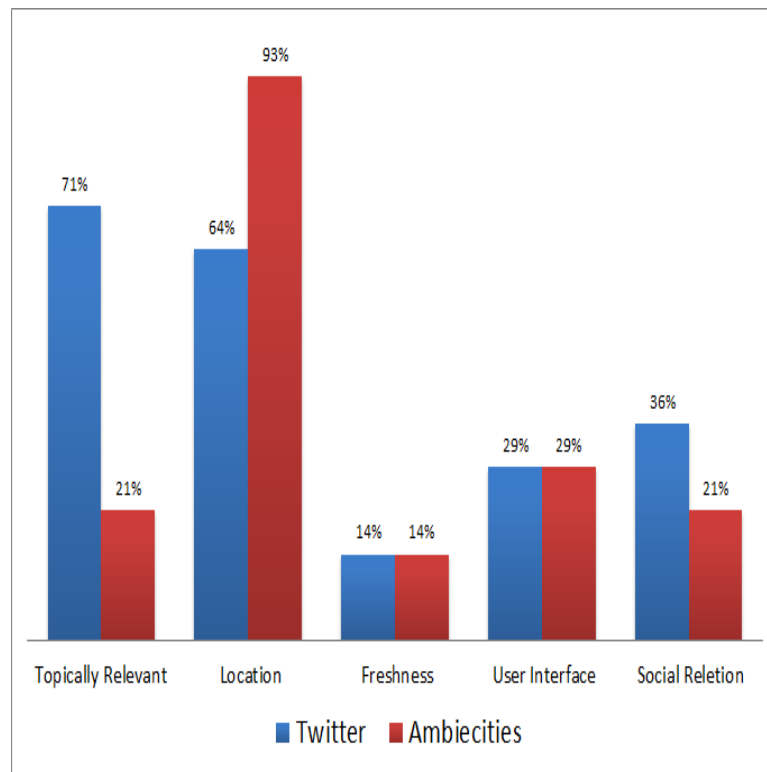


Figure 3.11: Which of the following features made the information displayed by the App enjoyable for you?

seconds using Ambiecities, stated:

“I was interested in knowing the type of content (photos & text) that people publish in remote places from where I live.”.

From his response, it is clear that Ambiecities achieves the goal of entertaining him whilst he was waiting for his friends in the simulated scenario. It is evident that Ambiecities influenced him to be interested in having an open ended and creative exploration rather than in locating a particular information or set of relevant documents.

The figure 3.9 illustrates that three of four participants that explore for more than 29 minutes had rated high FA. This figure also shows that 9 of the Twitter participants and 4 of Ambiecities searched for less than 8 minutes. Moreover participants, specially four participants using Twitter, rated FA and their overall UES as high, but perhaps as already brought out, they were biased Twitter search experience.

In the post-questionnaire, 71% of Ambiecities ($\bar{x} = 3.8$, $\sigma = 1.2$) and 50% of Twitter ($\bar{x} = 3.2$, $\sigma = 1.3$) participants answered ‘Agree’ or ‘Strongly Agree’ to the item “The time I spent searching just slipped away”. For example after finishing the task, Peter wrote on the text area where the users should put their results (e.g places or events, relevant tweets, news, etc):

“Absorbing and made the time go quickly with the map interface”.

Instead of submitting his results, he decided to explain how the application made him feel and described his search journey. Peter later described, why he felt this way when he answered the question, “while using Ambiecities, did you enjoy your exploration?”:

“Yes, it was great being able to visit different cities and regions following the twitter traffic”.

At the end of the session, after navigating around his locality, his place of birth and another city he had visited in South America, he asked the interviewer:

“Do I close the app?, Everyone will know where I have been”.

Spatio-temporal triggered social media content made some people focus on “being there” or being in some other spatio-temporal context rather than “being here”. So the participants who experienced cognitive absorption did not evaluate their interaction process in terms of finding relevant or useful tweets. Instead, they evaluated their search experience based how much they forgot their immediate surroundings (similar to escapism [16]).

3.6.6 Task Perception

Kuhlthau’s Information Search Process (ISP) [3] highlighted the fact that feelings and emotions such as uncertainty, confusion, and anxiety have an important place in the search process. However, there are some search scenarios do not hold the uncertainty principle of Kuhlthau because not all the search processes start from negative emotions, such as anxiety associated with the lack of knowledge [102]. Perhaps, this is true of casual-leisure scenarios where 1) not negative feelings, but positive might be the ones guiding the interaction; 2) the emotions and feelings might be more related with an hedonistic need than with conscious knowledge gap.

Moshfeghi & Jose [102] studied the cognition, emotion, and action aspects of different video search tasks, and discovered a significant variation between the search tasks characteristics. They found that for entertainment based search scenarios or task (i.e..entertainment by adjusting arousal level, and entertainment by adjusting mood) were perceived as easy and not-stressful.

Although during this experiment participants explored using one search task or scenario, this research wanted to investigate how the interaction with two different IIR-systems could affect the task or scenario perception.

Figure 3.12 shows that both Ambiecities and Twitter participants rated their task as easy,

interesting, clear and not stressful. This perception is similar to ENA and ENM participants in the study by Moshfeghi & Jose [102].

The variation in the answer provided by the participants for easy, interesting, familiarity and not-stressful are not statistically significant between applications.

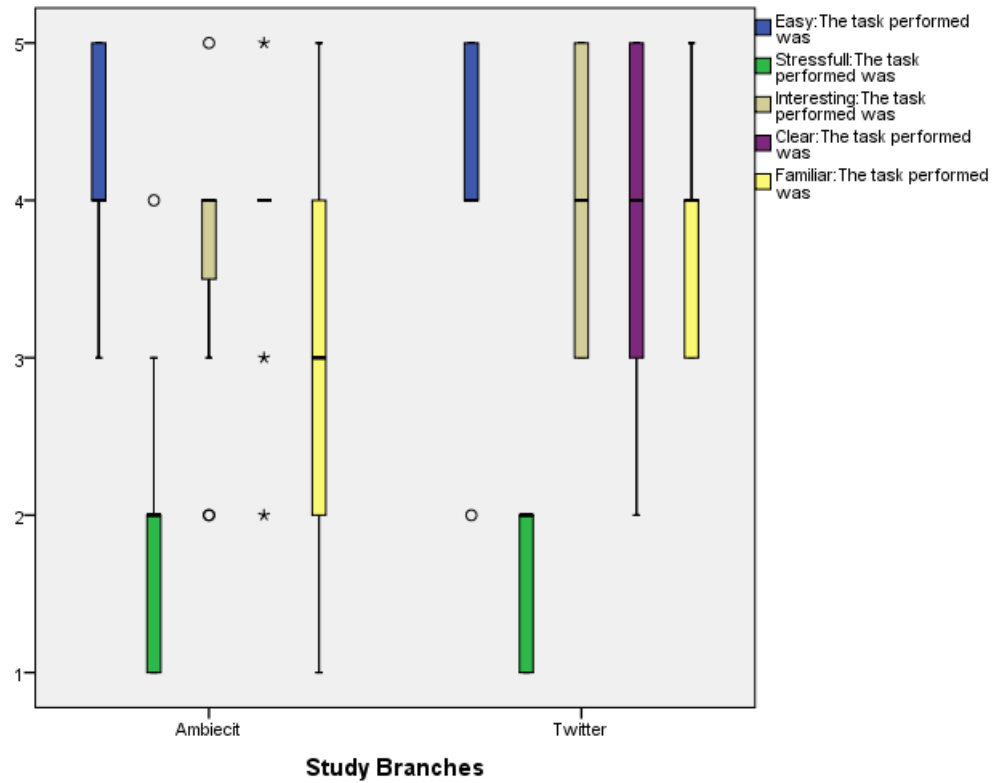


Figure 3.12: Task Perception Summary (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”)

3.7 Discussion

During this user study, the self-report measures were used to compare participants via non-parametric tests based on total time on task, and total estimated time on task. Using a similar mixed approached of measurements like O’Brien & Lebow for news information interaction with UES, behavioural and physiological data [117]. Because of the small sample in the study, this research is cautious about making inferences and conclusions from the data. However, there are some valuable insights regarding how curiosity-driven IIR-system might influence the UX in a social media scenario, and how hedonic aspects of search can become more important than the pragmatic or utilitarian aspects of information searching, specially in the context of leisure IIR research.

Instead of creating new ad-hoc self-report instrument to measure UX without an established

validity [28], a sub-set of O'Brien & Toms UES was used as main framework to measure the search experience [26, 24]. UES provided a holistic picture of the hedonic and pragmatic aspects of information searching UX between Ambicities (Curiosity-driven design) and Twitter (Query-response design). There was significant reliability and good internal consistency for most UES sub-scales even without eliminating any of the items for both applications. Although there were some correlations suggesting overlap among items of different sub-scales, this research did not perform a factor analysis like O'Brien & Toms study of UES as UX measurement for exploratory search [24] because of the small sample size. Overall, UES was key tool to evaluate different facets of the multidimensional UX, spanning from how participants assess the IIR-system (i.e. usability, aesthetic appeal, novelty), to their cognitive and affective states (i.e.. felt involvement, focused attention, enjoyment, endurability) [117].

From the self report data, recorded time on task and estimated time on task, this research was able to reject the null hypothesis stating that participants having a curiosity-driven search experience do not feel more focused attention (FA) than those using classical query-response paradigm. No other UES sub-scale had such a significant statistically difference between Ambicities and Twitter groups of participants.

There was also a significant association between FA high scores and total time on the task for participants in Ambicities. This research expected this result based on O'Brien & Toms (2010) e-Shopping study who reported that higher engagement participants would spend more time interacting with information system than those who were less engaged, because the pragmatic value of search could be less important to a user who is truly enjoying and immersed in an experience [113].

Another insightful result, from the different UES sub-scale scores for Ambicities and Twitter participants, is that participants' overall evaluation of the searching experience was differently built upon pragmatic and hedonic aspects between the two groups.

For example, the Twitter participants score higher FI (mean 3.43), PU (mean 3.84), and EN (mean 3.46). FA was the lowest sub-scale (mean 2.87). The values of these sub-scales and their correlation are similar to previous studies by O'Brien & Toms, and O'Brien & Lebow [113, 24, 117] where FA was always the lowest sub-scale. The EN, FI sub-scales results, and total time on task for the Twitter participant could be contradictory if they are analyzed using the leisure search scenario, and the a priori assumption about the purpose of interacting with IIR-system (i.e. to be entertained or engaged while waiting for their delay friends).

Although "users mental states are subjective, unobservable, and impossible for a researcher to infer, hence the measurement dilemma"[117]. For example, one participant interpreted that

the purposed of the search scenario was to get relevant information for their friends. So, the participant's evaluation was based whether or not the IIR-system gave high topical relevant information about the town or location in the simulated scenario.

In the Twitter group, PUs sub-scale was the highest scale score. Comments from Twitter's participants reflected "why" this sub-scale was vital for the overall search experience¹⁰ :

"This was interesting because I found my hometown, "Ewell". I found out that there is an antiques fair tomorrow and I think that I might go!. I also searched "Yuen Long", my mum's hometown in Hong Kong,I was interested to find out that there's an organic farm there [...]".

"I did some research about my favourite place in London, St Katherines Dock. Then I looked up Palais Brongniart because the company where I work is having a conference there next week, there was no info or recent posts. Then I came across 'Everything London' and it tells you about all the events happening in London and it is very useful so will be using this in the future .

PU sub-scale measures the pragmatic aspects of search, and Twitter high topical query-response paradigm helped them to find efficiently "interesting" and "useful" results based on their interest. PU scores resembled a relationship with what was presented in the section of relevance judgments where topically relevance was the most important factor to enjoy the exploration for Twitter participants. Therefore, query-response designed focused the complete user experience in the results and quality of content.

However, when the Twitter's participants were asked¹¹ if the enjoy their exploration, their comments reflected that some did not have fun while searching because of the type and quality of the content, but also due to the interaction. These are some comments from the Twitter participants regarding the UX:

"The exploration is not fun but practical, it is very easy to find information".

" Yes, but it was very basic. I would have liked to have other visualization tools.

Some had photos, but they were not always useful because they were selfies".

"Yes, normal".

Although high quality content is important to have an engaging and fun experience [24, 117] because fulfilling pragmatic aspects of search is vital to generate positive emotions (e.g. the satisfaction of finding an interesting tweet after exploring for long time) and the content itself can

¹⁰This comments are extracted from the text area below the following statement. "Please use Twitter for performing this task. Write or paste the results you find while making the task and take your time. Click here"

¹¹ While using the App, did you enjoy your exploration?[Please explain and give examples].

have a big hedonic value and influence user's relevance assessment (e.g. tweets with high quality pictures, funny tweets, etc). There are other ways to create a positive UX (e.g. interaction, aesthetic appeal), because both pragmatic and hedonic aspects of information searching are important in the overall evaluation even though they are perceived independently [33]. Therefore, IIR-systems should help searchers to reach a destination (i.e. precision, recall, usability), but what destination? and how? "Cold" query-response design seem to encourage minimum exploration and does not entice people to search beyond what is "normal", according to the total recorded time and subjective comments.

The last highlighted opinion (i.e. "Yes, normal") regarding the search experience reminded this research of previously cited words of Karen Spark-Jones, when commenting the weaknesses of an IIR evaluation purely based on situational judgments in work scenarios: "I liked what I got" from my search results could also imply "I am not worried about what I did not" [15].

In contrast, Ambiecities's participants rated NO and AE as the highest sub-scale. For example, some participants in this group answered regarding whether they have enjoyed or not their exploration the following¹² :

" Yes, It was easy the interaction and the user interface attractive".

"I liked the presentation of the map, and the easiness that you can navigate it. The colours were good, the interface was simple but good-looking. I liked the fact that you can view people's tweets".

"Yes, The visual interface was nice. (from desktop computer). I would be interesting to see it from a mobile interface".

"The exploration using a map is fun, but my personal interest decreased with the content of the tweets".

"Yes, It was fun to use this type of application because this kind of system can help people to explore and enhance their knowledge".

These comments suggest that Ambiecities fulfilled the aesthetic appeal of participants, one hedonic aspect of information searching. Is this important for an IIR-system?. According to O'Brien & Toms for typical search system, AE could be more "important at the beginning of an interaction in order to capture users attention" [24]. Nonetheless, this assumption must take into account the type of IIR-system, the search scenario and the information been retrieved. For instance, an e-Shopping search experience might need high levels of AE throughout the

¹²See previous footnote.

whole interaction in order to help users to explore, evaluate and compare products. Similar to the e-Shopping scenario [113], AE influenced the overall evaluation of UX for the Ambiecities' participants and it was key factor at the start of the search session. But are there more hedonic aspects of information searching?

The Ambiecities group of participants rated NO statistically higher than those in the Twitter group. NO sub-scales measures the curiosity evoked or satisfied by the experience, and interest in the task. In previous studies by O'Brien & Toms, the NO sub-scale had often a lower score than the other UES sub-scales, but it was a good predictor of user engagement. [113, 24].

In a previous study, O'Brien (2011) demonstrated that Novelty, AE and FI were key factor in maintaining user engagement with on-line news content [118]. During this laboratory study, novelty was noted "in headlines or content that evoked curiosity because they were outside the norm". Other times participants selected articles that they could relate to it personally whether or not the topic or issue was new. She summarized her paper saying: *"Overall, participants in this study may be underscoring two-types of engagement with on-line news: one that is rooted in the novelty and quality of the content - the need for story - and one dependent upon interactivity and aesthetic appeal.[...]Thus, if the interface fails to engage the user, the content may not matter; at the same time, the interface elements and the content may compliment each other and bring about an engaging experience"*[118] . Although O'Brien mentioned the importance of novelty (one collative variable of curiosity), in this particular paper she associates novelty only with quality of the content (i.e. the capacity of the content to trigger curiosity).

Nonetheless, Ambiecities tested that curiosity can be elicited based upon interactivity and aesthetic appeal as well as high quality content. Most of Ambiecities' participants acknowledged that they felt curiosity while interacting with the system whether commenting about their search experience or rating "Agree", "Strongly Agree" to the statement "I continue to used the application out of curiosity", even though the quality of the content was poor. According to the results section of Ambiecities group, the feeling of curiosity was vital to entice and encourage higher FA, total time on task, a higher absolute error between perceived total time on task and recorded total time, and the overall user engagement. This findings supports the hypothesis that mapping curiosity arousal principles to the design of IIR-system could create a better the search experience (RQ1) and increase users exploration (RQ2).

The fact that EN sub-scale is a bit lower for the Ambiecities participants than for the Twitter ones reflect that pragmatic aspects of information search such as retrieving high quality content can influence hedonic aspects as well. For example, a participant searching a long time and none of the found tweets have minimum of quality. Then, perhaps after initial engagement and

enjoyment, the participant could experience boredom, and frustration because the IIR-system is not given any noticeable utilitarian value. User's comments reflect that some of them thought they were wasting their time because of the content retrieved by Ambiecities .

However, Ambiecities had higher recorded total time. This suggests the existence of one or more interactive features in Ambiecities design that trigger curiosity and make experience engaging, and enticing for participants in comparison with the Twitter search interface with probably higher topically relevant content. Some of those interactive features, mentioned in the section about relevance judgments for casual-leisure, must have been more important than the perceived utilitarian and pragmatism value because most of Ambiecities' participants rated Ambiecities with better overall evaluation of UX than the ones using Twitter search service for the leisure search scenario.

Another important objective fulfilled during this study was that this research was able to study casual-leisure search; this research recorder and observe casual-leisure search behaviour such as "needles browsing" and "exploring for the experience" in an experiment setting. Instead of design the experiment around the creation of a "truly" hedonistic motivation, this research created the casual-leisure experiment around two assumptions: 1) search has some hedonic and utilitarian aspects related both to content and interaction that can trigger positive or negative emotions. Those 'casual' emotions could be the drivers of user experience that reflects what people normally do in order to get enjoyment; 2) the search experience can be design to generate or encourage the existence of hedonistic need such as curiosity. This methodological innovation was vital to the success and viability of the experiment.

3.8 Limitations

The user study had limitations regarding the sample size and it was affected in terms of the recruited subjects (i.e. most of them students). The pre-questionnaires assessed common familiarity with social media but not personality factors which seem to be very important in order to be in 'flow' [23]; and the possibility of experiencing the feeling of curiosity that could also be understood as a personality trait.

For one of the five laboratory subjects, the presence of an observer had a negative effect on their curiosity-driven searches and their total time. However, this research was able to record in the laboratory casual-leisure search sessions where participants laughed watching new tweets pop up, explored in depth and had similar scores as the rest of participants.

Regarding, the simulated casual-leisure search scenario , this research is going to define

a more realistic scenario of “everyday life” because according to Borlund’ approached [29], simulated search scenarios should be ‘real’, not something that users’ would imagine.

Also, although Twitter interface was the chosen baseline to compare Ambiecities, there were huge difference between them, and it is particularly difficult to assess the effect of particular components in Ambiecities. Moreover, the empirical observation has given this research some clues and ideas that will be validated in a further studies where the baseline would be visually similar to the curiosity-driven application avoiding bias towards Twitter.

Another limitation is that the only behavioural data recorded was total time on task. For future studies, this research plans to used other metrics such as mouse clicks, query, documents visited, etc.

Regarding the total time on task results and the interpretation of this metric, this research has showed from data total time can be an indicator of both positive and negative experience. So, the behavioural data should always be analyzed in the correct context (i.e. the search scenario, the system, the user) and with other subjective information such as questionnaires or open questions.

3.9 Implications

After the IIR experiment and the analysis of results in section 3.7, the implication for IIR are presented as follow.

3.9.1 Designing for the utilitarian and hedonic aspects of information searching

This research presented the design of an IIR-system built upon curiosity arousal mechanism. Drawing from the results, user’s aesthetic appeal, novelty and focused attention were significantly higher than those interacting with a query-response IIR-system even though some users like Twitter just because they biased their social network or the evaluated in terms of what normally is an IIR-system. So, were the actual search and exploration really different between both application?, Did the normal search behaviour change when users were prompted with curiosity-driven search interface?

Indeed, Ambiecities’s participants were enticed and encouraged to explore for more time and they kept interacting even after getting low content tweets. As already quoted, Caroline’s comments describing her search experience summarized the power of both interactivity and aesthetic appeal of Ambiecities to generate curiosity and a estate of deep involvement with the

IIR. She said, *“I click because I was feeling nosey. After 20 clicks I realized I did not care about them but I still wanted to read them. The tweet was there; I had to look at it”*. Peter comments is also insightful into key factors of the experience and the benefits upon the UX: *“Absorbing and made the time go quickly with the map interface”*. Therefore, the curious interaction trigger their emotions and feelings of curiosity and they did not stop the searching and browsing after exploring for almost 12 minutes. The Ambiecities’ search experience enabled also a more creative, and in depth involvement than those using Twitter. Charles and other participants described their search experience help them to go beyond a quick look up or locating search. Is it really important this finding for IIR outside of leisure search?

Absolutely, White & Roth (2009) mentioned that for exploratory search (an emerging are in IIR) is vital “learning, cognitive transformation, confidence, engagement and effect as well as result relevance and utility” [11]. For example consider that IIR-systems are not just tools that satisfies information needs, but environments in which humans interact with information content in order to learn. Following the basic assumption of UX that any product has pragmatic and hedonistic aspects [119]. How should IIR-system support learning? Perhaps following the “I liked what I got” from my search results and “I am not worried about what I did not” [15] “is not such a good idea, because usually some demotivated students are simply not “worry” about learning anything, and most of their findings could be biased the first three results in ranked SERP. Users, who often are untrained to use the IIR-system, need to develop confidence, curiosity and engagement in order to know their learning environment, and being able to use it to the full in other discover new valuable information. Perhaps through the generation of emotions and feelings, IIR-systems could guide participants to be “worry” for what is unknown to them; or at least for what is coming or next to happen in the interactive search experience.

The presented learning scenario is an example of the many fields in IIR that could benefit from understanding and mapping curiosity arousal mechanisms to the design of the search experience because there are many search scenarios where users need to explore unknown information space but they might lack the motivation to do it; on the other hand, search providers wants their users to be engaged with the search experience because they know “more time” and “enjoyment” could mean more money and better buying experience (e.g. e-commerce websites). O’Brien & Toms e-Shopping study showed that hedonistic aspects of search were very important during the whole experience.

Therefore, if IIR continues to design and evaluate based only on utilitarian information searching aspects, it would missed the “bigger picture” search, and new possibilities of improving the UX of IIR-systems [24, 113].

The remaining challenge for this research is to identify and extract if possible what are the key features of the interactivity, and the quality content for a curiosity search experience in social media?. What were the individual features that trigger the exploration and entice participants to keep searching? What are the most important arousal mechanism in UX? Is it possible to isolate them?. This last question seem to be impossible, reflecting for example in the interrelationship between the quality of content and the interaction for the UX. This research echos O'Brien words: "*if the interface fails to engage the user, the content may not matter; at the same time, the interface elements and the content may compliment each other and bring about an engaging experience*" [118]. Thus, the challenge could be, how does IIR-system achieve and support the user's engagement through out different stages of the search process?

After observing user's behaviour and responses in the study, this research has hypothesized that for many of the participants, the continuous interaction (i.e. push and pull) and the surprise effect of getting new markers into the map whether topically relevant, location relevant or "just saying good morning" generated uncertainty, and the need to click a new result (i.e. making people "worry" or nosey for the next tweet to pop up). Even when some felt frustration for what people were saying in Twitter and stopped browsing after getting a negative result (e.g. tweets about pornography, racism, or just in other language that the user did not understand), they acknowledge the feeling of curiosity and deep cognitive absorption while using Ambiecities. Therefore, this research will keep trying to make curiosity-driven experiences by generating uncertainty and surprise whilst the user's engagement is modeled by tracking the users' interaction in the searching process (e.g. mouse moves, clicks, etc.). This hypothesis are based on subjective responses, the NO and FA sub-scales scores and the total time on the task presented in the results section.

3.9.2 Measuring social media search experiences in a casual-leisure experimental setting

According to Elswailer et al. [16] user's with a strong hedonistic motivation follow search behaviours that break traditional IIR models after gathering user's self-reported search experience on social media and studying people's information seeking behaviours when using television. During these search information behaviour session, users might find topically irrelevant (e.g. low content) information but they may keep exploring because the system satisfies their current leisure need, or they might keep exploring because every result they get has content so curious and funny that they want to keep exploring even beyond what they have already plan. How could the the search experience be measured? and How could this "casual-leisure" behaviour

be study in a experimental setting?

For example, Colbert & Boodoo [120] reported that, in experimental environment, even if participants are instructed to browse and search an unlimited amount of time, they may still control their time in order to complete their task and make difficult to establish statistical difference in the user's engagement between two websites. Therefore, this research concluded that an open-ended leisure search scenario was not enough to generate "causal-leisure" behaviour.

As mentioned in the methodology section, Wilson & Elswailer [31] proposed the creation of an experimental setting where "users are provided with hedonistically motivated task" by making a faux-study (e.g. in a laboratory setting, users are told there is an unforeseen delay or problem, and instructed to use a computer while they wait. Then, after some time when they seem to be bored, then the study continues).

However, this research instead of providing users with "hedonistically motivated task" let participants interact with a social media IIR-system and observe how naturally they were able to enjoy their UX and how much they were distracted or entrained having in mind their current simulated situation. Then this research hypothesized that participants exploring with an IIR-system built mapping concepts of curiosity arousal mechanism from psychological theory into the search experience could cause casual-leisure search behaviours such "Exploring for the experience", "Needles Browsing" in a experimental setting. According to the literature review conducted by this research, there has not been presented in IIR a study casual-leisure search in action so far.

This methodological shift meant moving the hedonistic perceived value from the search scenario description and protocol planning to the interaction with IIR-system built around an hedonistic need (i.e. curiosity). Following the aforementioned assumption of UX that any product has pragmatic and hedonistic aspects [119]. Therefore, if the IIR-system highlighted the hedonic aspects of information searching, may be users would find the needed motivation to keep exploring for more time.

There were to ways of evoking curiosity in IIR-system: content and interactivity. For this experiment, this research choose to focus in the interactivity part, and the content was so how similar in the sense that the source of information was Twitter. As a result, various participants using Ambicities driven by curiosity show similar behaviours as the ones reported by Elswailer et al. [16] where users explore out of curiosity, and for more time than the previously estimated for them. This finding opens another methodological door to study casual-leisure search in action, and it is one of the novel contributions. Although the remaining challenge would be to define a set of recommendation to make the experimental setting reproducible in other IIR

fields.

Now, the other question was ‘How could the the search experience be measured?’. This research has shown at least two lessons. First, measuring the search experience should go beyond a binary decision (success/failure) and the consideration of pragmatic aspects of information searching. Although IIR should continue to improve pragmatic aspects of information searching because the quality of the content, and other usability aspects are important in the overall UX. The hedonic aspect of information search are vital to the UX as well.

O’Brien & Toms (2008) UES [26] seems to be a practical framework to show the “bigger” picture of search. As this experiment highlighted there much more than usability when interacting with an information system. The hedonic aspects of information searching were important even when participants interacted with the “cold” classical query-response paradigm of the Twitter search service. The hedonic aspects were the ones that make the difference between following a simple look-up search and moving on; or explore out fun for a large period of time. Harold and Eric, who were the Twitter participants with the longest search session wrote when describing their search experience:

“I enjoyed discovering new things. I found out things I didn’t know about certain areas of London. I also discovered links to restaurants and chef post. ”.

“People were publishing about the terrible storm that falls in Cali, and the resulting traffic jams and accidents. There were some news of Cali that I did not know and I must take them into account. It was easy, quick and making good use of the social platform”

Harold, and Eric felt motivated by his desire of finding fresh or new information. Eric, specially, felt a strong personal and social interest in the news about the storm, because this was happening at the same moment he was searching and where he lives. Social media is a source of information that it seems free from third parties like news providers such as newspapers and offers the potential of fresh relevant content by any citizen of the world. Twitter search service has an interactive feature that allows users to monitor the latest tweets after the user has performed the search query (i.e. first users get a pop-up message on the top of the rank list with the number of new tweets, then the user clicks and gets in top the most recent tweets).

The Twitter group of participants reported an UES rating and behaviour similar to O’Brien et al. [24, 113, 118], specially in her study of user’s engagement with on-line news (2011), she found that newsreaders engage with news stories they could relate to on a personal or societal level. Also McCay-Peet et al.[121], they reported that participants were distracted by content they found personally interesting but that was not relevant to their assign task, during their

investigation of how ‘catchiness’ (saliency) of relevant information affects user’s engagement (specially FA sub-scale). Although all these IIR- systems focus their complete attention in the results (i.e. good quality results) and they were based on the pragmatic query-response, what seems to make the difference between the user’s engagement and exploration was whether or not participants in the experiments were able to associate their personal or social interest with what they were getting from their searching process. In other words, the search process was not motivated by utilitarian factors, including efficiency and cost, but by their desire to satisfy hedonic needs, such as affect, social interaction and/or entertainment. Both qualities (hedonic and utilitarian) are key to the search experience and they can complement each other (i.e. one component is low, but the other high, and vice versa).

Therefore, IIR should advocate more interest in understanding hedonic aspects of search, designing IIR-systems around those hedonic aspect (both the interface and the interaction), creating metrics to capture those hedonic aspect, and maturing the metrics through the evaluation of different scenarios and experimental settings. IIR should start looking to other more mature fields of human behaviour like Psychology. Fore example, IIR should analyze how certain personality traits (e.g. autotelic personality) can affect the search experience, what similar scales and methods psychology and other fields have used to study the irrational but common behaviour of humans.

Second, measuring search experience should be done asking the “why” and “what”. This implies that the evaluation must take both psychometric questionnaires (i.e. UES, and not ad-doc satisfaction questionnaires), and subjective open questions; and behavioural metrics (i.e. interaction metrics). For example, Toms & O’Brien (2013) newsreaders user study showed that time on the task could be an indicator of both positive and negative experience. For example, some participants rated low their cognitive absorption and spent more time exploring because they were disoriented and feeling lost. Therefore, this mixed methods approached of evaluation is more holistic and brings more weight to the arguments and conclusions.

This research have stated that user’s performed behaviours similar to Elswailer et al. [16]. How did the mixed approached help this research to make this claim?. First, as already presented, subjective responses to the post-questionnaire show that specially users in Ambiecities explore aimlessly during the search scenario and focusing in enjoying the experience instead of finding anything in particular. Second, psychometric questionnaires revealed that the Ambiecities group felt more FA (i.e. cognitive absorption), and NO (i.e. due to the curiosity interaction design) than those interacting with Twitter. Both high FA and NO scores also suggest that participants where in a state of deep interaction with the IIR-system because their were driven

by curiosity. In other words, they were searching to satisfied an hedonistic need, not an informational. Third, Ambiecities group spent more time searching and had an average absolute error in the estimation of total time on task larger than those interacting with Twitter. According to the UES and the open ended questions, the total time spent on task and the average absolute error of total time on the task were related with high FA, AE and NO ratings. So, the aimless exploration was not the result of people frustration in a given task, but the feeling of curiosity evoked by the interaction and the aesthetic appeal of the user interface.

3.10 Conclusions

Understanding leisure search behaviour would help to go beyond the “cold” query-response paradigm and design search experiences, where the search journey itself is as important as reaching the destination.

This research gathered psychometric self-reports (i.e. UES), subjective responses to open questions, recorded total time on task (i.e. behavioural or interaction data), and absolute estimation error of total time on task in order to understand, and evaluate the effect of designing an IIR-system using curiosity arousal principles taken from psychology theory of curiosity in comparison to query-response in casual-leisure simulated scenario.

The evaluation analyzed both pragmatic (utilitarian) and hedonic aspects of the search experience were collected using UES of O’Brien & Toms [26]. Unlike in past research [113, 24], FA sub-scale values were on average higher and closer to the UES sub-scales mean value for the group interacting curiosity-driven IIR-system. FA sub-scale scores suggest that participants were in a state of deep interaction with Ambiecities and participants subjective responses showed that they might have explored beyond what they planned.

NO sub-scale scores were the highest sub-scale score for the Ambiecities group. This high values were empirical evidence that participants were driven by curiosity, and strongly entice to keep exploring. Their searching behaviour was similar to casual-leisure search behaviours described by Elsweiler et al. [16] (“wasting or killing time” or “exploring for the experience”) where the experience of searching and the satisfaction of an hedonistic need were more important than finding a particular piece of information or solving an information need. Although this research also acknowledged the obvious assumption that bad results may some participants prone to disengage during the exploration, there were other participants who kept exploring by an almost irrational desire evoked by the curiosity-driven system.

The recorded total time on task, and absolute average error of estimated total time on

task measurements corroborated FA and NO ratings and brought more evidence to support the hypothesis that curiosity-driven design of IIR increases cognitive absorption and exploration. There are many type of IIR-systems (e.g. browsing, searching, learning) would benefit from a better understanding of how design the search experience to evoked user's engagement.

The remaining challenge for this research is to investigate how should the hedonic value of both interactivity and content be used to create a better searching experience based on curiosity arousal theory? How could this mapping of curiosity theory be done in other IIR scenarios? (e.g. news reading), How would this affect the information searching behaviour?, and How could curiosity be model by the IIR-system in order to support an optimal engagement?

IIR field needs to stop evaluating the searching process in terms of the outcome (i.e. the search was successful because IIR solve my information needs) in order to develop a more holistic evaluation of search as an experience. During these user study participants UES ratings and behavioural data showed that the search experience of the Ambiecities participants was superior than the ones interacting with Twitter in the leisure scenario. According to AE, NO, FA sub-scales scores, the experience was highly ranked because curiosity-driven design focused the attention of participants in the hedonic aspects of information searching rather than pragmatical and utilitarian value.

Chapter 4

Curious query-response design for leisure news search

The previous chapter described the findings of a leisure search study from a combine pragmatic (utilitarian) and hedonic perspective of the search experience. The restrictive aspects of standard query-response paradigm and the typical rank list interface design were demonstrated by comparing and contrasting users information interaction between a “new” curiosity browse map software prototype (i.e. Ambiecities) and the “traditional” Twitter search service for leisure scenarios inspired by the question: What is happening around me? The study looked at a set of leisure scenarios derived from the literature reviewed in chapter 2 and showed evidence on how by evoking curiosity using Berlyne’s collative variables [18] (i.e. novelty and uncertainty) in the design of the search experience, participants were more likely to experience high cognitive absorption and perform an in depth exploration instead of a typical look-up behaviour. Also, individuals were enticed by the interaction, and compelled to explore longer to enjoy the experience (e.g. forgetting they were waiting for their friends, and to have some fun) rather than to find relevant information.

This chapter builds on this work by focusing in the study of how novelty and uncertainty, two important curiosity collative variables, can be applied in the well-known query-response information interaction paradigm and enhanced the search experience in a news leisure search scenario by supporting user engagement and exploration. As in the previous user study, this research adopts UX lens to explore and evaluate users search behaviour and systems performance.

The chapter is divided into four main parts. The first part of the chapter summarises the lessons learned, and discusses limitations from the first study to then provide a motivation for this second leisure news study. This second part describes how novelty and uncertainty curiosity arousal principles are applied in the information interaction and the software design

of the search prototype created for the leisure news search study. This section specially discusses the first version of the curious companion software module. The prototype uses Twitters social media API to get relevant news and come with two flavours of information interaction: cold query-response and curious query-response. The third part reports the experimental design, the data collection, and the search scenarios as well as a general demographic description of the participants. The fourth part presents the results and deals with the main aim of this work to determine if taking curiosity arousal principles into account in the design of IIR systems is advantageous in a leisure scenario from both utilitarian and hedonic perspectives of the search experience. This section also explores the role of personal interest and curiosity as trait in the searching process.

4.1 Reflection on experiment 1

The following portion explains both lessons learned and the limitations of the previous user study. Thus providing a rationale for a new user study with a better setup, tools, and metrics.

4.1.1 Lessons Learned

The initial experiment helped this research to have better insight on the positive effect of evoking curiosity by mapping novelty (e.g. surprise) and uncertainty into the information interaction design of IIR-systems [18] (RQ1).

As documented in the chapter 3, participants interacting with a curiosity driven IIR-system rated higher focused attention in the UES, and had longer search sessions. Therefore, Ambicities' participants were enticed to explore out of curiosity and performed an information searching behaviour similar to "Exploring for the experience" or "Needless Browsing" identified by Elswailer et al. [16] than those interacting with classical query-response paradigm (i.e. look-up search, then moving on) (RQ3).

Although IIR has focused the attention mainly in the pragmatic and utilitarian aspects of information searching, there is growing awareness in the field to understand search beyond the work oriented paradigm and measure the 'hedonic' aspects as well [53, 16, 26, 113, 102]. Several work and leisure scenarios could be greatly influence and perhaps improve by designing for this hedonic aspects because searching and browsing are human decision making processes (e.g. e-learning, e-commerce). Therefore, 'emotion' rather than 'reason' could have a large effect in the information searching behaviour and satisfaction of users. The experiment described in chapter 3 shows that when the IIR-system evoked participants' curiosity, they were enticed and

engaged to explore for more time than those who did not (RQ2).

The previous social media experiment also presented evidence that the interactivity factor of the user search experience could be as important as the content or the documents delivered by the IIR-system, especially for participants whose motivation was to seek pleasure and escape from a ‘boring situation’. This is an interesting finding because traditionally most IIR-systems based their whole user experience in the quality of the content they are able to give.

However, not all search scenarios are about locating and finding information [10, 11], there are some scenarios like learning where enticing and supporting exploration with interactivity and aesthetic factors could be as important as improving the average precision of the retrieved documents. For example, in education scenario a math teacher might not explain in detail all the steps to solve a homework if the teacher wants their students to learn and develop the necessary skills to resolve the problems. The teacher main challenge is not to give an answer but to motivate the students to research the subject, put effort in the solving process and perhaps enjoy the acquisition of knowledge. Following the same logic in an IIR learning scenario, the main challenge for IIR-systems is not to give a response but to entice and encourage a positive learning experience. One could imagine that an ideal IIR-system should probably be designed to penalise precision at the beginning of the search session, entice users exploration throughout the whole decision making process; and create an engaging (or possibly fun) search experience using interactivity and other hedonic aspects of information searching.

Does the data of the previous experiment provide evidence of the search journey being as important as reaching a destination? The previous experiment demonstrated that in some cases if the information interaction design (e.g. the aesthetics of the user interface, the interaction information features) engages the user and evokes curiosity, the content may not matter for the overall evaluation of the information system or at least the content might be less important for the overall search experience and outcome of the scenario (e.g. casual-leisure search, learning); at the same time, the interactivity elements and the content may benefit each other and bring about an engaging experience [118].

Overall, the main implication to IIR from this PhD work is that the previous study highlighted the possible benefits of designing for both utilitarian and hedonic aspects of information searching by mapping the curiosity arousal principles (i.e. novelty and uncertainty) into the information interaction design. This research also reported a group of suggestions of how to create a ‘casual-leisure’ experimental setting and measure the influence of a hedonistic need (i.e. curiosity) in the searching process.

4.1.2 Limitations

There were several limitations with the first experiment, as already indicated in chapter 3. Thus the research questions were partially answered, and this research needed more evidence to show the benefits of applying curiosity theory into IIR. These are the identified limitations:

- (a) **The information interaction design between the two apps was completely different.** For instance, Ambiecities had a map graphical user interface and triggered content based on the users' exploration of the desired spatio-temporal context; on the contrary, Twitter presented a ranked list of Tweets based on the standard query response. Therefore, 1) it is difficult for this PhD to conclude what particular feature or group of features were important to evoked curiosity in participants beyond some interesting comments when describing the search experience; 2) this research presents unclear pragmatic evidence of how existing IIR-systems could apply and benefit from curiosity driven design (i.e. IIR-systems based on query-response paradigm). Although the curiosity driven design section and results sections in chapter 3 can provide some insights and general ideas about how to map curiosity arousal mechanism in a standard IIR-system.
- (b) **The Twitter's familiarity factor.** The possible argument that participants stayed longer on Ambiecities because it was less familiar and therefore more interesting than Twitter, which most participants were fairly familiar with. Also, some Twitter participants rated the search experience as good despite the fact they did not explore for more than 1 minute, perhaps showing a possible previous favouritism with the social media website. Notwithstanding, the experiment designed was between subjects to avoid people guessing the new interface or system and behaving in favour of the new one.
- (c) **The simulated scenarios were not 'realistic'.** Although it is quite easy to imagine a scenario where one is waiting for someone and then becomes bored by the situation, and the idea of having touch screens in a smart city scenario, there were a few details in the scenario that someone could categorize as none 'realistic'. First, the user supposedly sees "big public touch screen with an app" which could be argued as not everyday life for many people and lucky discovery. Second, the touch screen was public, people might display a different behaviour when they are in a public environment. Specially, if they were accessing their private Twitter account.
- (d) **Curiosity types.** There are various forms of curiosity: perceptual, sensory, interpersonal, and epistemic curiosity [37, 41, 20, 8]. Although the user engagement scale (UES) of O'Brien & Toms enquires about curiosity in the novelty sub-scale [26], the wording of

the questions is too general and it makes difficult to report the specific type of curiosity motivating exploratory behaviour.

(e) **The experiment did not assess curiosity as personality trait in the IIR process.**

Despite the fact that curiosity can be evoked, Spielberger & Starr [39] theories that curiosity state individual differences could influence the arousal of other emotional states and behaviours (i.e. becoming curious). Litman & Jimerson reported that individuals vary not only in their predisposition to become curious, but also in their disposition to experience curiosity as a feeling of either “deprivation”(CFD) or “interest”(CFI) [19]. Thus, it will be important to examine how individual characteristics in CFD and CFI differentially predict the user’s engagement when interacting with an IIR-system. By measuring CFI and CFD, this research might find evidence about the moderator role of curiosity between users engagement, the different design features in the curiosity driven IIR-system and casual-leisure search scenarios.

(f) **Personal interest (PI) variable was not assessed.**

In a previous user study, O’Brien found newsreaders engaged with news stories they could relate to on a personal or societal level [118]. During their investigation of how ‘catchiness’ (saliency) of relevant information affects user’s engagement, McCay-Peet et al.[121] reported that participants’ interest in a topic was good predictor of focused attention, which in turn was a good predictor of positive affect. They reported that users were often distracted by the content they found personally interesting even when the topic of the content was not relevant to their assign task. According to some qualitative comments, Ambiecities allow participants to expressed their personal interest in terms of a geographical query pointing to a specific location or place such as the users home town, current location, holiday destination, etc. where they felt an emotional connection. However, the Ambiecities application did not perform any ranking or filtering based on the content or the topic of the Tweets, and neither the PI variable was measured or manipulated during the experiment.

(g) **The effect of type and quality of content, when evoking curiosity, was not measured.**

As already mentioned, Ambiecities did not retrieve and rank Twitter content based on the quality of the Tweet, or other social signals. However studies like McCay-Peet et al.[121] could suggest the importance of type of media (e.g. photos, video, text) and quality content in the user’s engagement

4.1.3 Motivation

After gathering both the lessons learned and the identified limitations, this research found enough motivation to conduct another user study to gain more empirical evidence supporting the main argument of this thesis, i.e. IIR systems in leisure scenarios benefit from applying curiosity arousal principles into their information interaction and search experience design. The following points explained the general guidelines for the second user study taken by this research, and their rational:

- (a) **‘Cold’ vs ‘Curious’ query-response for social media news search.** This research decided to design and develop an IIR-system for news reading based on social media streams of Twitter. The IIR-system would have two flavours: ‘cold’ (standard) query-response and ‘curious’ query-response. By designing both a baseline system and the ‘curious’ system, this research will removed the Twitter’s bias and familiarity factors in the participants’ psychometric questionnaires, as pointed in the limitations section. Therefore, this research will be able to compare and gather more evidence of how existing IIR-systems based on query-response interaction paradigm could benefit from mapping curiosity arousal principles into their design, and how to do mapping. This example of application of curiosity arousal principles into the query-response interaction ought to maximise the impact of this work in the field because as described in chapter 2, query-response is the standard paradigm when designing the search experience of IIR-system.
- (b) **Social media news search scenario** This research selected the vertical of news because 1) social media has been used extensively to share news, 2) other researchers have used news searching and reading to evaluate user engagement using simulated work scenarios [117]. Thus, it will be very interesting to compare UES ratings between work and leisure news reading and searching.
- (c) **Interactivity.** From the qualitative psychometric UES ratings (e.g. perceived usability sub-scale) and participants response to the open questions about the UX, this research have collected evidence of the key role of interactivity features (such as the continuous flow of markers in a given area of the Ambicities map, represented by a moving red marker) on evoking curiosity and providing an enticing search experience. Information systems are classified as interactive if users are allow to actively control within certain technological and temporal boundaries what content (information), and form they would like to access via the system. However, this research did not measure the interactivity directly in any of the open questions, nor used previous research on the concept of interactivity to create

a reliable and operationalize measurement. Although interactivity is pervasive concept when designing an engaging information system, there is not an agreed definition. An initial definition, provided by Steuers (1992) [122], is “the extent to which users can participate in modifying the form and content of the mediated environment in real time”. Liu & Shrum (2002) [123] view interactivity as the response to “the hardwired opportunity of interactivity brought by the digital environment. McMillan & Hwang (2002) [124] adopting experiential approach, defined interactivity as “a psychological state experienced by a site user during his or her interaction with the website. Their concept of “perceived interactivity” explains that there is not always a relationship between the interactive features of the website and the consumers appreciation of interactivity. For instance, some users are discourage by high levels of interactivity because the UX is cognitively demanding.

McMillan & Hwang (2002) [124] used three overlapping dimensions that are central to interactivity: perceived user control, direction of communication, and responsiveness (time). Therefore, perceived interactivity can be operationalized as the level to which the user perceives that the interaction or communication is two-way, controllable, and responsive to their actions. Following their framework, this research has adopted their definition of interactivity to create questionnaire to evaluate the ‘curious’ interactive features in the next user studies, and provide more evidence of the benefits of designing the information interaction with curiosity arousal principles.

- (d) **Curiosity as personality trait in the searching process.** Following Litman & Jimer-son curiosity trait definition as a feeling of either “deprivation”(CFD) or “interest”(CFI) [19], this research will study the relationship between the degree of user engagement and exploration with a given curiosity trait.
- (e) **Personal interest independent variable.** The news reading application had three main types of content: general, sports, and finance. Therefore, before presenting the application and the search scenario, this research could ask the future participants to choose which topic (among the three options) they perceived most and least personally interesting when they are reading news in their leisure time. Then, this research would be able to see how does personal interest variable affects user’s engagement in both the standard query-response paradigm and the ‘curious’ query-response.

The experimental design of ‘Cold’ vs ‘Curious’ in different news verticals had the intention of bringing evidence of whether type of content (e.g. general, sports, finance), novelty of

the content (e.g. last hour, last 24 hours, last week, last month) or novelty in the interface was the main driving force of the exploration.

Again to summarize, the purpose of the study was to bring more evidence to demonstrate the benefits of applying curiosity into an IIR-system (i.e. more user engagement, exploration beyond ‘normal’ and an UX with more cognitive absorption during the information searching process) by using psychometric questionnaires and interaction data to evaluate two similar social media news IIR-systems, one design using the standard query-response interaction paradigm and other using curiosity design arousal principles such as uncertainty, partial exposure and surprise. In other words, this second experiment has the goal of bringing more empirical data to answer the first two research questions (RQ1. Does curiosity-driven search design increase users’ engagement during casual-leisure search?; RQ2. Does curiosity-driven IIR-system increase users exploration during casual-leisure search?). The study also analyzed the role of curiosity as personality trait, perceived interactivity and personal interest in the searching process.

4.2 Newslabels. A ‘curious’ news search application based on social media

The following section relates the construction of a news search application based on Twitter Streaming API, and how this research applied curiosity arousal principles to the query-response information interaction paradigm through curious push notifications or search notifications. This part also describes the software design and architecture of the news search prototype.

4.2.1 Requirements

The software requirements describe in this section were defined as consequence of the user’s feedback during the first experiment and the idea of designing news search application based on social media news streams. The high level primary requirements of the news search application were:

1. **Almost real-time indexing:** The system must facilitate almost real-time indexing of an incoming Twitter Stream based on a seed-list of different curated news sources. Each different curated seed-list of news sources should be grouped by three news’ topics(i.e. General, Finance and Sports). For example, @BBCNews, @CNN ought to be in the general seedlist whereas @BBCSport, @SportsCenter, @espn should go in the sports seed-list.

2. **Search Service:** The system must allow standard query-response information interaction over the index data (i.e. a pull based interactivity). The search service should retrieve the results taking into account social features such retweets, saves, etc.
3. **Temporal Filters:** The system must provide different temporal filters to scope a search (e.g. the query “BarcelonaFC” could be filter by last hour, last 24 hours, last week, etc.). By default the system ought to display the top 20 tweets during the last 24Hours for a given news topic.
4. **Save & Share Service:** The news reading application should allow users to save tweets, and share what they have found to be interesting with other people. Before sharing, users would need to put a title and tags describing the group of saved search results.
5. **Natural:** The system must look and feel similar to other micro-blog search systems to mitigate learning effects with a main news feed.
6. **‘Curious subsystem:** A decoupled IIR sub-system should encourage exploration and extend the information interaction of the ‘cold query response paradigm by applying curiosity arousal principles in the search experience. This involves enabling push as well as pull information interaction.
7. **Feedback and Evaluation :** The system must record user interactions to support evaluation.

The figure 4.1 shows the layout prototype of the news searching and reading web application.

4.2.2 Software design, architecture and development of Newslabels

Each requirement deserved attention at the beginning of the design stage to reduce changes and problems when integrating all the sub-systems. However, this research decided to start with requirements 1 to 5, and elaborate an architecture design of a standard (“cold”) news search application, named “Newslabels”. When the baseline system was developed, this research started working in requirements 6, 7 and defined other decoupled sub-system named “curious companion” (the following section 4.2.3 and 4.2.4 discusses in more detail the design and architecture of this software module) .

During the design and development phase, this research followed XP Programing and TDD methodologies [30]. Every requirement was divide in detail user stories both for the server side and front end application. The ‘cold’ IIR system (or baseline) was finished around April 2015.

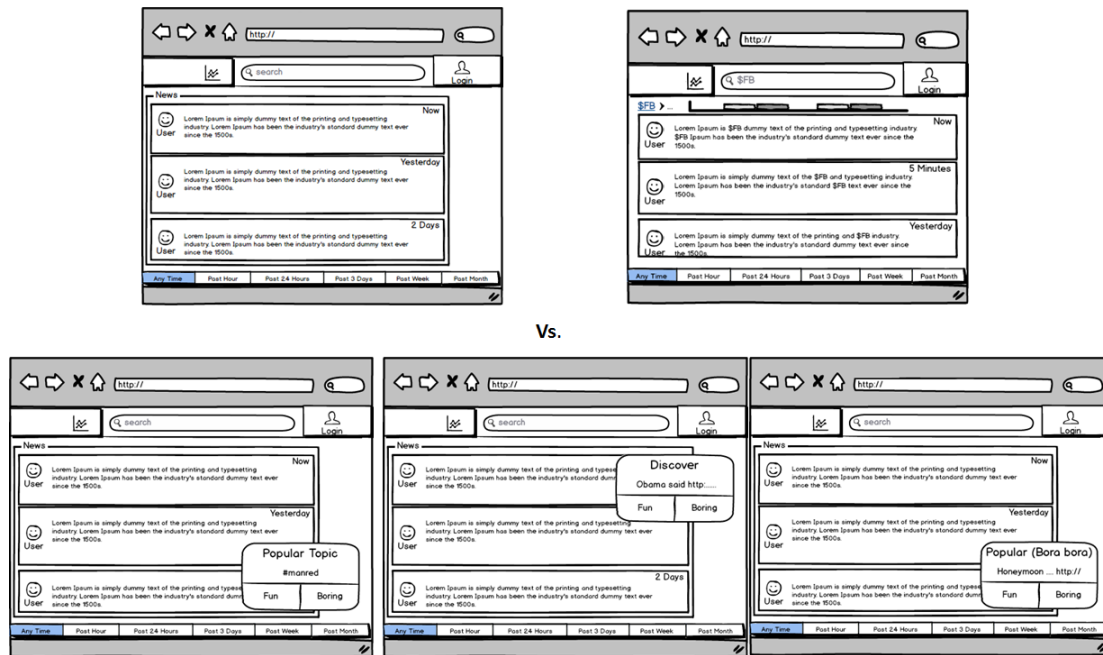


Figure 4.1: ‘Cold’ vs ‘curious’ query-response information interaction

Figure 4.2 shows the Newslabels design architecture. For the back-end component, the system has been developed using Dropwizard¹ (a Java web framework) and Lucene² (a Java full-featured text search engine library). In order to connect and get information from Twitter, this research have implemented a Java crawler with Twitter4J³. The crawler follows a list of relevant news providers and journalist.

Newslabels server-side may be broken down in a number of functional components:

1. **StreamCollector**: Connects to the Twitter streaming API using Twitter4J and filters the stream based on a seed-list. The different streams were executed as daemon jobs and restarted automatically every 15 minutes if an error was detected.
2. **Timeline Middleware**: Provides an interface for indexing, searching and filtering social media data in a Lucene index.
3. **Timeline micro-services**: Using Dropwizard and a web proxy server, several instances of the Timeline Middleware are created and wrapped with RESTful API layer.

The Single Page App (SPA) followed the Ember⁴ JavaScript framework and Twitter Bootstrap CSS⁵. These are the main components:

¹<http://www.dropwizard.io/>

²<https://lucene.apache.org/core/>

³<http://twitter4j.org/en/index.html>

⁴<http://emberjs.com/>

⁵<http://getbootstrap.com/>

1. **Adapter & Models:** Define the model of the web app such as NewsItem, Category, etc and the components to connect the SPA with the Newslabels server side API (supporting pull interaction model) and the browser local storage.
2. **Controllers:** Software components in charged of the orchestration and integration of the diverse features in SPA
3. **Web Components:** re-usable, and self contained pieces of software in the web app. For example, the Timeline chart using the D3.js⁶ library and handling the chart rendering, or the principal news feed component.

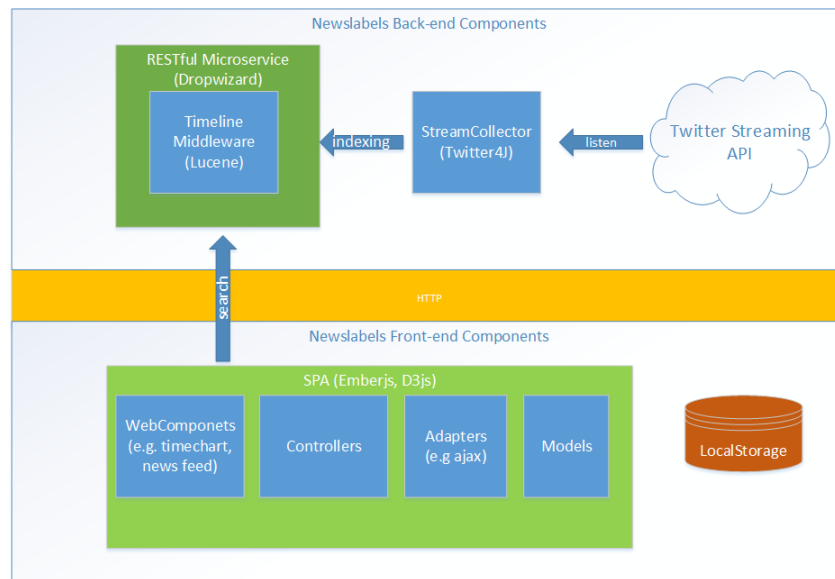


Figure 4.2: Architecture Newslabels.com

Figure 4.3 shows a screen-shot of the graphical user interface. The layout followed the previously shown prototype mockups in Figure 4.1. The baseline interface has basically three areas: the header menu bar, the timeline facet, time filter footer menu in the bottom of the page and the list of results. The header bar is made of an brand icon, a search box text field, search button and the button to go to the sharing stories section based on the tweets the user has gathered during the search session.

The timeline facet is a time-series chart showing the frequency of a query term or the total requested crawled Tweets during a selected range of time (e.g. a week, in the past 6 hours). Figure 4.4. illustrates the top news under the query ‘barcelona’ the 22th of April 2015. Figure 4.5 shows the Timeline facet of the past 20 days for the ‘barcelona’ query. Thus, the timeline help the search to get a glimpse that something has happen the day before and one week ago (i.e.

⁶<http://d3js.org/>

the champions league matches between PSG and Barcelona where Barcelona had the victory and moved to semifinals).

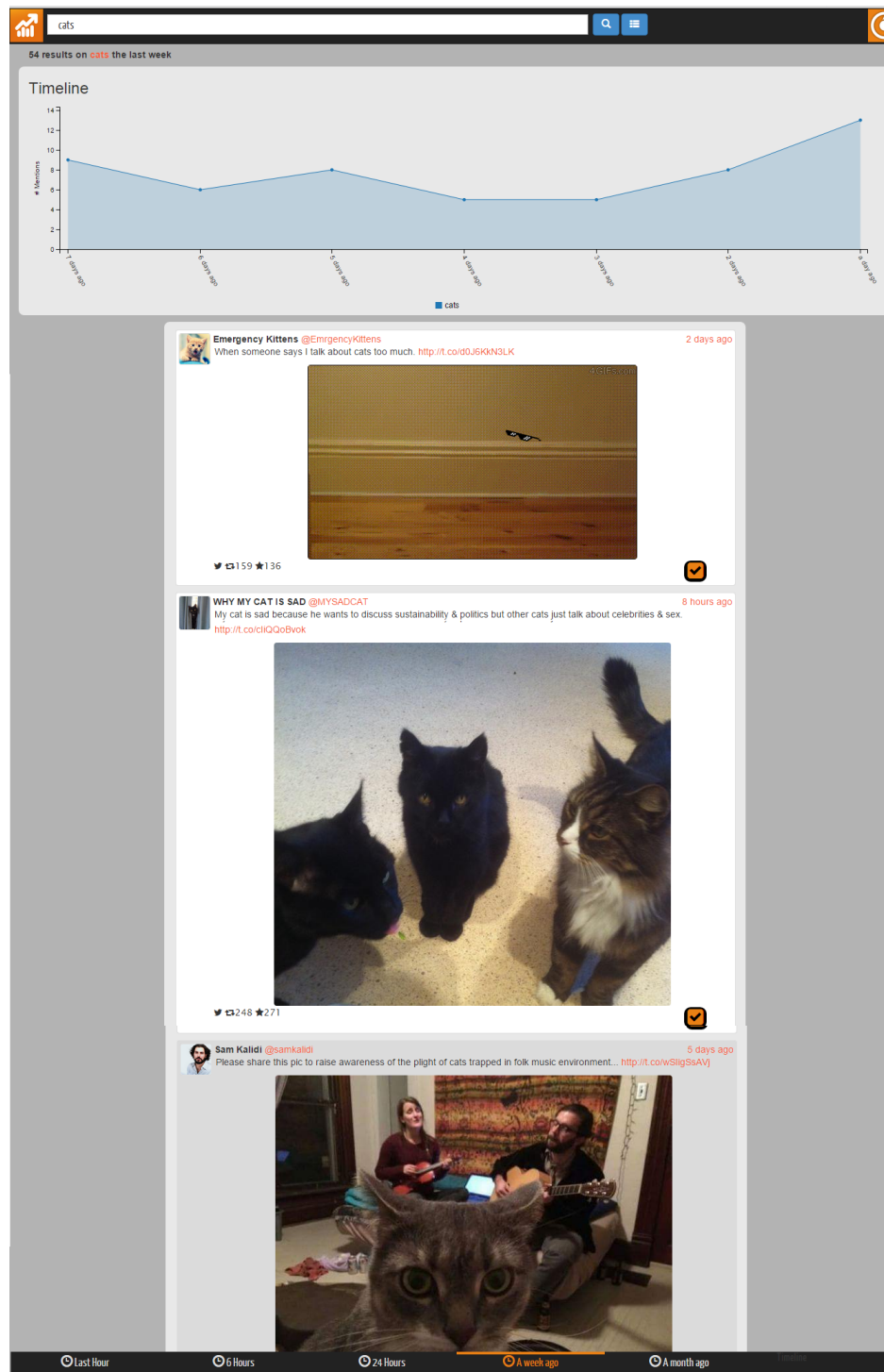


Figure 4.3: Newslabels Prototype. Standard graphical layout. Results for the query “cat” using the filter “week ago”

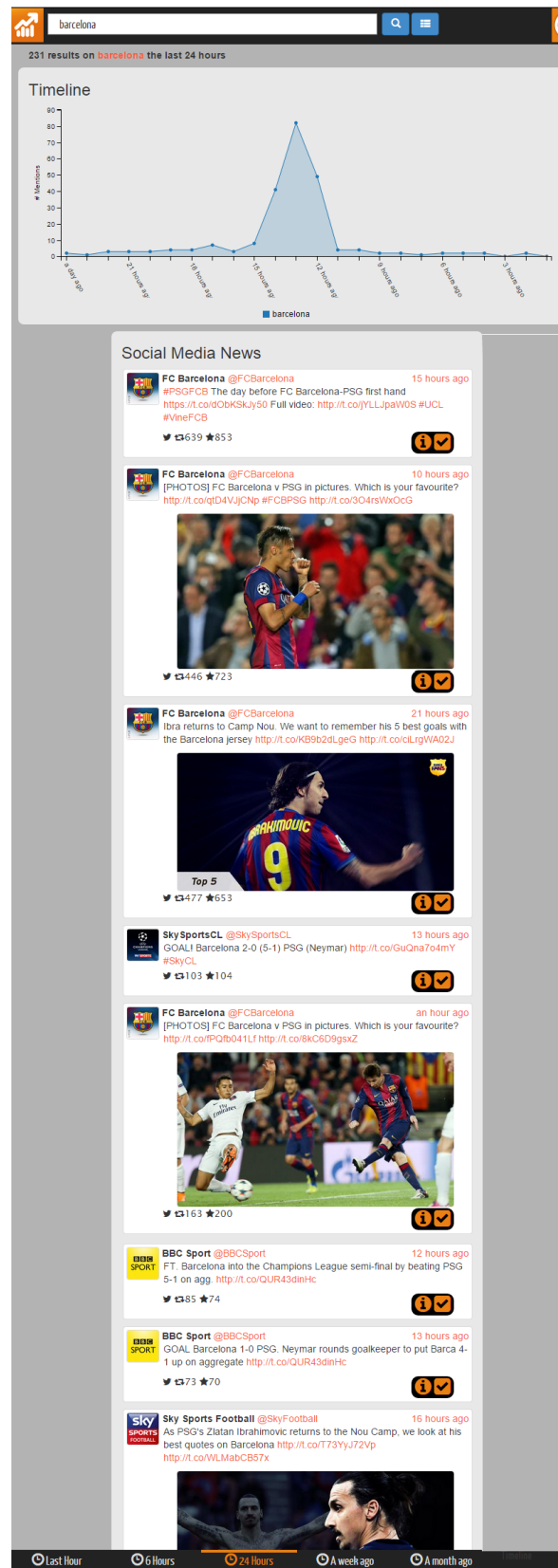


Figure 4.4: Newslabels Prototype. Results for the query “barcelona” using the filter “24 hours ago” the 22th of April 2015

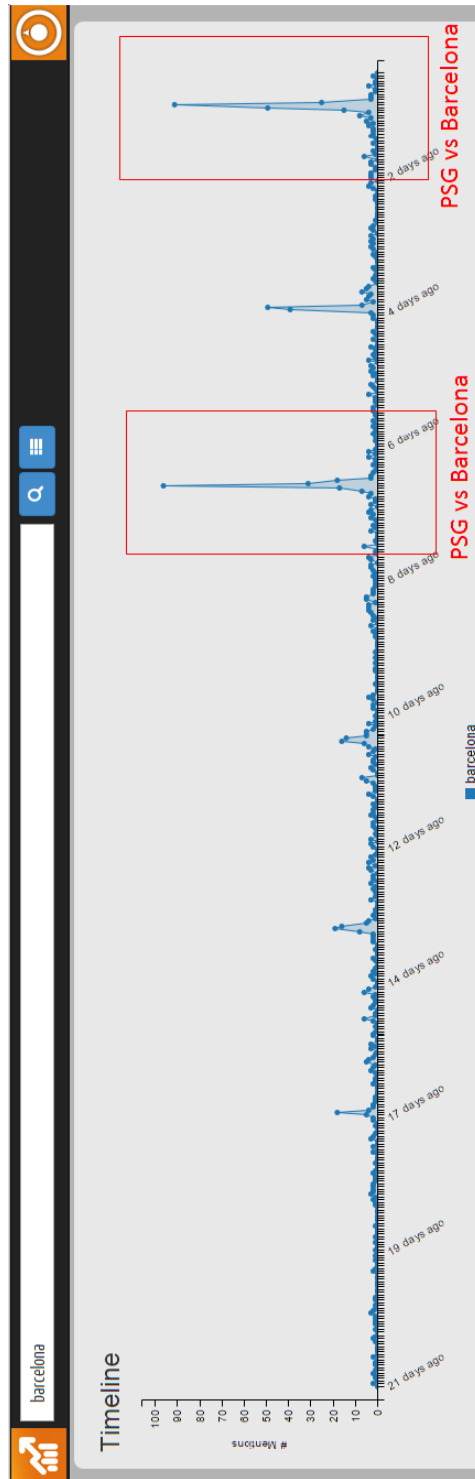


Figure 4.5: Newslabels Prototype. Timeline facet for the query “barcelona” using the filter “20 days ago” the 22th of April 2015

The time filter footer menu has four buttons that users can select to filter the information by a range of time. The filter selection and current query are notified to the searcher by modifying the CSS color of the selected time filter button and creating a search breadcrumb below the header menu bar.

The results section has the list of retrieve Tweets. The results are rank by a simple function that takes into account query textual similarity and social signals for each Tweet like number of re-tweets, and saved time. Each result in the list of results contains the user name, the label with created time-stamp, the tweet text, the social signals (e.g. re-tweets, saves), and a button to save the Tweet if the user want it. Sometimes if there are external links to other information sources providers (e.g. BBC News, Guardian), there will be another button besides the save result button and users could click on it to see the new article or video.

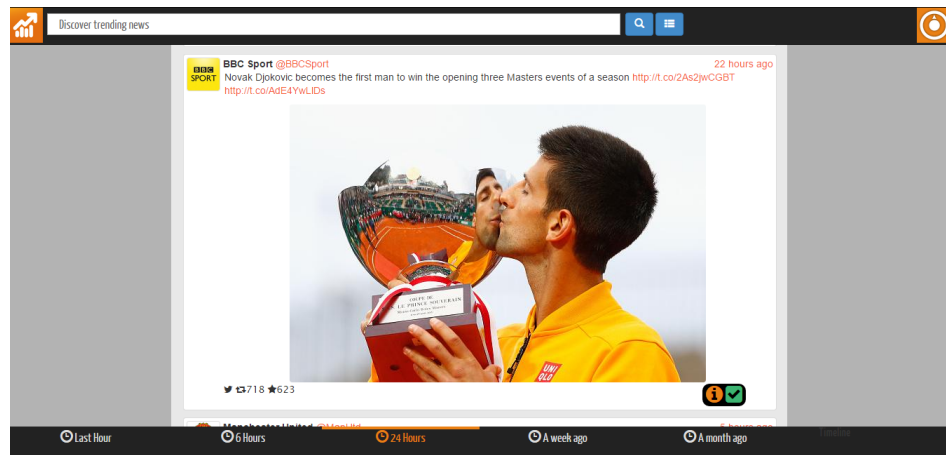
There is also the functionality to create a story based on the saved tweets, and share it on-line. Figure 4.6 outlines the different steps to create a story from your search session.

4.2.3 Mapping curiosity arousal principles into the query-response paradigm

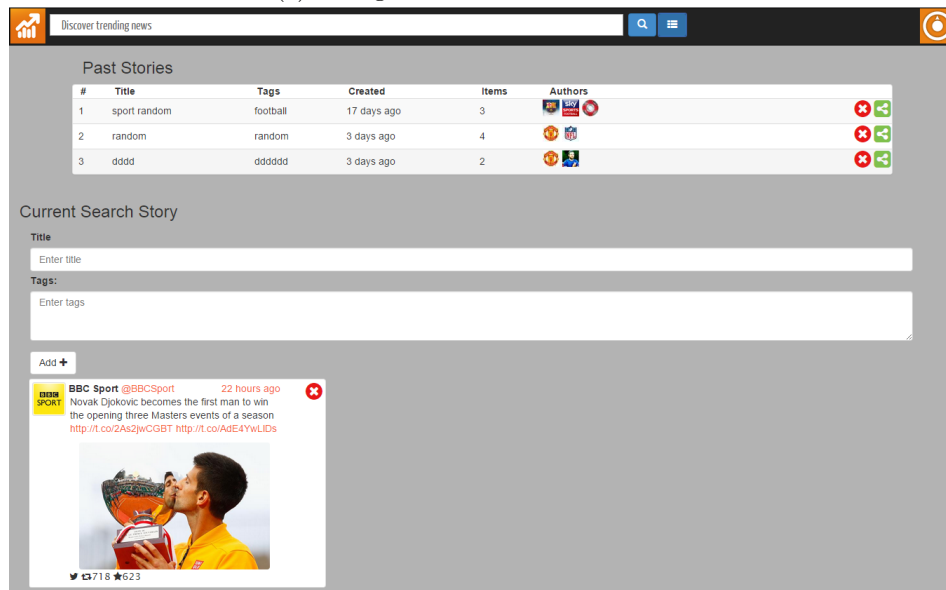
When making a public presentation, it is important to have informative material, however it is a lively, enthusiastic delivery that helps capture the attention of an audience. As already highlighted in chapter 2, search is a dialog between a user and system within a bounded context where the user is seeking ‘informative material’ and the system presents relevant information based on a given query. But in order to capture the attention and engage the user, the system should lively and enthusiastically deliver content, not just provide topical relevant content.

Following this observation, in the previous chapter, it was described how a map-based search application using push interaction and curiosity arousal principles was able to trigger more exploration and user’s engagement in a leisure social media scenario than the Twitter search service based on standard query-response paradigm. Inspired by the comments and psychometric ratings attributed to the Ambiecities ‘push’ interaction design during the previous experiment (e.g. engaged participants were happy looking new tweets pop-up and made them stay longer than they have anticipated), this research decided to mapped curiosity arousal principles into the query-response information interaction paradigm through the use of ‘curious search notifications that can be push from the IIR system or special sub-module of the IIR system. Thus, creating a search experience with both pull and push interaction supported by modern HTML5 Web-Socket standard for web browsers.

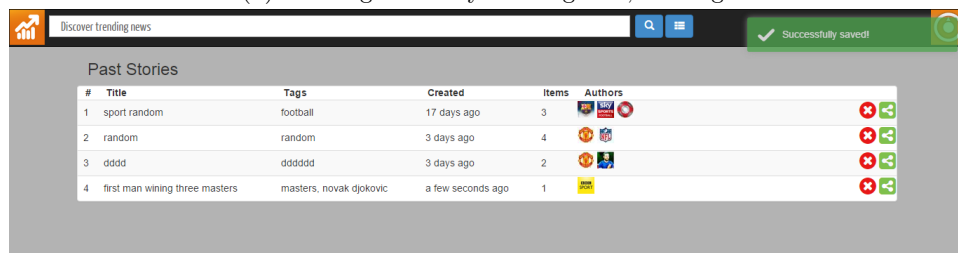
The idea of notification (also known as flash messages, or pop-ups) is not new in the web.



(a) Saving relevant news from Twitter



(b) Defining the Story. Adding title, and tags



(c) Adding the story, and then users decide if the want to share it

Figure 4.6: Sharing functionality

Although push notifications have a dubious reputation as annoying distractions, and many information interaction designers may be reluctant to create an application with pop-ups and popover ads. There is lots of evidence that popups increase online conversions in everything from newsletter signups to sales. There are businesses that used with great success without

alienating either existing or potential customers⁷.

Many on-line and social web applications use this feature to trigger more user's engagement, enhance usability (e.g. a brief notification message confirming a successful action done by the user) and influence decision making. For example, Facebook.com⁸ will send a push notification when a friend has liked a picture of a photo that a user has uploaded while the user is looking to their news feed. Booking.com⁹ will show push notifications about other users booking in a hotel display in the list of results, the amount of rooms left, or last booking (e.g. when and where was made the booking) for a particular hotel while the user is searching and browsing in their website (see Figure 4.7). This travel website provides social proof of the popularity for the area which users are looking at or a sense of urgency by showing how few hotel rooms are left because they want to influence their visitors decision making process and encourage them to act quickly.

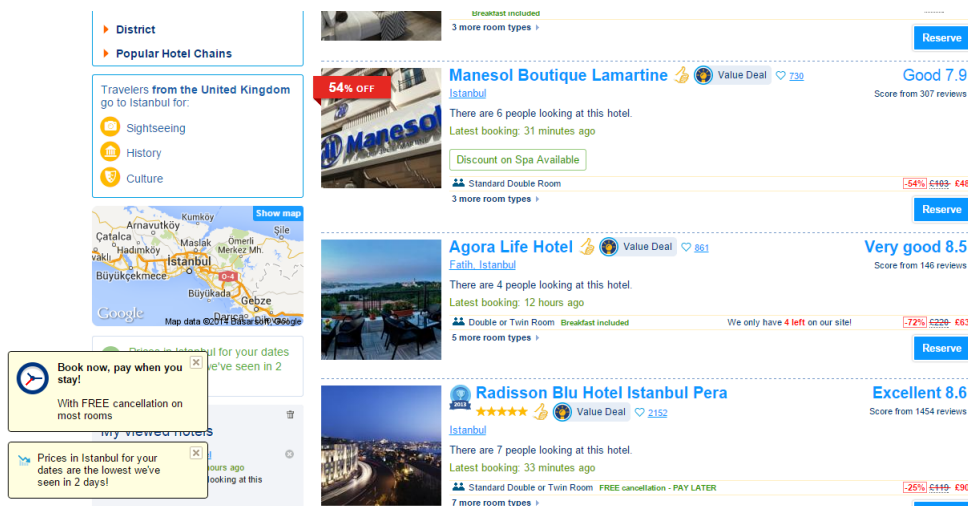


Figure 4.7: Flash notification message in Booking.com when searching a hotel for Istanbul, Turkey

Besides the standard query-response interaction for the news search prototype, this research planned the 'curious' search notification service to enable the following information interaction and search experience. Initially, the notification should appear in a particular location of the screen as a pop-up or modal dialog (a rectangle with some text). The main places would either be at the left or right of the central news feed because the notification should not interrupt the user view of any of the news items. Then the 'curious' search notifications with a limited time span (e.g. 5 seconds) would have started showing an appealing message in the screen. When the time expires, the notification should be clear from the interface. If the user clicks on a

⁷Yieldify. <http://yieldify.com/gb/>

⁸<http://www.facebook.com/>

⁹<http://www.booking.com/>

search notification before it disappears, the current list of results is clear, and the notification is used as the query to retrieve other results. The search notification should also provide a visual indication of how long before the notification expires.

To support this information interaction, each search notification defines the following attributes:

- **Id:** A serial number that identifies the notification.
- **Type:** There could be several types of notifications. For the moment there would be only a simple notification.
- **Topic:** A general category or field (e.g. Financial, Sports, General, etc).
- **Title:** A main title describing the notification.
- **Message:** The text description of the notification.
- **Keywords:** The assign query for this notification. When the notification is clicked, this is the query send to the server.
- **Location:** A coordinate where the notification is going to appear in the screen (e.g. top-left, bottom-right, etc)
- **Time Out:** The total time in milliseconds that the notification is going to be visible in the screen
- **Style:** The style and look and feel of the notification (e.g. CSS class)

A search notification service can be design with many different features that could potentially evoked curiosity by manipulating each of the attributes of a given search notification. This research has applied uncertainty, surprise and partial exposure concepts derived from curiosity arousal theory to the search notification service. In a future work, other types of curiosity such interpersonal curiosity could be used to trigger more engagement and exploration on collaborative search tools using search notifications, in a similar way that Booking.com encourage the purchase of hotel with notifications and messages based on the buying behaviour of other users.

For example, the search notifications could be generated according to the user’s context and analysing the crawled Twitter streams for the news search scenario. In the previous experiment, the curious driven map based system showed tweets trigger by the user’s spatio-temporal context (e.g. a tweet of something happening nearby right now). For the news scenario, the curious search notification can be based on a topic of growing popularity in a given spatio-temporal

and social context. Or notifications about other searchers behaviour (e.g. in the past hour five people has search for “apple watch”).

However, for the first version of the curious companion sub-module and for the purpose of the experiment, the notifications could be manually created based on relevant news because what this research wants to evaluate and test is the effect of surprising “push” search notifications in the information interaction over the search behaviour and user’s search experience. In other words, this research assumes that relevant and better search notifications, than static manually created notifications, would be generated using recommendation systems based on collaborative filtering or content based state of the art algorithms following social insights to decide what search notifications to show. The curious companion source code in Github ¹⁰ contains the total of notifications defined before the start of the experimental setup based on the headlines of news sources like BBC¹¹, CNN¹², ESPN¹³, Marca¹⁴ and Financial Times UK¹⁵.

Also the layout and the content of the search notification could be edited to partially display information and evoked curiosity (e.g. “Shakira’s has been tweeted 100 times during this minute. Do you want to see why?”, “UK has saved 200K times this news in the last hour” with a thumbnail of the Twitter user or a Tweet image masked with an almost transparent grey color). This idea is not new for the web too. Popular social networks and e-commerce sites have used the concept of partial exposure to generate uncertainty feeling, encourage exploration and affect decision making. For instance, LinkedIn.com¹⁶ service uses a personalized glimpse of what an user could be knowing from their professional network to encourage users to upgrade their account and pay for it. They try to move users into an unknown state and evoked curiosity by giving them bits of knowledge that could be fully known, as a paid members (Figure 4.8). Nonetheless, the search notification service was only design to produce a simple textual notification and only the message of a few notifications were created using partial exposure concept.

The push notifications should ideally be delivered by the system in different key contextual moments to evoke uncertainty and surprise. Specially when the user interest is decreasing and it’s going to leave the exploration. For instance, the push notifications can be display when the mouse leaves a particular area of the interface, the user has stopped scrolling for more than a time threshold, or other user interaction (e.g. clicks, mouse movements). For example,

¹⁰<https://github.com/juanmillan85/curiouscompanion.git>

¹¹<http://www.bbc.co.uk/>

¹²<http://edition.cnn.com/>

¹³<http://espn.go.com/>

¹⁴<http://www.marca.com/>

¹⁵<http://www.ft.com/home/uk>

¹⁶<http://www.linkedin.com/>

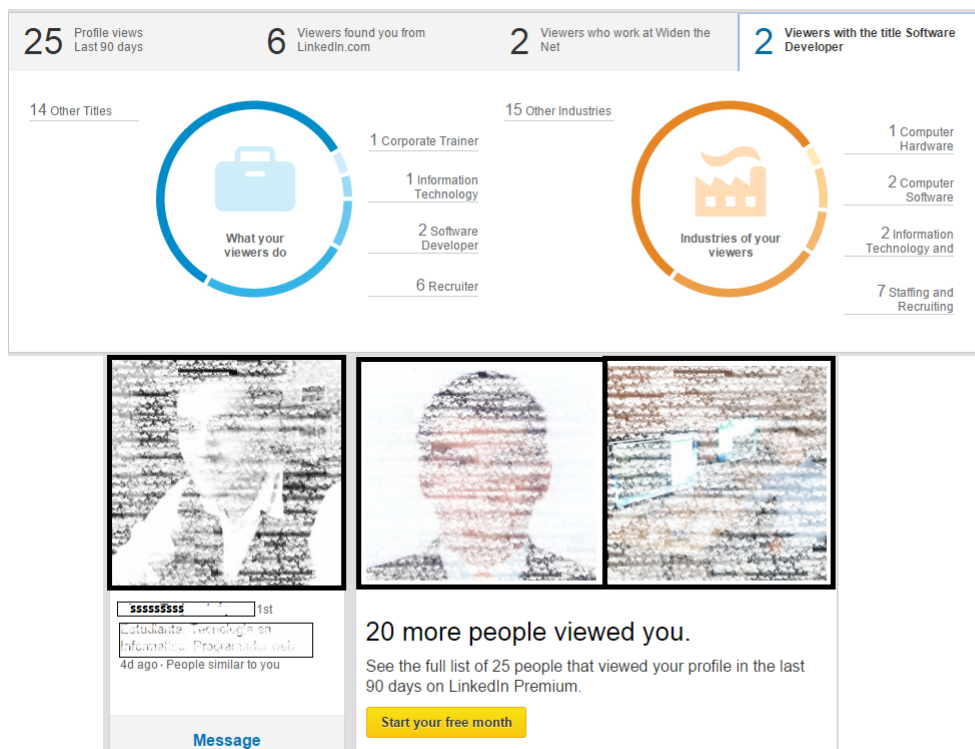


Figure 4.8: LinkedIn using partial exposure in the information interaction and interface design to evoked curiosity

Yieldify has marketing platform technology that anticipates when a user is about to leave a website without buying, and through a message or email can convert the abandoning visitors into repeating customers. Yieldify claimed that their marketing platform is able to re-engage visitors during an optimal browsing time period, before they abandon the website.

Although successful stories like Yieldify marketing platform show that tracking the optimal time of re-engagement based on implicit user's feedback and interactions, such as mouse movements, can have a powerful influence in the decision making of users. The first version of the search notification service would not track the users interaction as feedback to guess the optimal or key contextual moment to show the notification because this research assumed that using a random function to show a new notification could be enough to generate a surprising and engaging information interaction. At least, a more engaging than the one provided by the "cold" query-response paradigm.

This research planned that first version of the search notification service would push each notification following a predefined sequence with random gap interval of a minimum of 4 seconds and a maximum 44 seconds, assigning random screen locations (such as top-right or bottom-left, etc), a random time out between 3 to 8 seconds on the screen, and a random css style (e.g. red, yellow, green, and blue background color). The only deterministic function in the

notification service would be the filtering of search notifications based on the search vertical or news type that the users would be searching (e.g. if the user is looking into general news, the search service would send only search notifications for the general news topic, and not finance or sport). Therefore, none of the participants were able to predict where and when a new notification would appear manipulating both uncertainty and novelty in the information interaction.

4.2.4 ‘Curious’ companion software design and architecture

In order to match requirements 5 and 6, the curious companion module should established a two-way communication channel with the SPA, to enable both push communication from the server-side and the tracking of the user’s interaction.

Query-response information interaction for IIR-systems was designed around HTTP, the standard communicating protocol of web browsers [125]. HTTP is a request/response protocol which means that server cannot push messages to browser if there are no requests. However, HTML5 WebSocket [126] protocol brings real time communication in web browsers to a new level. Daily, new applications are designed to stay permanently connected to the web. Therefore, in a similarly to Ambicites prototype, this research decided to used the Web-Socket protocol to establish full-duplex channel of communication.

The implementation of the first ‘Curious’ companion was very simple and basic. For example, as already mentioned, the search notification service decided what search notification to push, and the values of other attributes by using a random functions rather than a complex model of artificial curiosity [111]. Although some may argued that the ‘curious’ companion ought to be named as random notifier, the module name (‘curious companion’) is based on the hypothesis that an surprising and unexpected stimulus such as push search recommendation can be perceived by an user as interesting and evoked perceptual curiosity during an individual search experience. Specially, when an user in a disengage state clicks the notification and then is rewarded with a new set of results or environment that maximises learning, and the likelihood of finding of other interesting information.

The ‘Curious’ companion defined three sub-components:

1. **WebSocketService:** This component is the gateway that wrapped the internal components of the system and exposed them using the WebSockets protocol. This modules acts as router of all incoming and outgoing messages. This component also uses a logging library to register and audit all the communication history through the web-socket channel.

2. NotificationService: The main component of the module responsible for recommending a search notification for a given topic news vertical.
3. StorageService: Allows CRUD operations over search notifications in the data storage.

In order to implement this module, this research used Akka Streams¹⁷, Spray.io¹⁸ and the Actor-based model of Akka¹⁹. Each of the components were instantiated as part of a stream pipeline or topology. Thus, each component was executed efficiently and using back-pressure, the capability of slowing down the producers message rate if the consumers cannot keep up. For the front end, the news application created a web-socket service using Socket.IO²⁰ and imported toastr.js²¹ library for rendering the notifications. The curious sub-system was finished around July 2015.

Figures 4.9, 4.10, and 4.11 show some examples of push search notifications.

Overall this section has described two systems to be evaluated in the remainder of this chapter. The greatest emphasis was placed on describing the design of ‘Curious’ Companion module and explaining the design decisions taken. Curious companion and the search notifications were designed based on the principles extracted from the earlier research and is hypothesized to help support exploration and user’s engagement when searching for fun. Newslab defined three verticals finance, general and sports.

4.3 Methodology

The section describes the summary of metrics, the experimental design, the data collection, search scenarios and the demographic description of the participants. As in the first experiment described in chapter 3, the experiment followed a user-centred approach [28] in a naturalistic scenario with mixed-methods perspective between IIR behavioural data and UX subjective self-reported variables.

4.3.1 Variables and constructs

The main conceptual constructs evaluated by this research are personal interest (PI), user’s engagement (UES), perceived interactivity (PIN), diversity (DIV), explorability (EXPO), pointers & contrasts (UNEX, uncertainty and surprise), I-type Curiosity (IC) & D-type Curiosity (DC), and Serendipity (SE).

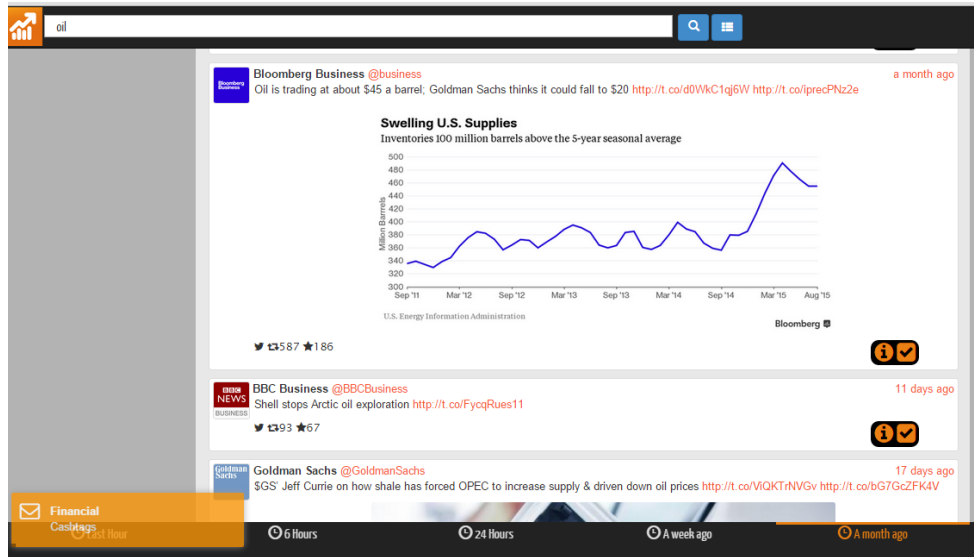
¹⁷<http://doc.akka.io/docs/akka-stream-and-http-experimental/1.0-M2/scala.html>

¹⁸<http://spray.io/>

¹⁹<http://akka.io/>

²⁰<http://socket.io/>

²¹<https://github.com/CodeSeven/toastr>



(a) Finance Vertical. Curious companion has set the cashtags search notification



(b) After clicking the cashtags notification, the query “\$AAPL \$AMZN \$GOOG” is triggered and a new set of results is presented.

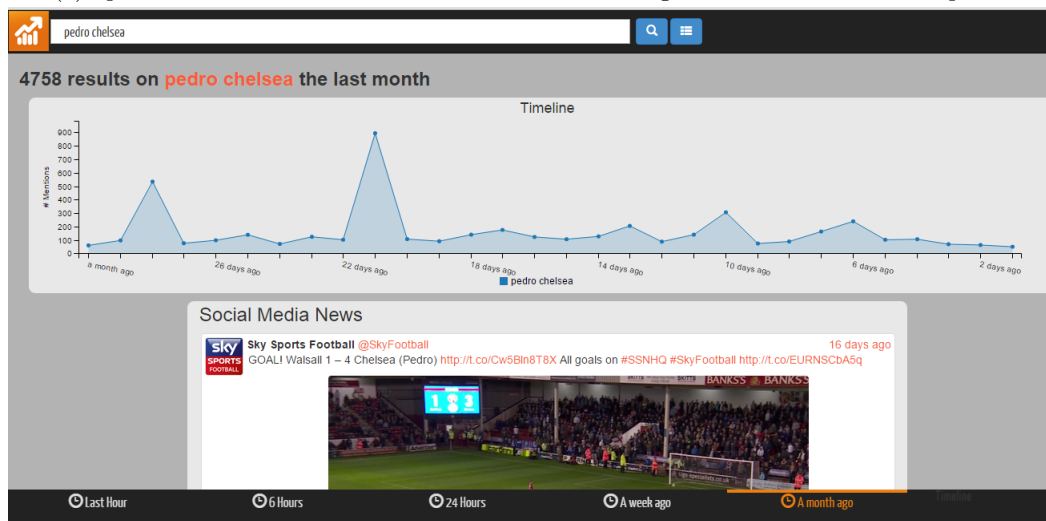
Figure 4.9: Search Notifications

PI variable was found by two studies McCay-Peet et al.[121] and O’Brien [118] to be a good predictor of focused attention, which in turn was a good predictor of positive affect. For the experiment, this research defined three news verticals: general, sports, and finance.

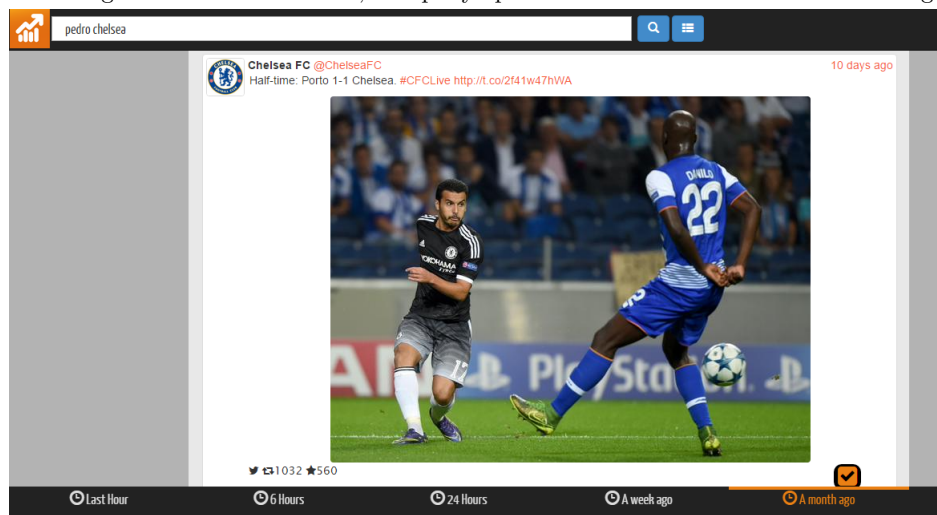
In a similar way to the previous study, user’s engagement variable was evaluated using O’Brien & Toms (UES, 31-items) [26] psychometric questionnaire grouped of six factors, Focused Attention (FA), Perceived Usability (PU), Aesthetics (AE), Endurability (EN), Novelty (NO), and Felt Involvement (FI). Participants rated all items using a seven-point Likert scale ranging from strongly disagree (1) to strongly agree (7) with a eighth option for not applicable similar to [118]. For the Novelty scale the item “My interaction with the App incited my



(a) Sports vertical. A search notification with the message “Blues’ transfers from Spain”

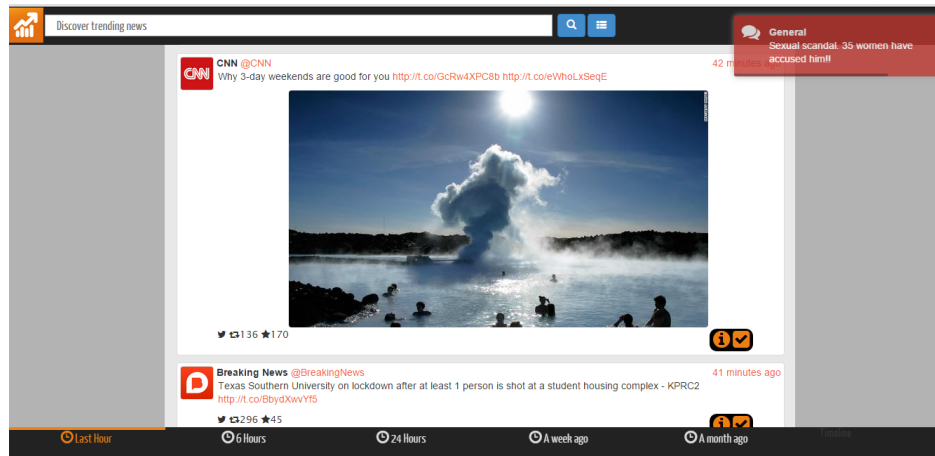


(b) After clicking the search notification, the query “pedro chelsea” and new set of results is generated

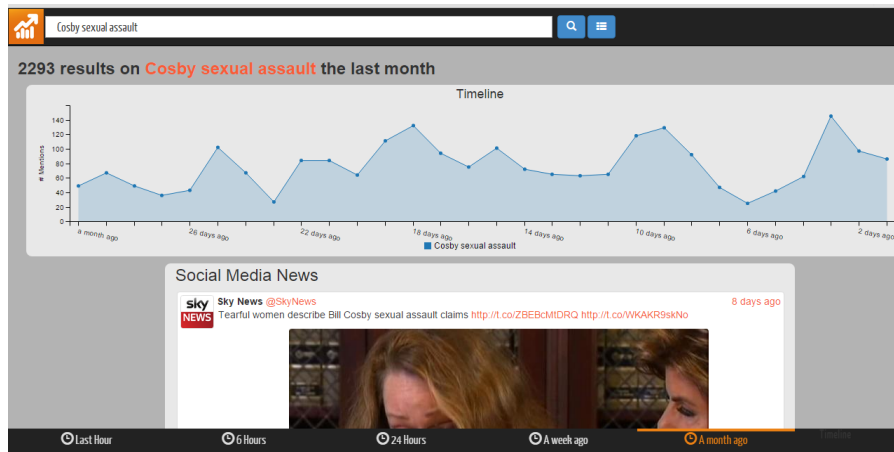


(c) Scrolling down for more news

Figure 4.10: Search Notifications



(a) General vertical. A partial exposure search notification about a sex scandal



(b) General vertical. After clicking the search notification, the query “Cosby Sexual Assault”

Figure 4.11: Search Notifications

curiosity” was added based on what user’s reported in the previous study regarding to what triggered their curiosity and for the Aesthetics scale, the one of items “The screen layout of the App appealed to my visual senses” was deleted because this research wanted to avoid repetition in the questionnaire. Serendipity was assessed with 2-items, “I found unexpected valuable information”, “The information presented has moved me to do something I have never thought doing”. Table 4.1 shows the total of 31 UES items, and 2 serendipity items, not 26-items UES as done in the previous experiment or 28-items in WikiSearch study by O’Brien & Toms, 2013 [24]. The wording of the UES items was also adapted to news search scenarios. No more adaptations were done for the purposes of maintaining the validity of the scale to measure the UX.

Perceived interactivity, explorability, diversity and unexpected pointers & contrasts constructs were evaluated by participants using a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5). The perceived interactivity was evaluated with 3-times using

Questions	UES Sub-scale
The time I spent searching just slipped away.	FA
I felt in control of the searching experience.	PU
I was really drawn into my searching tasks.	FI
I felt involved in the searching tasks	FI
Searching using the App was worthwhile	EN
I felt frustrated while using the App	PU
I was absorbed in my searching task.	FA
Using the App was mentally taxing.	PU
My search experience was fun	FI
I could not do some of the things I needed to do using the App.	PU
I consider my search experience a success.	EN
This search experience did not work out the way I had planned.	EN
I would recommend the App to my friends and family.	EN
I was so involved in my searching task that I lost track of time.	FA
I felt annoyed with using the App.	PU
This search experience was demanding.	PU
The App interface is attractive.	AE
My search experience was rewarding.	EN
I felt discouraged while using the App.	PU
The App interface is aesthetically appealing.	AE
I found the App confusing to use.	PU
I felt interested in my searching tasks.	NO
The content of the App incited my curiosity.	NO
I continued to used the application out of curiosity.	NO
The screen layout of the App appealed to my visual senses.	AE
I blocked out things around me when I was using the App.	FA
I lost myself in this searching experience.	FA
When I was using the system, I lost track of the world around me.	FA
My interaction with App incited my cuiosity.	NO
I liked the graphics and images used on this App.	AE
I forgot about my immediate surroundings while searching in the App.	FA
I found unexpected valuable information.	SE
The information presented has moved me to do something I have never thought doing.	SE

Table 4.1: User Engagement Scale Questionnaire and Serendipity items for the casual-leisure news scenario

McMillan & Hwang (2002) [124]. The addition of perceived interactivity is supported by the results and comments of the previous user study regarding the role of interactivity in the user's exploration and engagement.

The EXPO, DIV and UNEX metrics were taken from Björneborn (2008) [107] dimensions in the physical library that supports users' divergent information behaviour, and affect the likelihood for serendipity when users find materials and information not planned for. EXPO joined Björneborn unhampered access, display, and explorability dimensions to allow users to judge their exploration in the search news prototype. UNEX metric combined pointers and contrast dimensions which resemble a relationship with the curiosity constructs of uncertainty and surprise. Thus, pointers and contrast dimensions could be a good subjective metric to eval-

uate how surprising, unexpected and “eye tempting” were the notifications for the information interaction. The DIV scale evaluated the variety of topics, genres, resources, activities, sections in the App and remained unchanged. Table 4.2 shows a summary of the items.

Questions	Construct
The App presented content in ways that invited me to explore across topics.	DIV
Unexpected words and phrases caught my eye.	UNEX
Unexpected words and phrases sparked my thinking.	UNEX
With the App, I explored many topics that normally I do not examine	DIV
The search App enables concurrent communication	PIN
Unexpected visual features of the App caught my eye	UNEX
I was able to examine a variety of topics.	DIV
I wanted to click on things to see where they would take me.	EXPO
The search App is interactive	PIN
The App encouraged me to browse and explore.	EXPO
I was interested in the content of the App.	EXPO
I wanted to find out more about the topics that I encountered on the App.	EXPO
The search App enables two-way communication.	PIN

Table 4.2: Questionnaire for PIN, DIV, UNEX and EXPO constructs

I-type Curiosity (IC) and D-type Curiosity (DC) concepts were assessed using an 8-item questionnaire by Litman et al. [19, 20, 8]. Using this questionnaire, curiosity was evaluated as a personality trait. Table 4.3 shows the 8-item questionnaire.

Questions	Curiosity Type
I enjoy exploring new ideas.	I-typeEC
I find it fascinating to learn new information.	I-typeEC
I enjoy learning about subjects that are unfamiliar to me.	I-typeEC
I spend hours on a single problem because I can’t rest without answer	D-typeEC
I brood for a long time to solve problem	D-typeEC
Conceptual problems keep me awake thinking	D-typeEC
I usually frustated if I can’t figure out problem, so I work harder.	D-typeEC
I work like a fiend at problems that I feel must be solved.	D-typeEC

Table 4.3: Litman et al. questionnaire for I-type Curiosity (IC) and D-type Curiosity (DC)

The search behaviour metrics follow classic IIR summarized by Kelly’s (2009) [28]. The interaction data such as Total Dwelling Time in task (TDT), Guessed Total Dwelling Time in task (GTDT), number of queries issued, number of documents viewed, query length, number of saved documents, number of click documents, and number of share documents. The Table 4.4 shows the summary of the variables and concepts to be measured during this user studies.

Construct (Abbreviation)	No. Items	Operational Definition	Reference
User Engagement Scale (UES)	31	the quality of user experience that describes a positive human-computer interaction	O'Brien & Toms (2008, 2013) [26, 24]
Perceived Interactivity (PIN)	3	the psychological state experienced by a site user during his or her interaction with the website	McMillan & Hwang (2002) [124]
I-type Curiosity (IC) and D-type Curiosity (DC)	8	Curiosity is the intrinsic desire for new information that will stimulate interest or reduce uncertainty	Litman & Jimmerson (2004) [19]
Personal Interest (PI)	4	General News, Sports, Finance, Fashion.	O'Brien (2011) [118], McCay-Peet et al. (2012) [121]
Total Dwelling Time (TDT)	1	Total time in a given task	Kelly (2009) [28]
Serendipity dimensions	13	Explorability (EXPO), Diversity (DIV) and unexpected pointers & contrasts (UNEX)	Björneborn (2008) [107]
Guessed Total Dwelling Time (GTDT)	1	Total time guess after a given task	Kelly (2009) [28]
Interaction IIR	5	Number of queries issued, number of documents viewed, query length, number of saved documents, number of click documents, number of share documents	Kelly (2009) [28]

Table 4.4: Variables and metrics summary based on previous research

4.3.2 Search scenarios

Instead of creating a leisure search scenario from scratch again, as done in the previous user study, this research has decided to based the simulated search scenario on Borlund' approached [29], Moshfeghi & Jose's entertainment tasks [102] , and The Social Book Search Lab Interactive Track of CLEF 2015 ²².

These are the two scenarios for this user study:

TASK A

'Imagine you are waiting to meet a friend in a coffee shop or pub or the airport or your office. While waiting, you come across this website and explore it looking for any "news" that you find interesting, or engaging or relevant... Explore anything you wish until you are completely and utterly bored. When you find something interesting save it by clicking the tick button. (Click here). Please describe your search experience.'

TASK B

'Imagine your boyfriend/girlfriend/best-friend is travelling and communication ac-

²² <http://social-book-search.humanities.uva.nl/#/interactive> (Access April 2015)

cess is very limited. It is now a few days since he/she has gone and you are missing him/her very much. You are feeling very sad and in order to change your mood, you have decided to search some interesting news. Find as many relevant news as possible that makes you feel happy using this news reading application, and save it by clicking the tick button (Click here). Please describe your search experience'

The first scenario was taken from The Social Book Search Lab Interactive Track of CLEF and it was adapted to the news reading scenario (e.g. replace books for news). The second scenario was taken from Moshfeghi & Jose's Entertainment by adjusting Mood (ENM) [102] and the wording was partially change because the original scenarios were created in the context of video retrieval. According to Moshfeghi & Jose, the first scenario could be characterised as Entertainment by adjusting Arousal (ENA).

Although each task presented a differently story, both tasks defined a very similar open-ended leisure search scenario where each participant was challenge to do what comes natural in order to have fun and joy while interacting with IIR-system. The tasks were focus in the "experience" rather than in finding specific news, event or answering a particular question.

4.3.3 Data Collection and Procedure

SurveyGizmo was used to guide users through the experiment and store their responses to the questionnaires. The whole protocol was made available in English and Spanish. From the lessons learnt in the first user study, this research decided to conduct the study in a naturalistic environment where users would be able to explore and search for fun with a greater probability than in a laboratory setting.

Before the start of the study, this research conducted a short pilot study with 2 participants in order to test the correct functionality of the on-line designed survey. The participants were able to comprehend and perform the search scenarios without difficulties. Some questions were reworded due to some grammatical and spelling errors. The on-line link for the experiment was opened from the 18 August of 2015.

At the beginning of the experiment, an ethical consent form described the purpose of the experiment. Once participants had consented, they completed a demographic questionnaire that enquired them about their gender, age, level of education, familiarity with social media sites, and their personal interest (PI) in news reading by answering two questions: '*Which news category do you find the most interesting?*' and '*Which news category do you find the least interesting?*'. Each question about PI, users were given three possible options: General, Finance or Sports.

Next, the experimental setup followed a mixed within subject design. For the first session, subjects were randomly assigned to the ‘cold’ or ‘curious’ system, and the most or least interesting chosen topic. Leisure search scenario remained the same through the whole experiment because both of them described in essence a similar leisure setting. Each open-ended leisure search scenario contained a bottom link to the randomly assigned IIR-system. When participants clicked the link a new tab in the browser was created. Below the link they were presented with a long area section where they could write and paste their comments about the experience. During the search session, there was no minimum or maximum time. Users were recommended to use a browser such as Google Chrome, or Mozilla Firefox.

Subjects should complete two tasks scenarios using a counterbalanced design among subjects. The counter balanced design of system assignment and personal interest had the intention of showing whether content topic or system assignment was the main driving force of exploratory behaviour. The content quality and novelty remain unchanged for all task scenarios in the different news verticals because both ‘Curious’ and ‘Cold’ used the same search indexes. There were four possible branches in the experiment and each of them had an 25% of probability of being assigned to a given participant:

1. **COPI vs CUNPI:** Personal Interest with Cold System in scenario A and No Personal Interest with Curious in scenario B
2. **CUPI vs CONPI:** Personal Interest with Curious System in scenario A and No Personal Interest with Curious System in scenario B
3. **CONPI vs CUPI:** No Personal Interest with Cold System in scenario A and Personal Interest with Curious System in scenario B
4. **CUNPI vs COPI:** No Personal Interest with Curious System in scenario A and Personal Interest with Cold System in scenario B

After a participant finished their browsing session in the other tab and clicked the done button, Surveygizmo computed recorded their total dwelling time in seconds. Subjects were then asked to guessed the total time they have spent during the search session, and they were presented with a Likert scale block of questions based on the described UES 32-items, and SE 2-items [26]. All of the questions included in these questionnaire will be a forced-choice type and they are going to be randomly order. After the first block of Likert scale two open questions were required: *While using the App, did you feel the information presented was relevant?* and *While using the App, did you enjoy your exploration?*

After this section another set of Likert scale were presented to evaluate perceived interactivity, explorability, diversity, pointers and contrast constructs. Then participants should ranked the features or functionality of the App from the most (top) to the least (bottom) important

during the search task. The exposed features were interactivity, relevant content, timeline graph, time filter buttons, GUI (layout, colors), and content topic.

At the end of each search scenario evaluation, every participant assessed their perception of their completed task in terms of the difficulty of the task, the familiarity of the participant with the task, the extent to which they found the task stressful, interesting and clear. These concepts were measured by asking the following question ‘*The task we asked you to perform was [easy/ stressful/ interesting/ clear/ familiar] (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”)*’. Again, the task perception questionnaire was taken from Moshfeghi & Jose [102]. This section of task perception and the open questions were key to validate the assumption of both tasks being ‘very similar’.

Finally, after completing both the search scenarios and their evaluation, an exit questionnaire was introduced to each participant. In this questionnaire this research gathered information about the curiosity personality trait using Litman et al. [19, 20, 8] as well as the following open and multiple option questions:

- ‘Which task do you enjoy the most? Explain why. (answer: “Task 1”, “Task 2”)’
- ‘Which task was the longest? Explain why. (answer: “Task 1”, “Task 2”)’
- ‘Did you notice the Search Notifications or Dialogs? (answer: “Yes”, “No”)’
- ‘Did you click in the Search Notifications or Dialogs? (answer: “Yes”, “No”)’
- ‘What did you like about the search experience? (For example a particular feature or functionality)’
- ‘Did you learn something unexpected about a particular topic?’
- ‘Please give us any suggestions you have for improving the search experience using the previous application’

Figure 4.12 illustrates the complete protocol. Following the previous user study, user’s behavioural data tracked during the experiment was interpreted using subjective questionnaires where classic IIR metrics answer the “what” and self-reported UX metrics the “why”.

4.3.4 Participants User Study

Participants were recruited through various means such as email (e.g. the internal list of emails of City London University for students, and staff members), social media sites (e.g. Twitter, Facebook, LinkedIn). The participants accessed to the on-line study directly through

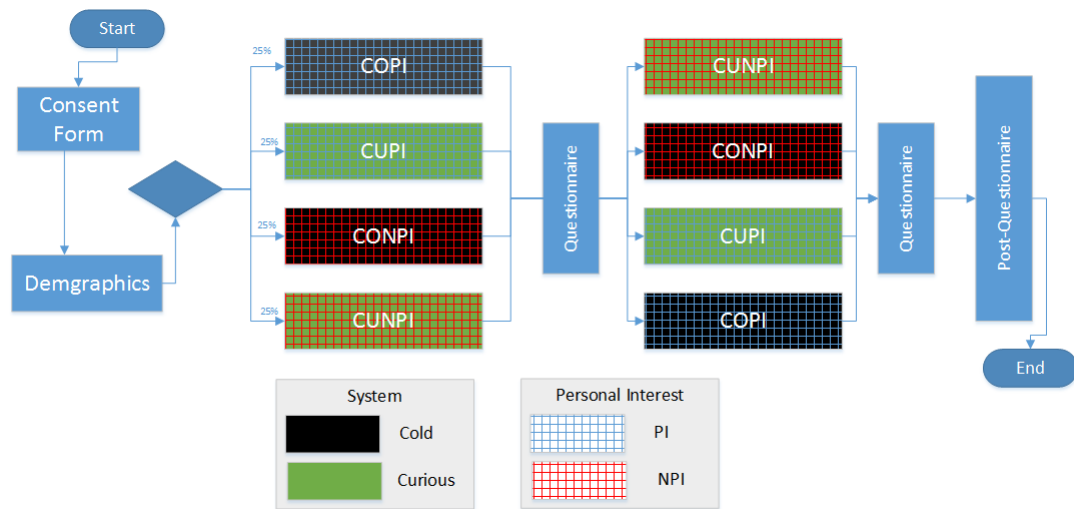


Figure 4.12: Protocol for the user study

SurveyGizmo. The duration of the experiment was expected to be between 50 minutes and 1 hour for each participant. Because the questionnaires were also provided in Spanish, all the answers to the open questions and comments were translated to English by a native Spanish speaker.

Until the 2th of October 2015, there were around 218 access to the link of the experiment. However, the sample used by this research comprised 37 participants who accessed and completed the on-line study by signing the ethical form. Four participants reported technical problems with the search applications probably because their network firewall blocked Socket IO WebSocket requests. Therefore, they were removed from the sample and only 33 participants were used in the data analysis. Nine participants were assigned to branch *COPIvsCUNPI*. The other three branches had groups of eight participants.

Although 37 participants is a small number in comparison to 218 possible participants, the study participation was voluntary. All the rest of people who accessed the SurveyGizmo link stop the user study after reaching the the first post psychometric questionnaires or the second task page. Discerning this possible response from people and to collect more data, this research created a module in the IIR-system to record and create interaction logs of many different users search actions. Therefore, the search behaviour capture and analysed by this research was not limited to 33 participants who signed the ethical forms.

There were 19 males (57.6%) and 14 females (42.4%) in the sample. Participants ranged in aged from 18 - 64, though 45.5% were between the ages of 25-34 and the second largest group was between the ages of 18-24 (30.3%). The remaining participants were between 35 and 54 (n=7, 21.2%), or over 55-64 (n=1, 7.1%). Half participants were employed (n=17, 51.5%),

33.3% students (n=11) and 15.2% self-employed.

The sample in the study were well educated: 11 participants (33.3%) held an undergraduate degree, 12 got a masters degree (36.4%), 6 had a trade or other technical college degree (18.2%), 1 pursued a Doctoral Degree (3.0%) and 3 (9.1%) reported high school diploma. Twenty reported that they spoke more than 2 languages.

Most reported daily use of social media information and familiarity with popular social networks sites through desktop computers or mobile phones. Participants were members of different social media sites: Facebook 87.9% (n=29), Twitter 81.8% (n=24), Linkedin 69.7% (n=23), Google+ 72.7% (n=24), Instagram 75.8% (n=25).

The most interesting news category selected by the majority of the 33 participants was General (n=18, 54.5%); the second place was for the Sports news (n=12, 36.4%) and Finance news (n=3, 9.1%). Eleven (91.7%) participants who selected Sports as most interesting were males. In contrast, the top least interesting category chosen in the sample was Sports (n=17, 51.5%); second, Finance (n=9, 27.3%) and third, General (n=7, 21.2). Twelve (70.6%) participants that selected Sports as the least interesting were females.

4.3.5 Data Preparation

There were two sources of data for the study. First, the psychometric questionnaires and open-ended questions were collected through Surveygizmo.com and exported to SPSS for analysis. The negatively worded UES items were reverse coded. As already explain three participants were taken out because they reported problems with the app during the study.

Next, for each item descriptive statistics and inter-item correlations were examined. There were not missing values for the 33 participants. Cronbach's alpha was employed to investigate the internal consistency of each sub-UES, PIN, UNEX, EXPO, DIV, and SER based on DeVellis' guidelines [115].

The total time of the task was recorded in seconds by the SurveyGizmo tool. Therefore, the values were converted into minutes and seconds using the format hh:mm:ss because this scale was more appropriate to read and analyse the data; the total time estimated after the search task was given by participants in minutes too. After screening the data, 33 participants remained who on average completed each task in approximately 00:09:25 (Std. Error 00:00:49).

The second source of information was the user's interaction log generated by the curious companion module. The curious companion module was able to track different users' actions. For example, clicking the search notification, saving results, making a query, the connection of a new user, and the disconnection of an user from the Newslab web applications.

WebSockets were employed as protocol of communication to track the users' actions. Every time the Newslabels app loaded in the browser, the application generated a random hash value to identified a search session as well as anonymized a given participant. In other words, every time an user opened a tab on the browser or refreshed the web application, a new session id would be created.

```

2531 09:31:49.90 ,NewParticipant,General,,,,,,,,,,,,,
2532 09:31:50.00 ,ReceivedMessage,Financial, 1041q , , Open , 2015-09-10 15:31:48 ,
2533 09:31:52.23 ,ReceivedMessage,Financial, 1041q , enable , Query , 2015-09-10 15:31:51 , , , , filter=twenty,
2534 09:38:32.23 ,ReceivedMessage,General, 1041q , , Close , 2015-09-10 15:38:21 ,
2535 10:22:40.13 ,NewParticipant,Sports,,,,,,,,,,,,,
2536 10:22:40.32 ,ReceivedMessage,Sports, 50NAH , , Open , 2015-09-10 09:22:30 ,
2537 10:22:44.32 ,ReceivedMessage,Sports, 50NAH , , disable , Query , 2015-09-10 09:22:34 , , , , filter=twenty,
2538 10:23:11.59 ,ReceivedMessage,Sports, 50NAH , , disable , Save , , , , , 10 , Twitter#641815904318128133 , " Justin Timberlake & Jinny Fallon stole the show at the US Open by doing the Single Ladies dance. , via @RaddJaffy ,http://t.co/
2539 10:23:12.07 ,ReceivedMessage,Sports, 50NAH , , disable , Unsave , , , , , 10 , Twitter#641815904318128133 , " Justin Timberlake & Jinny Fallon stole the show at the US Open by doing the Single Ladies
2540 10:23:13.39 ,ReceivedMessage,Sports, 50NAH , , disable , Save , , , , , 10 , Twitter#641815904318128133 , " Justin Timberlake & Jinny Fallon stole the show at the US Open by doing the Single Ladies dance. , via
2541 10:29:01.79 ,NewParticipant,General,,,,,,,,,,,,,
2542 10:29:02.01 ,ReceivedMessage,General, Lncvd , , Open , 2015-09-10 09:28:52 ,
2543 10:29:03.79 ,ReceivedMessage,General, Lncvd , enable , Query , 2015-09-10 09:28:54 , , , , filter=twenty,
2544 10:46:38.87 ,ReceivedMessage,General, Lncvd , enable , Click , 2015-09-10 09:46:29 , , , , , topic:"General" , "Tech Stories", "toast-top-right", "info", "Google Yahoo Facebook Microsoft Mozilla Twitter Oracle
2545 10:46:45.55 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 0 , Twitter#641850285463629824 , " The Internet is buzzing that someone might drop a new album at today's #AppleEvent. Who do you think
2546 10:46:46.31 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 1 , Twitter#641822082491650048 , " This kid "nose" a good tune when he hears it. Quite a performance. https://t.co/w8h2w0dyu", John
2547 10:46:48.23 ,ReceivedMessage,General, Lncvd , enable , Query , 2015-09-10 09:46:38 , Google Yahoo Facebook Microsoft Mozilla Twitter Oracle apps , filter=month,
2548 10:46:51.33 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 3 , Twitter#639320923036889089 , " Twitter has built a large ad business, but faces fierce competition from Facebook and Google: http://
2549 10:46:52.69 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 4 , Twitter#639563535710683138 , " Twitter, Facebook, Google and Microsoft all reported similarly disappointing diversity numbers this
2550 10:46:54.51 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 6 , Twitter#636585424673222656 , " Details behind Facebook and Twitter suspending the accounts of the Virginia shooting suspect: http://
2551 10:47:01.89 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 12 , Twitter#633942558802493440 , " How to donate your Twitter or Facebook account for World Humanitarian Day: http://t.co/DRKa2Jew08
2552 10:47:08.57 ,ReceivedMessage,General, Lncvd , enable , Save , , , , , 19 , Twitter#637800809321960080 , " How to disable video auto-play on Facebook and Twitter http://t.co/VDCL30e0 http://
2553 10:47:54.17 ,ReceivedMessage,General, Lncvd , enable , Click , 2015-09-10 09:47:44 , , , , , topic:"General" , "Ebola stories", "toast-bottom-right", "error", "ebola", "Thu Sep 10 10:47:50 EDT 2015", 7884,
2554 10:47:57.96 ,ReceivedMessage,General, Lncvd , enable , Click , 2015-09-10 09:47:48 , , , , , topic:"General" , "Sexy, funny, accessible, enigmatic. Who was she?", "toast-bottom-left", "warning", "Marilyn Monroe",
2555 10:47:58.74 ,ReceivedMessage,General, Lncvd , enable , Query , 2015-09-10 09:47:49 , , , , , filter=month,
2556 10:53:26.88 ,NewParticipant,General,,,,,,,,,,,,,
2557 10:53:27.10 ,ReceivedMessage,General, LxoyG , , Open , 2015-09-10 15:53:26 ,
2558 10:53:28.07 ,ReceivedMessage,General, LxoyG , enable , Query , 2015-09-10 15:53:27 , , , , filter=twenty,
2559 10:53:32.35 ,ReceivedMessage,General, LxoyG , , Close , 2015-09-10 15:53:31 ,
2560 10:53:32.71 ,ParticipantLeft,General,,,,,,,,,,,,,
2561 10:53:32.71 ,Terminated,,,,,,,,,,,,,
2562 10:58:57.54 ,ReceivedMessage,Sports, 50NAH , , Close , 2015-09-10 09:58:47 ,
2563 10:58:57.80 ,ParticipantLeft,Sports,,,,,,,,,,,,,
2564 10:58:59.52 ,ReceivedMessage,General, Lncvd , , Close , 2015-09-10 09:58:49 ,
2565 10:58:59.66 ,ParticipantLeft,General,,,,,,,,,,,,,
2566 10:59:57.79 ,Terminated,,,,,,,,,,,,,

```

Figure 4.13: Example of the recorded interaction logs

Figure 4.13 shows an example of interaction logs' content. The logging service recorded 4 types of events:

- NewParticipant:** This event meant that a new WebSocket connection was crated from the Newslabels wep application. Beside the timestamped, the record had information about the topic of the session
- ReceivedMessage:** This event wrapped the main domain functionality of the IIR-system or user's actions. The actions represented were Open, Close, Query, Click Search Notification(Click), Save Result(Save) and Unsaved result (Unsave). Each action was related to a topic, and the type of system (i.e. 'Curious' vs Cold; notifications enable, or disable). The Open action had information about the session id generated for the search session, and the local timestamp of the web client when the action was executed. In a similar way, Close action recorded the session id and the local timestamp when the web application was closed(e.g. user refreshes, closes the browser or browser tap. The Query action contained the keywords (i.e. text on the query), session id, local timestamp, and timefilter used in the query (e.g. filter=twenty to get only records on the last 24 hours). The Click action had the information about the local timestamp when the notification was click, message display in the notification,location in the screen (e.g toast-bottom-right means the bottom right corner of the screen), keywords in the search notification, timestamp

when the notification was created in the server, session id, type of notification (e.g. error is red, info is blue, warning is yellow and success is green), and the time duration of the notification. The Save and Unsave actions had information of the timestamp, session id, tweet id, position in the result page, text, and source of information.

- **ParticipantLeft:** This event indicated that a WebSocket connection had been closed from the Newslabels client and the topic of the session.
- **Terminated:** This event described that the curious companion agent in the server was terminated.

According to the interaction logs, there were 494 unique sessions during the period starting the 18th August of 2015 and ending the 2th October of 2015. Those ids do not necessarily imply different users, but exploration periods. Thus, one user could have multiple session ids. For instance, if the user opened two tabs of the Newslabels web application, there would be two different session ids for each them. Or if the user refreshes the page or loaded again the page, the user would get a new session id. The logger was able to track concurrent users of the Newslabels web application.

The final interaction log file had 5064 lines and the size was 542.1 KBytes. The interaction log contains data from 218 who accessed the user study, not just 33 participants used in the questionnaires. The file was created by merging the logs of three servers running both versions of the IIR-system (“Curious” vs “Cold”). Each server was responsible for a given topic (e.g General, Financial, Sports). There were three servers because each server indexed a different stream of Twitter with a different seed list of news sources. It was not possible to have multiple streams in single machine due to Twitter’s Streaming API policies (e.i. Twitter Streaming API does not allow two parallel stream connections with the same IPv4 address).

After the file was merged, this research used Apache Spark²³ as data processing engine and Apache Zeppelin²⁴ for the data visualization.

4.4 Results

Several analyses were performed on the captured data. The section presents the findings. First, the results of the psychometric questionnaires are described along with other recorded metrics such as participants’ total time on the task, total time on task estimation, perception of the task, relevance assessment and other subjective questions about the search experience. Then,

²³<http://spark.apache.org/>

²⁴<https://zeppelin.apache.org/>

this research compares within subjects and between subjects factors to establish the role of variables such as personal interest (PI), exposed system (“Cold” vs “Curious”), and curiosity as a personality trait. In the second part of this section, this research investigates the interaction logs of both IIR systems and shows interesting patterns in the data between different actions such as queries performed, search notification clicked, and the documents saved.

4.4.1 Reliability analysis of UES

Tables 4.5 and 4.6 depicted that all the UES sub-scales were highly reliable and internally consistent scales across sessions.

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	7	3.68	1.78	0.91
PU	8	5.04	1.15	0.81
FI	3	3.95	1.77	0.91
EN	5	4.30	1.48	0.87
AE	4	4.47	1.41	0.85
NO	4	4.54	1.89	0.92
UES	31	4.33	1.34	0.92

Table 4.5: Session 1 UES with 31-items

UES	N	$\bar{x}$	σ	Cronbach's Alpha
FA	7	3.65	1.65	0.95
PU	8	5.13	1.21	0.91
FI	3	4.29	1.74	0.92
EN	5	4.49	1.42	0.85
AE	4	4.67	1.58	0.91
NO	4	4.15	1.77	0.93
UES	31	4.39	1.23	0.87

Table 4.6: Session 2 UES with 31-items

Although DeVellis [115] recommends that sub-scales over 0.90 should reduce the number of items because an alpha too high may suggest that some items are redundant in the scale. Removing an item from NO, FI and FA would not have reduced the Cronbach's alpha significantly in both search session sessions.

Perceived Interactivity (PIN) [124] and the sub-scale dimensions of Björneborn [107] such as Explorability (EXPO), Diversity (DIV) and unexpected pointers & contrasts (UNEX) showed a good internal consistency as well. Tables 4.7 and 4.8 summarized the descriptive statistics for this scales as well.

Means were determined by summing participants ratings of items within each sub-scale and dividing by the total number of items for that sub-scale; these individual scores were then computed to obtain means and standard deviations for each sub-scale. ShapiroWilk test of

UES	N	$\bar{x}$	σ	Cronbach's Alpha
PIN	3	3.23	0.87	0.78
EXPO	4	3.34	1.09	0.82
UNEX	3	3.08	1.36	0.95
DIV	3	3.09	1.19	0.84

Table 4.7: Session 1 Other Scales

UES	N	$\bar{x}$	σ	Cronbach's Alpha
PIN	3	2.86	1.02	0.78
EXPO	4	3.30	1.37	0.92
UNEX	3	3.39	1.33	0.91
DIV	3	2.94	1.29	0.84

Table 4.8: Session 2 Other Scales

normality was done for each sub-scale . The total results are reported in Figure 4.14

Tests of Normality			
	Shapiro-Wilk		
	Statistic	df	Sig.
NO1	.902	33	.006
FI1	.929	33	.032
FA1	.938	33	.061
AE1	.919	33	.018
EN1	.933	33	.043
PU1	.942	33	.078
UES1	.955	33	.191
UNEX1	.885	33	.002
DIV1	.944	33	.091
EXPO1	.899	33	.005
PIN1	.905	33	.007
FA2	.965	33	.363
PU2	.953	33	.163
FI2	.918	33	.016
AE2	.909	33	.009
EN2	.928	33	.030
NO2	.924	33	.024
UES2	.938	33	.060
EXPO2	.871	33	.001
PIN2	.940	33	.068
DIV2	.932	33	.039
UNEX2	.864	33	.001

Figure 4.14: Normality ShapiroWilk Test all scales. The red color cells indicate that ShapiroWilk Test accepted the assumption of Normality

The ShapiroWilk test of normality refuted the assumption of normally distributed data for the majority of sub-scales. FA1, DIV1, UES1, FA2, PU2, UES2, and PIN2 were the only scales that accepted the assumption of normal distribution.

4.4.2 Search Scenario Perception

This section discusses an important assumption of the study design that both search scenarios Task A and Task B, although with different wording, were perceived as similar leisure search scenarios. Thus, neither one of them affected more than the other the motivation of participants nor their search behaviour, or emotions.

A paired t-test was conducted to determine whether there was a statistically significant mean difference between the UES scores when participants performed Task A and Task B. UES scores were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). Participants UES scores were higher in Task B ($M = 4.39$, $SD = 1.23$) as opposed to Task A ($M = 4.33$, $SD = 1.34$); the null hypothesis of equal distribution is accepted $t(32) = -0.289$, $p = 0.774$. Therefore, there was not evidence of enough variation of user's engagement across tasks.

The other sub-scales were compared with Wilcoxon signed-rank nonparametric test because the normality assumption was not met using ShapiroWilk Test. This comparison showed that only PIN sub-scale was statistically significant across tasks ($Z = -2.387$, $p = 0.017$). The mean PIN rating was higher ($M = 3.23$, $SD = 0.87$) in Task A than in Task B ($M = 2.86$, $SD = 1.02$). For this research, it is unclear 'why' PIN had statistical difference across tasks A and B. Perhaps users exposed to the 'curious' system at the beginning rated significantly low PIN in the second task, whereas the participants who got the curious system in Task B had already rated PIN high in the baseline system.

In order to test task perception, this research used Moshfeghi & Jose [102] task perception questionnaire. The ShapiroWilk test of normality refuted the assumption of normally distributed data for all the items. Wilcoxon signed-rank nonparametric test to compare related samples showed that the null hypothesis of equal distribution was accepted for all task perception items (e.g. Easy, Stressful, Interesting). Therefore, there is not enough evidence in the sample to support the idea that Task A and Task B were neither perceived differently nor affected heavily the search behaviour of participants.

In addition, as already mention in the section 4.3.3, participants were asked in a post questionnaire about 'Which task do you enjoy the most? Explain why' and 'Which task was the longest? Explain why'. Thus, allowing participants to report the preference and subjective explanation to their behaviour.

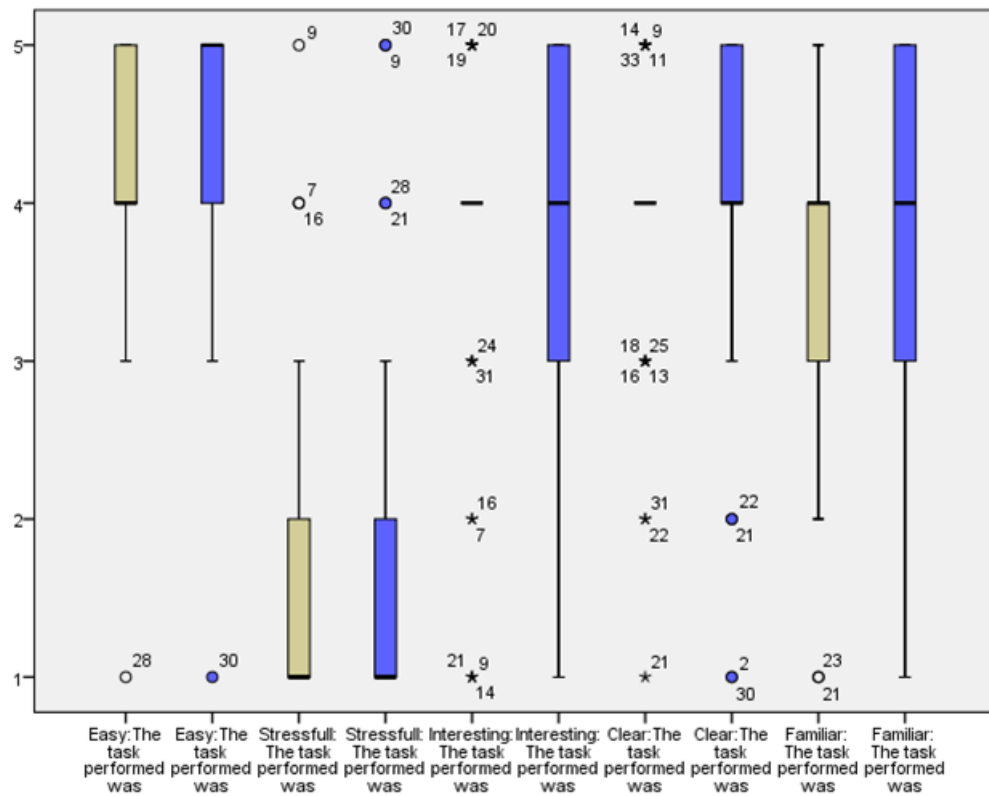


Figure 4.15: Task Perception between tasks (answer: 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”). Task A is brown and Task B is blue

A simple descriptive statistic over the first question (‘Which task do you enjoy the most? Explain why’) shows that 48.5% (n=16) liked more Task A, and 51.5% enjoyed more Task B. Therefore, this data and their provided explanations (i.e. the fact that none of the explanations mentions the type of task, the story, etc.) supports the idea that both tasks were perceived and performed with the same motivation and goal.

Regarding the second question question (‘Which task was the longest? Explain why’), 63.6% (n=21) participants thought their longest search session was Task A whereas 36.4% (n=12) guessed their longest exploration was Task B. Although this shows some kind of bias towards Task A, after reviewing each of the explanations there is not a single comment mentioning the type of task or particular goal of the task as being the reason of their selection or dwelling time on the task. Four participants participants comments reported that their longest search scenario was Task A because it was the first time they were exposed to the system.

Total Dwelling Time (TDT) [28] tracked by SurveyGizmo is another measurement that supports the assumption of both Task being very similar. TDT for Task A and B were analysed with ShapiroWilk test of normality and both rejected the hypothesis of normal distribution. So, this research employed Wilcoxon signed-rank nonparametric to test related samples and

found that there was not enough evidence in the sample to argue that either Task A or Task median B had statistically different total time, $z = -0.205$, $p < 0.05$. Figure 4.16 shows some descriptive statistics of TDT in Task A and B.

Descriptives				Statistic	Std. Error
TASKA_TIME	Mean			0:09:44	0:01:14
	95% Confidence Interval for Mean	Lower Bound		0:07:11	
		Upper Bound		0:12:16	
	5% Trimmed Mean			0:09:26	
	Median			0:06:44	
	Variance			185371.468	
	Std. Deviation			0:07:10	
	Minimum			0:01:02	
	Maximum			0:24:10	
	Range			0:23:08	
TASKB_TIME	Mean			0:09:07	0:01:06
	95% Confidence Interval for Mean	Lower Bound		0:06:52	
		Upper		0:11:22	
	5% Trimmed			0:08:39	
	Median			0:07:34	
	Variance			144617.127	
	Std. Deviation			0:06:20	
	Minimum			0:00:35	
	Maximum			0:26:23	
	Range			0:25:48	

Figure 4.16: Total Dwelling Time (TDT) for Task A and Task B

In summary, the assumption about that Task A and B did not influence strongly the search behavior and motivation of participants during the experiment is well supported by the analysis of the subjective responses of users as well as TDT.

4.4.3 Relationship among different constructs and total time on the task

For Task A, TDT had a weak moderate relation with NO and strong relation with UNEX. Among the sub-scales there were a strong relationship (above 0.60) between NO and FI, FA,

Item	TASK TIME	NO	FI	FA	AE	EN	PU	PIN	UNEX	DIV	EXPO
TASK TIME	1	.									
NO	.37*	1									
FI	0.20	.819**	1								
FA	0.23	.783**	.833**	1							
AE	0.26	.598**	.514**	.514**	1						
EN	0.03	.803**	.849**	.727**	.476**	1					
PU	0.01	.546**	.487**	.403*	.346*	.711**	1				
PIN	0.22	.617**	.602**	.703**	.626**	.552**	.344*	1			
UNEX	.68**	.575**	0.296	.431*	.403*	0.172	0.052	.451**	1		
DIV	0.249	.749**	.609**	.708**	.355*	.666**	0.337	.551**	.383*	1	
EXPO	0.282	.845**	.775**	.714**	.528**	.686**	.367*	.599**	.431*	.745**	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.9: General correlation between constructs and TDT in Task A

EN, PIN, DIV, EXPO; FI and FA, EN, PIN, DIV, EXPO; AE and PIN; EN and PU, DIV, EXPO; DIV and EXPO. Moderate high association (above 0.50) was noted between NO and AE, PU, UNEX; FI and AE; FA and AE; AE and EXPO; EN and PIN; PIN and DIV, EXPO.

PU relation with the other sub-scales such as FI was similar to previous research where they found that PU influenced FI [113]. However, in a similar way to the Ambiecities participants in the previous task, the strongest correlations of FI, EN and EXPO were with NO and FA. This relationships among participants show that hedonic features of the UX were perceived as more important than utilitarian values in the leisure scenario (e.g PU).

PIN and FA relationship relation supports the idea that interactivity fosters cognitive absorption and states of flow. PIN's high correlation with NO adds to the idea that interactivity features employed by the search application might stimulate users' curiosity, encouraging them to carry out more exploration with the site. The moderate relationship between UNEX and PIN also shows the role of interactivity as generator and influencer of the "cue effect"[127]. For example, certain cues (such as the presence of an attractive search notification message) are probably going to influence the persuasibility of the website appeal. For example, advertisers are well known to use them often to trigger positive associations with the advertised product.

Other important correlation was observed between NO and UNEX, this can be related to the fact that novelty, and uncertainty are collative variables of curiosity[18]. The high relationship gives indication that some of the items within these two scales may load the other scale.

For Task B, TDT had strong association (above 0.6) with NO and moderate (almost 0.5) with PIN, and UNEX. There was strong association (above 0.60) between NO and PIN, UNEX, EXPO; FI and FA, EN, PIN, DIV, EXPO; FA and EN; EN and PU, DIV, EXPO; PI and DIV, EXPO; DIV and EXPO. Moderate high relationship (above 0.50) was noted between NO and EN, DIV; FI and AE, PU; AE and EXPO, EN; EN and PIN; PU and EXPO, PU, UNEX; FI

Item	TASK TIME	NO	FI	FA	AE	EN	PU	PIN	UNEX	DIV	EXPO
TASK TIME	1										
NO	.649**	1									
FI	.159	.477**	1								
FA	.101	.310	.798**	1							
AE	.205	.433*	.581**	.509**	1						
EN	.103	.549**	.906**	.664**	.514**	1					
PU	.205	.414*	.575**	.226	.216	.675**	1				
PIN	.472**	.602**	.648**	.411*	.455**	.545**	.456**	1			
UNEX	.494**	.650**	.231	.151	.348*	.253	.142	.477**	1		
DIV	.243	.537**	.654**	.420*	.397*	.669**	.433*	.656**	.386*	1	
EXPO2	.330	.652**	.744**	.506**	.514**	.723**	.507**	.696**	.540**	.786**	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.10: General correlation between constructs and TDT in Task B

and AE; FA and AE, EXPO; AE and EXPO; EN and PIN; PU and EXPO; PIN and DIV, EXPO; UNEX and EXPO.

In both search scenarios NO, and UNEX are related to TDT. This posits that user's curiosity 'feeling' evoked by novelty and surprise, and total time exploring are closely related. This is an important validity aspect of the experimental setting because this suggests that participants had a hedonistic motivation (i.e. curiosity) rather than a utilitarian view when they were doing the tasks and the hedonistic motivation was more likely to trigger participants exploration than the perceived usability.

Tables 4.9 and 4.10 show Spearman's correlation between sub-scales in Task A and Task B. These statistics do not imply always causation, but provide an overview of the multidimensional nature of the search experience. In other words, these statistics provide a better picture of the user's search experience and the interrelation between the different psychological constructs measured during the experiment.

4.4.4 Personal Interest Variable Analysis

Personal Interest variable was manipulated during the experiment using a complete factorial rotation between the most (PI) and least interesting (NPI) news topics for a given participant. The main hypothesis was that similar to O'Brien newsreaders study [118] and McCay-Peet et al.[121], PI should predict and facilitate users' engagement and exploration.

Pairwise comparisons were performed within-subjects (repeated measures) analysis of medians of TDT using PI and NPI as independent variables with Wilcoxon signed-rank nonparametric test. TDT sub-scale of PI was statistically significant higher than NPI ($Z = -2.35$, $p = 0.019$). Another nonparametric test to check if PI and NPI independent variables affected the TDT distribution is Firendmans Two Way Analysis of Variance. ANOVA was not applied be-

cause the dependant variable TDT was not normally distribute. Friedmans' test demonstrates that TDT was significantly different between PI and NPI tasks during the exercise intervention for the same subjects, $\chi^2(2) = 10.939$, $p < .001$. Therefore, PI had a strong effect in TDT.

UES pairwise comparison using t-test for within subject analysis reported that PI tasks had higher scores ($M = 4.69$, $SD = 1.21$) as opposed to NPI ($M = 4.02$, $SD = 1.36$); the null hypothesis of equal distribution is rejected $t(32) = 3.36$, $p = 0.002$. So, there is also evidence of the PI role to enable user's engagement. Moreover, using Wilcoxon signed-rank nonparametric test, NO and UNEX sub-scales accepted the hypothesis of equal distribution for PI and NPI tasks. Therefore, a PI did not always by itself evoked curiosity .

4.4.5 System Assignment Variable Analysis

Following a similar analysis as the previous section, the effect of 'Curious' and 'Cold' system independent variables were analysed using a within subject design.

Pairwise comparison of TDT with Wilcoxon signed-rank nonparametric test was statistically significant higher for the 'Curious' than for the 'Cold' system ($Z = -3.31$, $p = 0.002$). Friedmans' test demonstrated that TDT distribution was significantly different between 'Curious' and 'Cold' tasks for the same subjects, $\chi^2(2) = 5.12$, $p < .001$. Therefore, System exposure independent variable had a strong effect in TDT within tasks.

UES pairwise comparison using t-test for within subject analysis reported that 'Curious' tasks had higher scores ($M = 4.45$, $SD = 1.36$) as opposed to NPI ($M = 4.26$, $SD = 1.21$). But the null hypothesis of equal distribution was accepted $t(32) = 0.837$, $p = 0.409$. There was not enough evidence of the role of system assignment and user's total score of engagement.

However, a Wilcoxon signed-rank nonparametric test for NO, and UNEX sub-scales demonstrated statistically variation within subjects between 'Curious' and 'Cold'. Wilcoxon signed-rank test for NO sub-scale showed higher scores for 'Curious' than the 'Cold' system ($Z = -2.97$, $p = 0.003$). UNEX sub-scale had higher scores for 'Curious' than 'Cold' system ($Z = -4.45$, $p < 0.001$). This results supports the idea that interactive and surprising search notifications had an important role in making participants curious and explore for more time. Figure 4.17 illustrates the relation between TDT and UES for the search tasks grouping by system.

4.4.6 Analysis Between Subjects for each Tasks

Because most sub-scales did not follow a normal distribution, this research tested between subjects difference in each task with Kruskal-Wallis test.

For task A and the system assignment independent variable, Kruskal-Wallis showed that

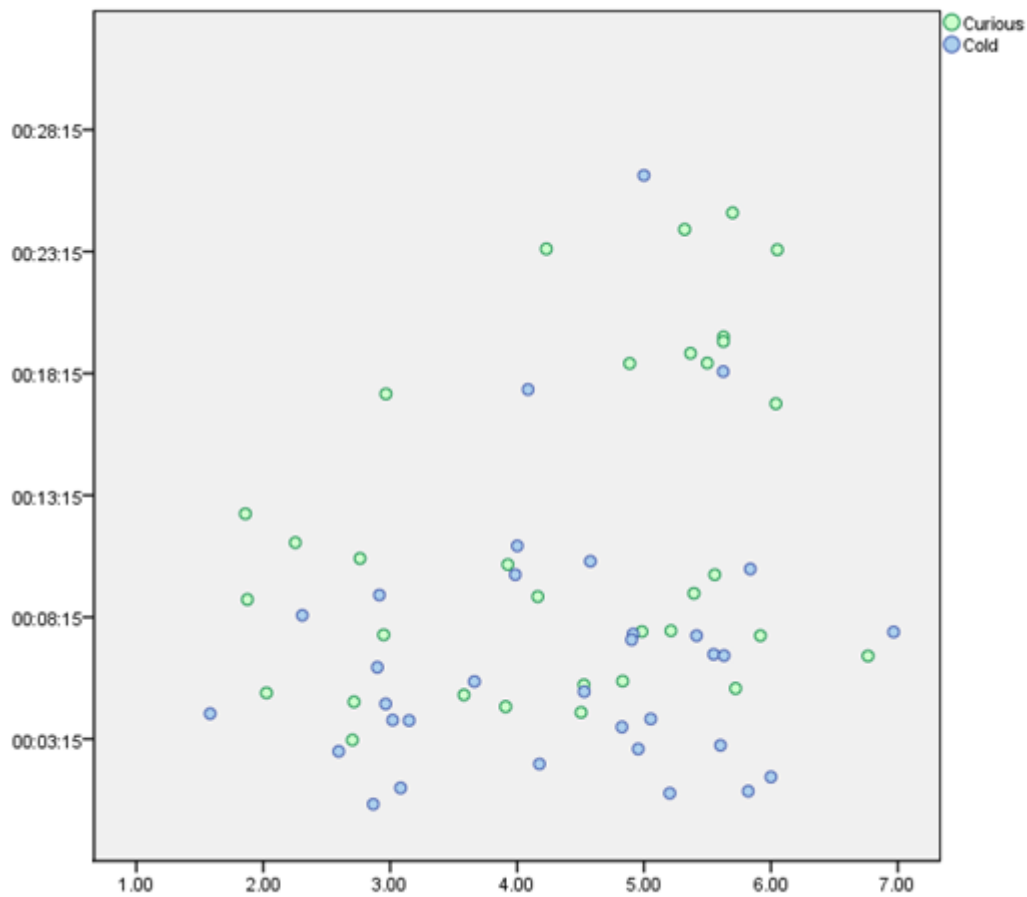


Figure 4.17: Scatterplot of total Dwelling Time (TDT) vs UES for both Task A and Task B. Grouped by system assignment: Cold is blue and 'Curious' is green

TDT, NO, AE, PIN, UNEX, and EXPO had statistically higher distributions when participants were exposed to the 'Curious' system rather than the 'Cold' system. These statistical significance are similar to the ones reported in the previous experiment with Ambicities and Twitter search interface. The only sub-scale that did not change significantly as in the previous experiment was FA.

In contrast, task A and the PI assignment independent variable, Kruskal-Wallis showed none statistically different distribution among participants assigned either to PI or NPI tasks. This is an interesting result because it seems that for Task A, PI was not that important between subjects and supporting the idea that participants in NPI task stayed longer because they were curious about the notifications and the information interaction. Tables 4.11 and 4.12 depicts both Kruskal-Wallis tests for task A.

In task B and the PI independent variable, Kruskal-Wallis noted that EXPO, PIN, FA, EN and FI had statistically higher distribution between subjects. For the same task and with system assign independent variable the test highlighted again that NO and UNEX were higher when

Test Statistics a,b

	TASK TIME	NO1	UNEX	EXPO	DIV	PIN	FI	FA	AE	EN	PU
Chi-Square	11.711	6.049	22.997	3.974	1.516	5.592	.522	1.781	3.214	.016	.105
Asymp. Sig.	.001	.014	.000002	.046	.218	.018	.470	.182	.073	.899	.745

a. Kruskal Wallis Test

b. Grouping Variable: system assignment

Table 4.11: Between subject analysis for system assignment in Task A

Test Statistics a,b

	TASK TIME	NO	UNEX	EXPO	DIV	PIN	FI	FA	AE	EN	PU
Chi-Square	1.682	.448	.076	.380	0.000	.593	2.528	.255	.189	.334	.005
Asymp. Sig.	.195	.503	.783	.538	1.000	.441	.112	.614	.664	.563	.942

a. Kruskal Wallis Test

b. Grouping Variable: interest

Table 4.12: Between subject analysis for system assignment in Task A

Test Statisticsa,b

	TASK TIME	NO	UNEX	EXPO	DIV	PIN	FI	FA	AE	EN	PU
Chi-Square	1.636	5.270	6.133	.477	.258	.008	.917	1.374	1.032	.074	.749
Asymp. Sig.	.201	.022	.013	.490	.611	.928	.338	.241	.310	.786	.387

a. Kruskal Wallis Test

b. Grouping Variable: system

Table 4.13: Between subject analysis for system assignment in Task B

Test Statisticsa,b

	TASK TIME	NO	UNEX	EXPO	DIV	PIN	FI	FA	AE	EN	PU
Chi-Square	1.682	2.826	.642	4.215	2.786	3.848	4.860	4.153	.553	5.369	.718
Asymp. Sig.	.195	.093	.423	.040	.095	.050	.027	.042	.457	.021	.397

a. Kruskal Wallis Test

b. Grouping Variable: interest

Table 4.14: Between subject analysis for system assignment in Task B

participants interacted with the ‘Curious’ system. Tables 4.13 and 4.14 shows Kruskal-Wallis tests for task B in both independent variables.

The results of PI in task B are interesting for leisure search research and the role of emotions in IIR because although both task were perceived in similar way, the result seem to suggest that after searching in NPI related task and perhaps being bored, in the second task participants PI had a more important effect when they were searching and evaluating the UX. For example, some users after being assigned to NPI (No Personal interest) topic commented that the task was not successful and show their frustration by quitting quickly the task, then for the second task with their PI (Personal interest), the users explored for a greater time and had positive comments. The overall TDT for both A and B scenarios also suggest that this search behaviour effect was stronger when users had CONPI (‘Cold’ NPI) and then CUPPI (‘Curious’ PI) than when users had CUNPI (‘Curious’ NPI) and COPPI (‘Cold’ PI).

In other words, for some the leisure search behaviour (or the desire to have fun) was reinforced in the second task by manipulating a NPI in the first task. Wilson & Elswiler [31] proposed the creation of an experimental setting where users are bored at the beginning of the experiment, then they received “a hedonistically motivated task” and a search tool to alleviate this feeling. But for them it was unclear how to trigger such a motivating setup with “ecological validity”. Perhaps as this results suggest personal interest (PI) manipulation can be an interesting way of studying “casual leisure search” behaviour in action.

However, as mentioned in section 4.4.2, TDT for Task A and B failed Wilcoxon signed-rank nonparametric test with related samples and found that there was not enough evidence in the sample to argue that either Task A or Task B had different TDT distribution.

In summary, the between subject analysis shows that system assignment affected in both task the NO and UNEX (Surprised) sub-scales suggesting that the system with push search notifications evoked participants curiosity even when the news were not from personally interesting topics. Figure 4.18 shows the total dwelling time compare to personal interest variable group by system assignment.

4.4.7 Search Notification Perception and UX

From the previous analysis, this research has been able to test that the search notifications system affected significantly the probability of being curious in a given task according to TDT, NO and UNEX sub-scales. Moreover this section presents the subjective comments and other items in the questionnaire related to the ‘curious’ search notifications. This section compares the effect of them in the UX with previous research as well.

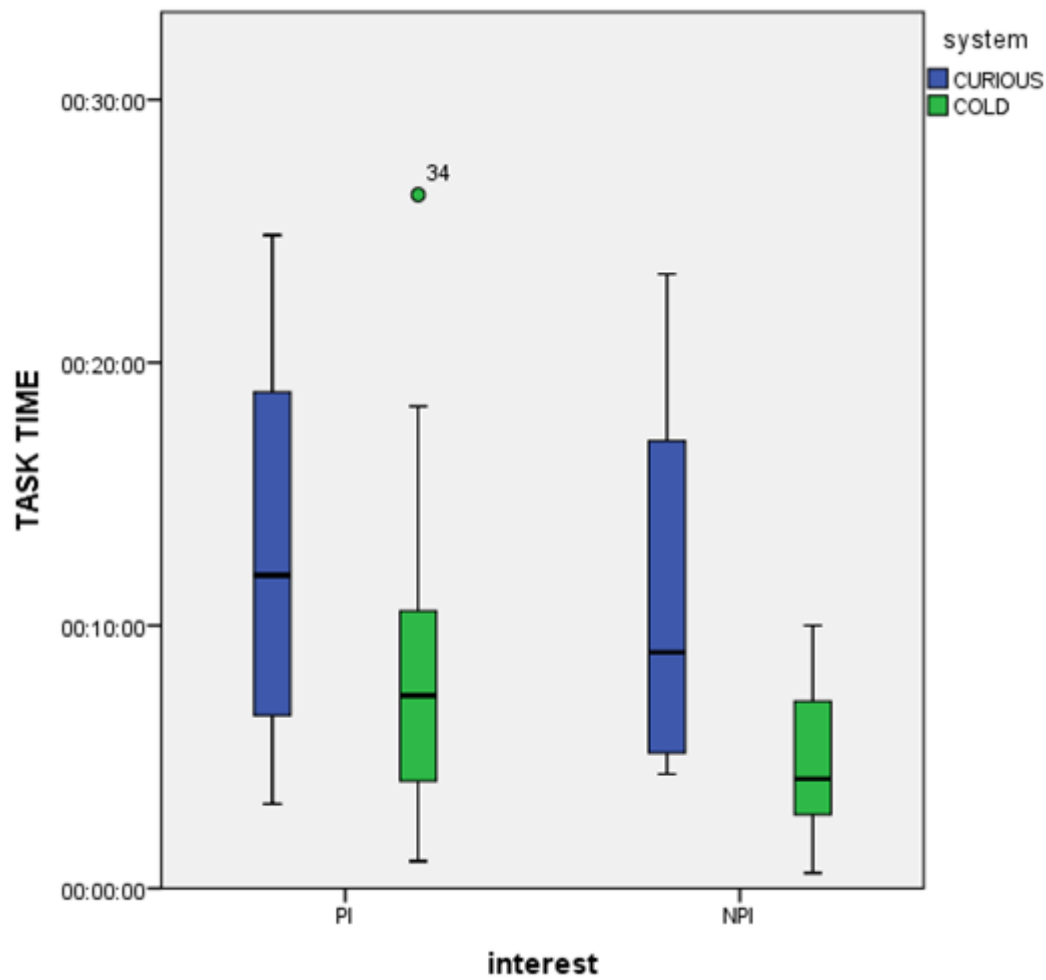


Figure 4.18: Boxplot of total dwelling time (TDT) vs personal interest (PI) across tasks. Grouped by system assignment: 'Cold' is green and 'Curious' is blue

The following comments are a sample of the comments given by participants to the question 'What did you like about the search experience? (For example a particular feature or functionality)', 'While using the App, did you enjoy your exploration?' or other comment sections in the questionnaire:

- Participant A: CUPI (0:18:40) vs CONPI (0:01:15)

"I had no search criteria, I was just browsing. In that I felt inspired enough to keep scrolling/clicking, it was relevant (i.e. interesting)".

"The text in the popup boxes was vague enough to click on, and I actually stumbled on interesting news "

"Popup boxes had me clicking on things and learning about new topics. However, this is the same reason I would never actually use this site, too much time-consuming. You feel compelled to keep clicking on popup boxes."

- Participant B: CONPI (00:04:42) vs CUPi (0:24:51)picu

“The GUI is very streamlined and easy to use. It is also very attractive. I liked being able to open news items in a new tab (and I shared a news item on Facebook). I liked the pop ups too, they let me open up new lines of information I could dig into.”

- Participant C: CUNPI (0:17:01) vs COPI (0:06:41)

“I like the popups and the possibility to go to the source of the news”.

“I discover a lot of interesting news and the interaction with the popups was great ”

“It was interesting to see the headings of different types of news and how my input on the pop ups had an effect on the articles that came up on the screen. It was also interesting how one article could take me to another one that was related.”

- Participant D: CONPI (0:02:45) vs CUPi (0:03:13)

“I like the popups. When I was searching and it didn’t generate results. A red pop-up appeared and I clicked. Just in time when there was no results, this help me to keep searching”

- Participant E: CUPi (0:24:10) vs CONPI (0:00:35)

‘I like the popups. They were related to topics of my interest’

- Participant F: COPI (0:01:02) vs CUNPI(0:19:34)

‘I enjoy the notification windows, they guide me to find better news articles.’

‘The notifications or suggestions keep me going and searching’

- Participant G: COPI (0:03:45) vs CUNPI (0:07:40)

‘News do not necessarily make me happy but at least the interactive features of the App can keep me occupied by offering topics I might like. I like that the pop-up offers disappear after a while so that they do not disturb too much.’

‘I liked the possibility to save news and look at them later. Also the popups.’

These are just a sample of the interesting comments regarding the popups. There are not negative comments regarding the role of the search notifications during the search experience

because they did not interrupt the reading flow and some of the participants kept clicking them to find new information before they were gone. This interactive feature made them explored more and evoked their curiosity, even when they were not personally interested in the news topic.

However, when they were asked to give suggestions some participants mentioned that the search notifications should be personalized and the frequency control (e.g. too engaging is also bad for some users). For example, one male participant wrote a feedback email that partially mentioned: *“The curiosity messages are great! They make you aware of stuff that is useful and you may not know. However, I am wondering if you could keep those messages on the screen for longer. One option would be to keep them altogether in a separate section once you have shown them. You may want to have a menu option (or a separate section), where users can see all the curiosity messages displayed so far. The reason is that users may not have time to click on the curiosity messages at the same time they are doing something else. In my case, I was looking at information about yesterday’s earthquake in Chile and suddenly I got a message that I thought it would be interesting to read. Though I wanted to click on that message, I had to let it go, because I had not finished reading about Chile”*. Although this comments suggest a possible enhancement in the search notifications, it is interesting to note that the participant is implicitly highlighting that the search notifications did not interrupt the reading flow but they were visible enough to made him wonder or ‘curious’ about what would happen if he had clicked the search notification. This description of UX is frequent pattern in the study.

For this research, it is very interesting to see that some of Ambiecities participants’ comments and opinions about the search experience are very similar to some of the comments of participants in the news searching study when they were evaluating the their search experienced in the “curious” system. Therefore, the compiled qualitative data supports the hypothesis that applying curiosity arousal principles such as novelty and surprise to the search experience by using search notifications increased both user’s engagement and user’s exploration.

Some comments also suggested that the information interaction provided by the search notification, made people more ‘distracted’ or ‘inspired’ to look for new things and some found unexpected interesting information (e.g. one of the used the word ‘stumble-upon’, to emphasised a serendipitous discovery). Other comments such as the one made by participant highlights the role of “partial exposure” in the design of some search notifications and their curiosity effect because when reporting Participant A mentioned that the notification message “was vague enough to click on”. In other words, the incomplete and vague information moved the participant to find out more about what the recommendation or suggestion in the push

notification.

Another important information about the effect and potential of ‘curious’ search notification is evident after a simple statistical analysis over these questions: ‘Did you notice the Search Notifications or Dialogs?’ and ‘Did you click in the Search Notifications or Dialogs?’. 29 (87%) participants saw the notifications and 21 (71%) of those participants clicked on them. When this responses are compared to previous news searching studies such as O’Brien et al. (2011) [118] where in sample of thirty participants only seven (23%) clicked or used the recommended links, this research finds enough evidence to highlight the power of interactivity and the curiosity principles used to design the search notifications. Because it is not only the content a recommendation (e.g. a link or search notification) and the algorithm behind it what it is important. It is also key the way in which the communication flows (e.i. the information interaction) between the system and the user; and how diverse human and interactive factors can be manipulated in the UX to change user’s decision making or encourage particular behaviours.

In summary, the qualitative feedback explains the high psychometric ratings of the Novelty (NO), uncertainty or surprise (UNEX) sub-scales and the distribution of total dwelling time (TDT) across search sessions where participants were exposed to the “curious” search notifications.

4.4.8 Total Dwelling Time and Guessed Total Dwelling Time

In both tasks, participants summed up a total dwelling time (TDT) of 10 hours, 22 minutes and 27 seconds; a guessed total dwelling time (GTDT) of 8 hours, and 36 minutes; an error guessed total dwelling time (EGTDT) of 1 hour, 46 minutes, and 27 seconds; and absolute error guessed total dwelling time (AEGTDT) of 3 hours, 43 minutes and 31 seconds.

Figure 4.20 shows guessed total dwelling time compare with personal interest and group by system assignment. This results seem to suggest despite of the personal interest when participants were assigned to search with the curious system they underestimated their dwelling time. In other words, they thought they spent less time searching than what they actually did. Perhaps this result could suggest that they were absorbed searching and exploring.

However, in contrast with Ambicities’ study, the relation between guessed total dwelling time, error guessed total dwelling time, absolute error guessed total dwelling time, and focused attention sub-scale (FA) did not show that the interaction induced most of participants to state of deep cognitive absorption. In a following section, this research hypothesize that a specific difference between Ambicities and Newslabels interactivity and search experience could be responsible for the variation in FA sub-scale ratings.

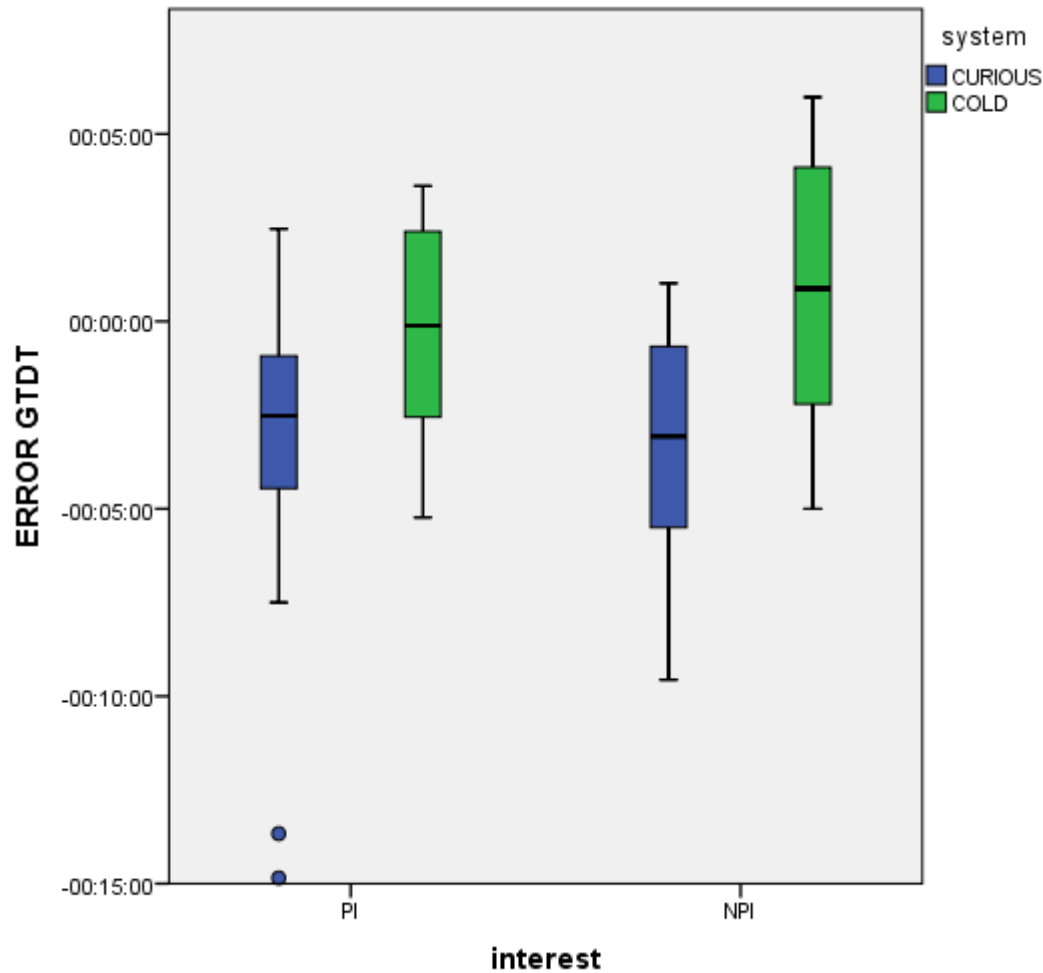


Figure 4.19: Boxplot of error guessed total dwelling time (EGTDT) vs personal interest (PI) across both tasks. Grouped by system assignment: ‘Cold’ is green and ‘Curious’ is blue

4.4.9 Curiosity as a trait

Using Litman et al. [19, 20, 8] curiosity trait definition and psychometric questionnaires, this research studied the relationship between the degree of user engagement and exploration with a given curiosity trait, and whether the assumption that individuals with high interest type curiosity (I-type) would be more likely to explored more ‘for fun’ than deprivation type curiosity (D-type), specially when exposed to the curious system.

Figure 4.20 shows the descriptive statistics for each type of curiosity trait. The brief, 5-item I-type curiosity scale ($M = 16.3$, $SD = 2.93$; and D-type curiosity scale ($M = 13.6$, $SD = 3.69$) both had acceptable internal consistency (I-type: $\alpha = .79$; D-type: $\alpha = .81$) and were moderately positively correlated ($r = .53$). This results are similar to previous research done by Litman et al. [8, 9] where the validity of the I/D model was tested with university students as well as non-student samples.

Interest type Curiosity (I-type)**Scale Statistics**

Mean	Variance	Std. Deviation	N of Items
16.30	8.780	2.963	5

Deprivation type Curiosity (D-type)**Scale Statistics**

Mean	Variance	Std. Deviation	N of Items
13.64	13.676	3.698	5

Figure 4.20: Descriptive statistics for I-type and D-type scales of curiosity as trait [8, 9]

After trying several statistical analysis, this research did not find enough evidence that curiosity as a personality trait had a strong influence in the user's engagement, total dwelling time, and the overall information seeking behaviour perceived during the experiment.

For this research, this was an interesting result because the weak relation between curiosity as a personality trait, and the other psychometric measurements implies that it was the interaction with the 'curious' IIR-system, not the personality of the participant, what affected their search behaviour, and perception of the UX. Thus, it is possible to design to trigger curiosity in spite of the user's personality or curiosity trait.

This result also shows that there must be a relation between the mechanisms supporting extrinsic reward curiosity and the intrinsic curiosity. Gruber et al. (2014) [128] found a similar relation in a psychological study, and highlighted the importance of stimulating curiosity to create more effective learning experiences. They found that within a learning scenario if an extrinsic mechanism is used to evoked curiosity simultaneously even when unrelated to the goals of the scenario, the mechanism has the potential of boosting the subjects intrinsic motivation and improve the subject's learning process. In accordance with this research, the psychometric ratings, and the total dwelling time suggest that the search notifications even when unrelated to the topic of interest boosted the participants' intrinsic motivation to find out more and explore for more time than when they were presented with "normal" query-response search experience.

4.4.10 Analysis of Interaction Logs

Several performance indicators were recorded during the users' search sessions. As already mentioned, there were 494 unique sessions during the period starting the 18th August of 2015 and ending the 2th October of 2015.

Search Notification IIR-System	Total Session IDs	Percentage Total Session IDs
Enable or "Curious"	182	36.84%
Disable or "Cold"	312	63.16%
Total sessions	494	100%

Table 4.15: Total sessions grouped by system type. The system with "Curious" search notification (enable) and the "Cold" search system (unable).

Topic	Total Session IDs	Percentage Total Session IDs
Financial	155	31.37%
General	215	43.52%
Sports	124	25.11%
Total sessions	494	100%

Table 4.16: Total sessions grouped by topic (Financial, General, Sports).

Table 4.15 shows the total number of sessions grouped by system assignment. From the total number of sessions, 312 (63.16%) were performed with the "Cold" IIR-system (baseline), and 182 (36.84%) were done with the "Curious" IIR-system. Table 4.16 depicts the total number of sessions grouped by topic (Financial, General, Sports). The general news vertical had 215 (43.52%) sessions, the financial had 155 (31.37%) sessions and the sports 124 (25.11%).

For each session, this research consolidated the total number of individual actions such as total number of queries, total number of click search notifications, total of saved documents, and total of unsaved documents. Table 4.17 shows the descriptive statistics for total number of actions group by system assignment, topic and type of action.

Figure 4.21 illustrates max, min, and mean for each action grouped by topic. Although Figure 4.21 shows that the general topic had the session with the maximum number of queries (118), the rest of statistics for each action across topics followed a similar distribution. For instance, the mean number of queries grouped by topic were very similar (General=3.56, Sports=3.40, Financial 2.58). Thus, the interaction logs suggest that topic was not the most important factor to determine users' search behaviour and their affective state while using the Newslab web application. This result is similar to the one describe in section 4.4.4 regarding the variable of personal interest.

There was a total of 3300 event actions in the interaction log. Figure 4.22 describes the mean number of actions grouped by system assignment and type of action. This graph portraits that

Notifications	Topic	Action	Total Actions	Mean Actions	Min	Max	Stddev
disable	Financial	Query	181	1.88	1	9	1.78
disable	General	Query	248	1.66	1	9	1.30
disable	Sports	Query	127	1.89	1	6	1.47
disable	Financial	Save	111	6.52	1	26	6.13
disable	General	Save	118	6.21	1	24	6.98
disable	Sports	Save	35	2.69	1	10	2.62
disable	Financial	Unsave	1	1.0	1	1	NaN
disable	General	Unsave	3	1.0	1	1	0.0
disable	Sports	Unsave	2	1.0	1	1	0.0
enable	Financial	Query	193	3.27	1	70	9.01
enable	General	Query	360	5.45	1	118	14.81
enable	Sports	Query	280	4.91	1	32	6.28
enable	Financial	Save	50	6.25	3	15	4.065
enable	General	Save	191	8.6	1	30	7.96
enable	Sports	Save	130	7.2	1	26	8.40
enable	General	Unsave	2	1.0	1	1	0.0
enable	Sports	Unsave	3	1.5	1	2	0.70
enable	Financial	Click	55	3.6	1	25	5.97
enable	General	Click	131	4.85	1	40	7.87
enable	Sports	Click	91	3.37	1	13	3.18

Table 4.17: Descriptive statistics for total actions group by system assignment, topic, and type of action

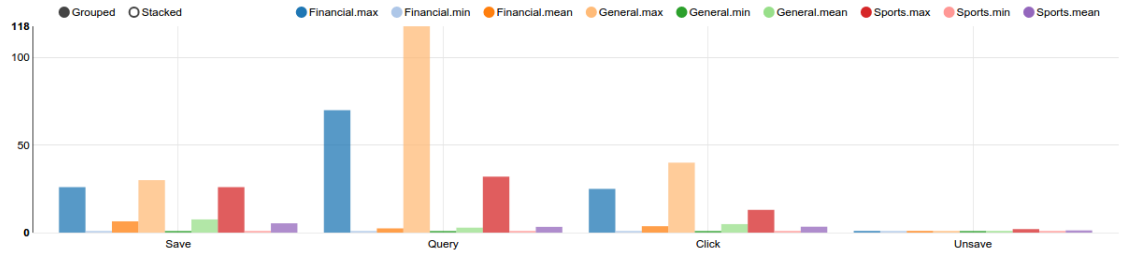
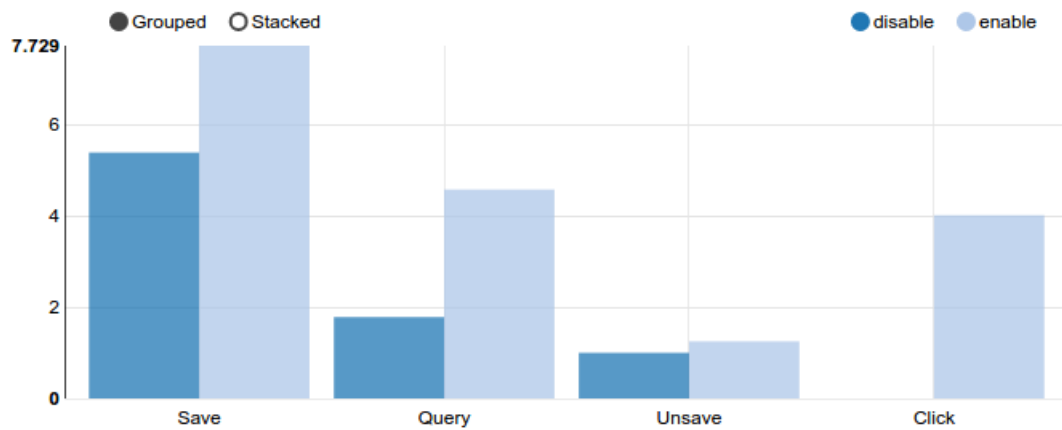


Figure 4.21: Descriptive statistics for each action group by topic

the mean number of queries for the search sessions with “Curious” IIR-system ($M = 4.57$, $SD = 10.85$) is almost three times more than the sessions with the “Cold” IIR-system ($M = 1.78$, $SD = 1.50$). Table 4.17 depicts that a similar relationship between system assignment holds true for each topic. For instance, the number of queries in the financial, general and sports vertical are 1.74, 3.28, and 2.59 times greater for the search sessions with curious notifications than the sessions with “normal” query-response paradigm. Table 4.18 shows the descriptive statistics for each action group by system assignment.

There was a total of 1389 queries in the interaction logs. 833 (60%) were done with the “Curious” IIR-system and 556 (40%) with the “Cold” IIR-system. The numbers should be understood by taking into account that there were more search sessions with the baseline (312) than with the IIR-system with search notifications. Then, it is clear that users of the curious system performed more queries in total. Therefore, this research can state that users searching

with the search notifications enable were more likely to made more queries than the baseline users.



h

Figure 4.22: Mean number of actions group by type of action and system assignment

Notifications	Action	Mean Actions	Stddev	Min	Max	Number Sessions	Total Actions
disable	Query	1.78	1.50	1	9	312	556
disable	Save	5.38	5.94	1	26	49	264
disable	Unsave	1.0	0.0	1	1	6	6
enable	Query	4.57	10.85	1	118	182	833
enable	Save	7.72	7.56	1	30	48	371
enable	Unsave	1.25	0.49	1	2	4	5
enable	Click	4.01	5.94	1	40	69	277

Table 4.18: Descriptive statistics for each action group by system assignment

Another interesting insight from the data analysis was evident after counting how many unique queries there were across sessions group by system assignment. The “Cold” query-response had a total of 34 unique queries and the “Curious” had a total of 233 unique queries. Figure 4.23 illustrates the relationship. This finding highlights a large difference in the search behaviour between these two groups. Users interacting with the IIR system with search notifications elaborated more diverse explorations over the different social media news, and not just performed more queries. It is important to note that the statistic of unique queries does not take into account the time filters used in every query, it was simple count on the text o keywords employed in the query.

Table 4.19 shows the top 50 most common queries grouped by system. The observed data suggest that most of the users interacting with “Cold” IIR either scroll-down using the infinite list functionality or they query the IIR-system using only the timeline filters (e.i. popular news in the last hour, popular news in the last 24 hours) because Table 4.19 describes that the two top queries are empty string enable, and empty string disable. The empty Query event was

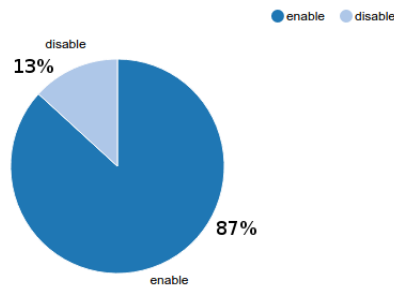


Figure 4.23: Total unique queries grouped by system assignment

trigger every time someone opened the Newslabls web application, and every time an user clicked in a timeline filter button without elaborating a text query.

The “Cold” IIR-system had only one query in the top 50 of most common queries (Rank 43th). Whereas the users’ interacting with the curious system were inspired to explore and search more topics. This data describes that 33% of the queries issued in the Curious IIR-system were trigger by the action of clicking a ‘surprising’ search notification (general, sports, financial). Other 48% of the queries had the same text of the keywords in the search notifications. This finding suggest that after clicking in a search notification and browsing for while, users’ click in the different timeline filters to change the content presented to them. The percentage of unique terms not included in any of the search notifications is the 3%. Therefore, this data demonstrates that most of the query actions were influenced by the appearance of a search notification or a previously clicked search notification. This research got this conclusion based on the correspondence between the keywords used in the search notifications and the following text employed in the queries. This behavioural pattern found in the data logs, it is supported too by the participants comments regarding the search notifications UX and the psychometric scores.

The analysed data posits that users interacting wit the “Cold” IIR-system scroll down until they would disengaged, or clicked on a interesting link and left the website. Whilst people interacting with the “Curious” IIR-system could received a surprisingly search notifications which would potential re-trigger users’ engagement with the IIR-system, and made them search longer time. Table 4.20 presents the descriptive statistics of total time in a search session group by system assignment. The mean total time in a search session for users exposed to the IIR-system with search notification is almost four times higher than those interacting to the baseline. However, the “curious” group had a high standard deviation in comparison with the “Cold” group. This shows that the total time values are widely spread in the log.

Table 4.21 shows that 69 (38%) search sessions with search notifications had at least one click

Rank	Notifications	Query	Frequency
1	enable	""	399
2	disable	""	283
3	enable	Clicked Search Notification - General	130
4	enable	Clicked Search Notification - Sports	89
5	enable	Clicked Search Notification - Financial	55
6	enable	pedro chelsea	14
7	enable	research science geo	13
8	enable	travel places	13
9	enable	specialolympics LA2015	11
10	enable	transfer gossip	11
11	enable	chelsea manchester arsenal premier league manutd	11
12	enable	world	10
13	enable	clinton trump walker bush carson cruz huckabee rubio paul polls	10
14	enable	Marilyn Monroe	9
15	enable	calais migrant	8
16	enable	Cosby sexual assault	8
17	enable	mexico donald trump	8
18	enable	tube strike	8
19	enable	corruption platini	8
20	enable	galaxies earth nasa satellite asteroid universe	8
21	enable	debt revenue deficit tax income	7
22	enable	war	7
23	enable	falcao chelsea	7
24	enable	di maria psg	7
25	enable	best	7
26	enable	nuclear explosion	6
27	enable	baseketball lakers knicks	6
28	enable	kazan2015	6
29	enable	kenya	6
30	enable	cats kitten	6
31	enable	quebec ottawa canada	6
32	enable	animal photo pet underwater	6
33	enable	iran	6
34	enable	CEO CTO CIO	5
35	enable	war hate terrorism	5
36	enable	champions winner gold	5
37	enable	balotelli millan	5
38	enable	google yahoo facebook microsoft mozilla twitter oracle apps	5
39	enable	glasgow	5
40	enable	price market stock	5
41	enable	modern family	5
42	enable	ebola	5
43	disable	higuaita	5
44	enable	cincinnati masters	5
45	enable	terrorism yemen tunisia ISIS attack	5
46	enable	galaxies earth nasa satellite asteroid	5
47	enable	Nino	4
48	enable	jokes	4
49	enable	cuadrado	4
50	enable	labour jeremy corbyn	4

Table 4.19: Top 50 most common queries grouped by system.

Notification	Mean (Minutes)	Stddev (Minutes)	Max (Minutes)	Min (Minutes)
Enable	21.40	68.31	745.12	0.02
Disable	5.17	12.79	140.27	0.0

Table 4.20: Descriptive statistics of total time in a search session group by system assignment.

event. This result must take into account that there were 218 accesses reported by SurveyGizmo, and only 37 completed both task. There were 4 participants that reported problems with NewsLabels. Therefore, this number must be considered having in mind that 1) multiple search sessions ids could be related to an individual participant, 2) sometimes when a searcher came back from visiting another website the result page was refreshed because the internal links in the tweet were not created using Emberjs tolink component ²⁵ and 3) search notifications could have encouraged curiosity even when the users did not click on them.

There was total of 635 click events in the log. This meant on average 4.01 clicks for all the search sessions with the search notifications.

Notification	Did Click	Total Number of Sessions
enable	Yes	69
enable	No	113

Table 4.21: Top 10 most common news sources in the saved Tweets

The total recorded time for all session was 5444 minutes (90.7 Hours). The total time metric was calculated after subtracting the date recorded in the Close action and the date logged in the Open action for the same session id. This means that after searching for a while an user could open a different tap, and start doing other things with the browser without ending the search session. The system gets a Close event when the user closes the page, refreshes the page or closes the entire browser. Thus, if an user forgets the web application opened in one of tabs, the session total time will include this time as part of the search session. The same can be said for the users who refresh the tab by mistake or due to the Embers linkto issue when returning from a visited news source website because every time the web app is loaded or refreshed, the user would be assigned to a new session id. This problem appeared when the users click on a twitter link in the text of the tweet, but not when they click in the source button. Thus, it is possible that a real search sessions could combine two or more recorded sessions in the interaction logs.

Table 4.22 shows the top 10 longest sessions. For instance, the longest search session (5dHBk) started at 2015-08-24 22:01:18.0 and ended at 2015-08-25 10:26:25.0. Although the interaction logs show that the user was engaged searching and exploring (e.i. 118 Queries, and

²⁵<http://stackoverflow.com/questions/27016241/emberjs-routing-page-reloads-on-every-link-click>

SessionID	Open	Query	Close	Click	Save	Unsaved	Total Actions	Open Date	Close Date	TDT(Min)	TDT(Sec)	Notification
5dHBk	1	118	1	40	0	0	160	2015-08-24 22:01:18.0	2015-08-25 10:26:25.0	745.117	44707	enable
E5jO9	1	4	1	2	0	0	8	2015-08-19 09:46:33.0	2015-08-19 15:41:40.0	355.11	21307	enable
XX5IA	1	1	1	0	0	0	3	2015-08-18 18:29:07.0	2015-08-18 22:37:37.0	248.5	14910	enable
tTU9O	1	17	1	2	0	0	21	2015-08-25 18:38:19.0	2015-08-25 22:43:35.0	245.26	14716	enable
b7163	1	4	1	3	0	0	9	2015-08-18 14:36:27.0	2015-08-18 18:02:05.0	205.63	12338	enable
Eofv5	1	1	1	0	0	0	3	2015-09-18 11:57:49.0	2015-09-18 14:18:05.0	140.26	8416	disable
VwlTH	1	8	1	4	4	0	18	2015-08-30 19:09:21.0	2015-08-30 20:48:53.0	99.53	5972	enable
7nbCv	1	2	1	0	4	0	8	2015-09-11 12:15:39.0	2015-09-11 13:40:33.0	84.9	5094	disable
obaCq	1	70	1	25	0	0	97	2015-08-25 11:04:51.0	2015-08-25 12:28:12.0	83.35	5001	enable
tRbvR	1	1	1	0	0	0	3	2015-08-23 22:50:51.0	2015-08-24 00:02:16.0	71.41	4285	enable

Table 4.22: Top 10 longest sessions.

40 Clicks), the dates suggest that the user forgot to close the web application before going to sleep. There were two search session without search notification among the 10 longest sessions.

SessionID	Open	Query	Close	Click	Saved	Unsaved	Total Actions	Open Date	Close Date	TDT(Min)	TDT(Sec)	Notification
5dHBk	1	118	1	40	0	0	160	2015-08-24 22:01:18.0	2015-08-25 10:26:25.0	745.12	44707	enable
obaCq	1	70	1	25	0	0	97	2015-08-25 11:04:51.0	2015-08-25 12:28:12.0	83.35	5001	enable
Vo9ms	1	26	1	16	30	0	74	2015-09-13 01:07:43.0	2015-09-13 01:24:43.0	17.00	1020	enable
fV5Ig	1	27	1	11	25	0	65	2015-09-16 14:26:28.0	2015-09-16 14:59:41.0	33.22	1993	enable
TdHp3	1	18	1	11	27	0	58	2015-09-17 14:59:28.0	2015-09-17 15:11:06.0	11.64	698	enable
u2DnQ	1	32	1	6	0	0	40	2015-08-25 10:36:54.0	2015-08-25 11:04:27.0	27.55	1653	enable
OO68b	1	13	1	9	11	0	35	2015-09-13 22:51:59.0	2015-09-13 22:57:29.0	5.50	330	enable
Au9Ql	1	16	1	10	3	0	31	2015-09-19 08:31:43.0	2015-09-19 08:37:37.0	5.90	354	enable
sbtza	1	8	1	1	19	0	30	2015-09-15 14:39:43.0	2015-09-15 14:55:02.0	15.32	919	enable
eu5TE	1	1	1	0	26	0	29	2015-09-20 21:11:34.0	2015-09-20 21:32:08.0	20.57	1234	enable

Table 4.23: Top 10 longest sessions.

Table 4.23 reports the top 10 sessions order by total actions. There was not a single “Cold” search session within the ten sessions with most actions. For instance, the 3th search session (Vo9ms) suggest that the user had an engaging UX with the Newslabel web app because the user performed 26 Queries, 16 Clicks to Search Notifications, and 30 Saved documents within 17 minutes of exploration. This means almost one search notification click every minute. It seems that the user instead of clicking to look the source of the news, and visit other websites, the user search and browse within the Newslabels domain.

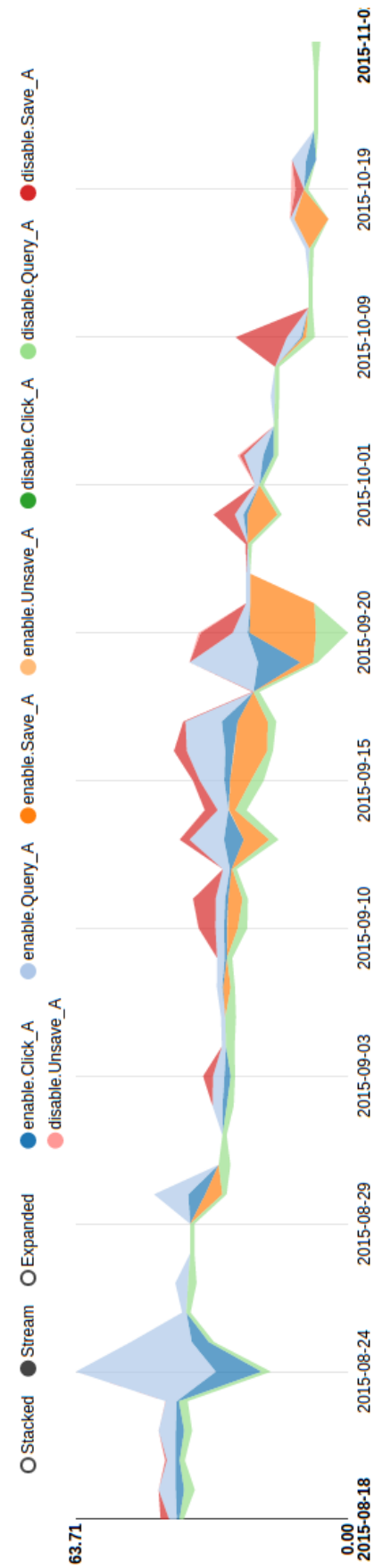


Figure 4.24: Stream graph representing the relation between different actions on both systems group by date



Figure 4.25: Stream graph representing the relation between different actions in a sample of search sessions



Correlations					
		query	click	saved	TDT (minutes)
query	Pearson Correlation	1			
	Sig. (2-tailed)				
	N	494			
click	Pearson Correlation	.919**	1		
	Sig. (2-tailed)	.000			
	N	494	494		
saved	Pearson Correlation	.137**	.206**	1	
	Sig. (2-tailed)	.002	.000		
	N	494	494	494	
min	Pearson Correlation	.679**	.610**	.034	1
	Sig. (2-tailed)	.000	.000	.456	
	N	494	494	494	494
**. Correlation is significant at the 0.01 level (2-tailed).					

Table 4.24: Correlations within the recorded actions and TDT

In contrast, 10th search session (eu5TE) shows that either the user did not understand the role of search notifications in the UX (e.g. when the user saw the search notifications, the user thought the purpose of the notifications was to just present information, and not to be click) or the user was not able to see the search notifications in the screen because at the moment of presenting the search notifications, the user was probably exploring the news source website (e.g. the user finds an interesting news from CNN, and clicks the source button. Then a new tab is created where CNN website is loaded. While the user is in the CNN website, the user is not able to see the notifications presented by the Newslab web application loaded in another tab).

The data indicates that clicking the search notifications trigger more user's interaction and exploration. Figure 4.24 shows a stream graph representing the relation between different actions such as clicking the search notifications, saving news, and querying the IIR system group by date. The graph shows how every time there was a grow in the quantitative value of Clicks (dark green), there was a quantitative grow in the number of Queries (pastel green) for the sessions interacting with the "Curious" IIR-system.

Figure 4.25 illustrates a stream graph representing the relation between the different actions in a sample of interaction logs group by session id. The graph shows an interesting pattern that when there is a pick in the number of queries (Blue), there is a pick in the number of clicks for most sessions (Pastel Orange). This sample contains sessions with the "Cold" system as well as the "Curious". Other pattern is the relation between total number of documents saved (Green) and total number of queries (Blue).

Figure 4.26 portraits the the different search sessions group by system assignment in a timeline. The graph shows how at beginning of the experiment, there were more search sessions.

The data also reflects that specially in the middle of September this research sent emails to various City University's email lists requesting students and staff to participate in the experiment. Figure 4.27 depicts the total amount of time of all search session group by system assignment and date. The observed data shows that even though there were more search sessions with the baseline system per day than with the "Curious" system, the total amount of time per-day was almost always higher for the "Curious" search sessions. The only time when the total amount of time per-day is higher for the baseline, there were not search sessions with the "Curious" system.

Table 4.24 depicts the correlation within the main action events (Click, Query, Save) and total dwelling time recorded by logger in the IIR-system. TDT had a strong association (above 0.60) with Click and Query. Therefore, it seems that every time a participant click in a search notification the probability of making more queries and searching for more time time grow.

There was a strong correlation between Click and Query events too. Save events had low correlation with Click and Query events. Nonetheless, the correlation with Click was stronger than with Click.

The highlighted data indicates that users who clicked the search notification made more queries and explore for more time than those who did not clicked on the search notifications, or did not have search notifications enable in their search experience. Figure 4.28,4.29, 4.30 and 4.31 illustrates the effect of clicking between all search sessions group by system assignment and the fact that an user clicked on a search session (1=Clicked, 0: No Clicked). Figure 4.28 and 4.29 explain that the distribution of queries and dwelling time grew considerably when users clicked on the search notifications. In fact the number of queries increase only when users clicked in the search notification. Thus, again the data visualization suggests that clicking in the search notifications meant more queries, and diverse exploration.

For this research, this data analysis is consistent with the collected psychometric scores and the participants opinions regarding their search experience with the "Curious" system because both the presented data and the psychometric questionnaires posit that users interacting with the search notifications explored for more time and diverse topics (e.i. made more and diverse queries). With the interaction logs, this research has confirmed how the "curious" search notification influenced the search behaviour and made the participants more active in their exploration.

Now, this research presents information regarding the click search notification actions. Table 4.25 indicates the top 10 search notifications most commonly click. The most common search notifications reflect common topics during the elaboration of this experiment. For instance,

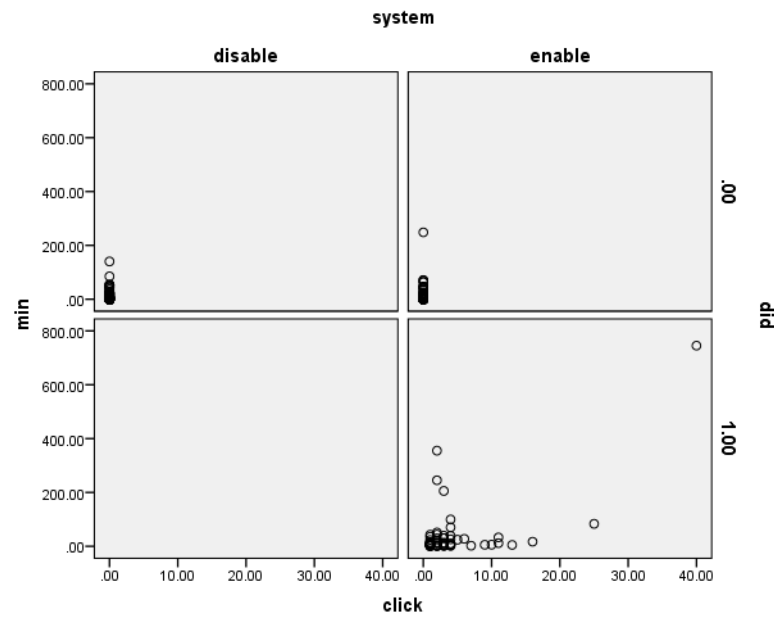


Figure 4.28: Matrix Scatter Plot between DTD (Minutes) and Click events

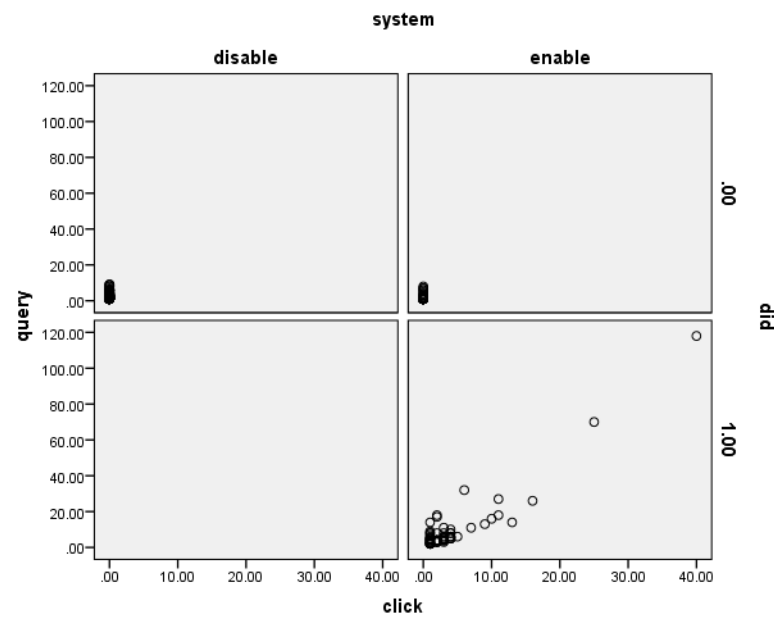


Figure 4.29: Matrix Scatter Plot between Query and Click events

the 4th most common click search notification had the message “sexy, funny, accessible and enigmatic. Who was she?”. This search notification was designed having in mind the concept of partial exposure to evoked the users curiosity. But there were other important factors that made the notifications a curious surprise. The ‘curious’ search notifications were presented for a limited time span. The time span was random value between 3000 to 8000 milliseconds. Table 4.27 shows a descriptive statistic for the duration of the click search notifications. The data

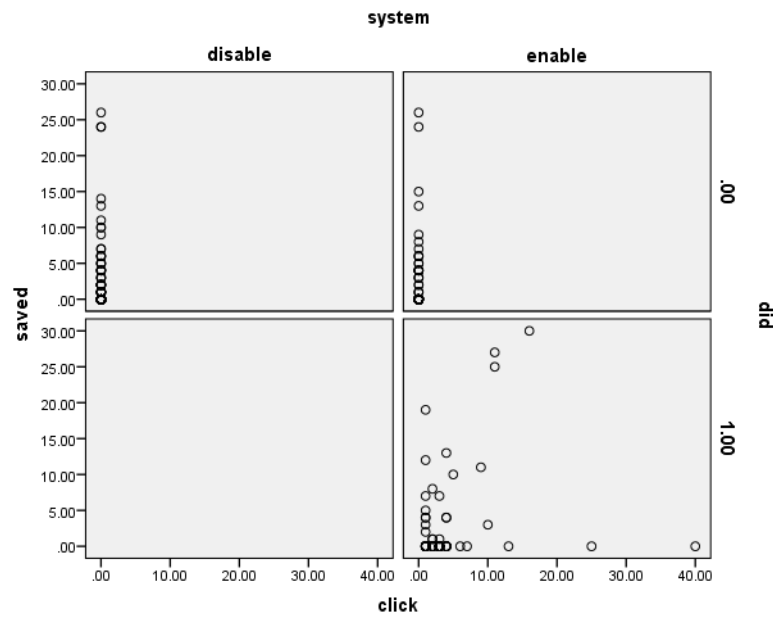


Figure 4.30: Matrix Scatter Plot between Query and Click events

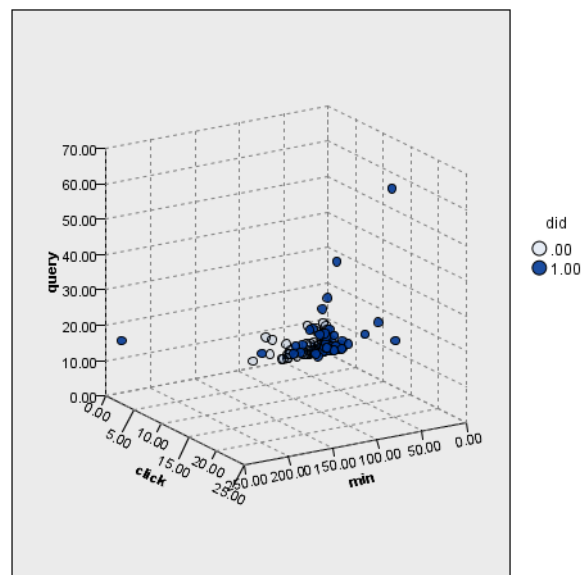


Figure 4.31: 3D Scatter Plot between TDT(Minutes), Click and Query events

reveals that the mean duration of the click search notifications was 5773 milliseconds. Thus, the data suggest that the best duration to display the search notifications was around 5 to 6 seconds.

The location of the search notifications in the screen was also a random function between three possible locations. The Table 4.26 shows that the most common position where the search notifications have been clicked was the top right corner. So, the data seems to suggest that the

Rank	Term	Frequency
1	11	30
2	today	29
3	world	27
4	new	26
5	#yearinspace	24
6	what	22
7	years	21
8	how	20
9	#mufc	20
10	@space_station	19
11	now	17
12	day	17
13	september	17
14	real	16
15	best	15
16	@cristiano	15
17	lion	14
18	#hala	14
19	thanks	14
20	people	13
21	14	13
22	madrid	12
23	night	12
24	cecil	12
25	photo	12
26	#rmucl	12

Table 4.28: Most frequent 26 terms int the saved Tweets

visualization was created using D3.js and Jason Davies word cloud library ²⁶. The visualization depicts that users saved tweets about remembrance 9/11 (e.g. 11, September, 14, years), football (e.g. real madrid, #hala, #mufc, @cristiano), universe (e.g. #yearinspace, @space_station, pluto) and the killing of Cecil the lion.

There were a total 635 saved actions during all search sessions. 371 (58%) saved actions were done by users interacting with the “Curious” IIR-sytem in 48 search sessions whereas 264 (42%) were performed by the users searching in the baseline IIR-system in 49 search sessions. This means that 26% of all the curious sessions had at least one saved event whilst 15% of all the baseline sessions had at least one saved event.

Table 4.28 presents the top 26 most common terms in the saved Tweets. The most common term in the saved tweets was the number 11 which it is related to the commemoration news related to 9/11 attacks in 2001 (14 years later). The 2th, 4th, 11th and 12th terms are linked to the temporal dimension of the news. The 5th and 10th most frequent terms are related to the space and the universe. The 9th, 14th, 15th, 18th and 22th term are about football. The 17th and 24th term are associated with the news regarding the killing of Cecil the lion in Zimbabwe.

²⁶<http://bl.ocks.org/ericcoopey/6382449/>

Table 4.29 depicts the top 10 sources among the saved tweets in all search sessions. The most common source was NASA's Twitter Account. All their tweets contained beautiful photos of earth, pluto, the sun and other starts in the universe.

The same can be said about the astronaut Scott Kelly tweets. Manchester United, SportsCenter, and Champions League are information sources of sport news. The Economist and CNNMoney are the top financial sources of information.

It is interesting that the most saved Tweets were not related to a particular classical news sources like CNN, BBC and Times.com (e.g. politics of countries) when participants were using the general news vertical. They decided to save tweets with beautiful pictures and landscapes of the universe from NASA, and Scott Kelly's Twitter account (8% all saved tweet). This suggest that some participants in order to complete the open ended question in the task decided to save appealing and relaxing pictures instead of the classical media news which could have been more popular and important. Therefore, this pattern in the search behaviour suggests that the relevance assessment was not based on the news item's topic relevance, but on how they felt when they were looking to a particular image. Further research should be done to understand the effect of content and type of content in people's search behaviour within a leisure search scenario.

Rank	News source	Frequency
1	NASA	30
2	The Economist	22
3	Scott Kelly	20
4	Manchester United	16
5	SportsCenter	13
6	CNN	13
7	TIME.com	10
8	Champions League	8
9	CNNMoney	7
10	The Verge	6

Table 4.29: Top 10 most common news sources in the saved Tweets

Figure 4.33 and 4.34 shows a collage visualization of the most saved tweets. The most saved tweet was related to the day of world suicide prevention with a total of 7 saves. The fourth was related to Cristiano Ronaldo with a total of 4 saves. The fifth was related to 9/11 with a total of 4. The 6th (4 saves), the 7th (3 saves) and the 10th (3 saves) were related to photos of the earth and some news regarding pluto. Table 4.30 shows the top 10 most saved tweets. The twitter ids are ordered by the total count of saved actions. The low amount of saved actions over a particular news item reflects the nature of social media and the news vertical where one important news today may not be as important tomorrow or even in the next two

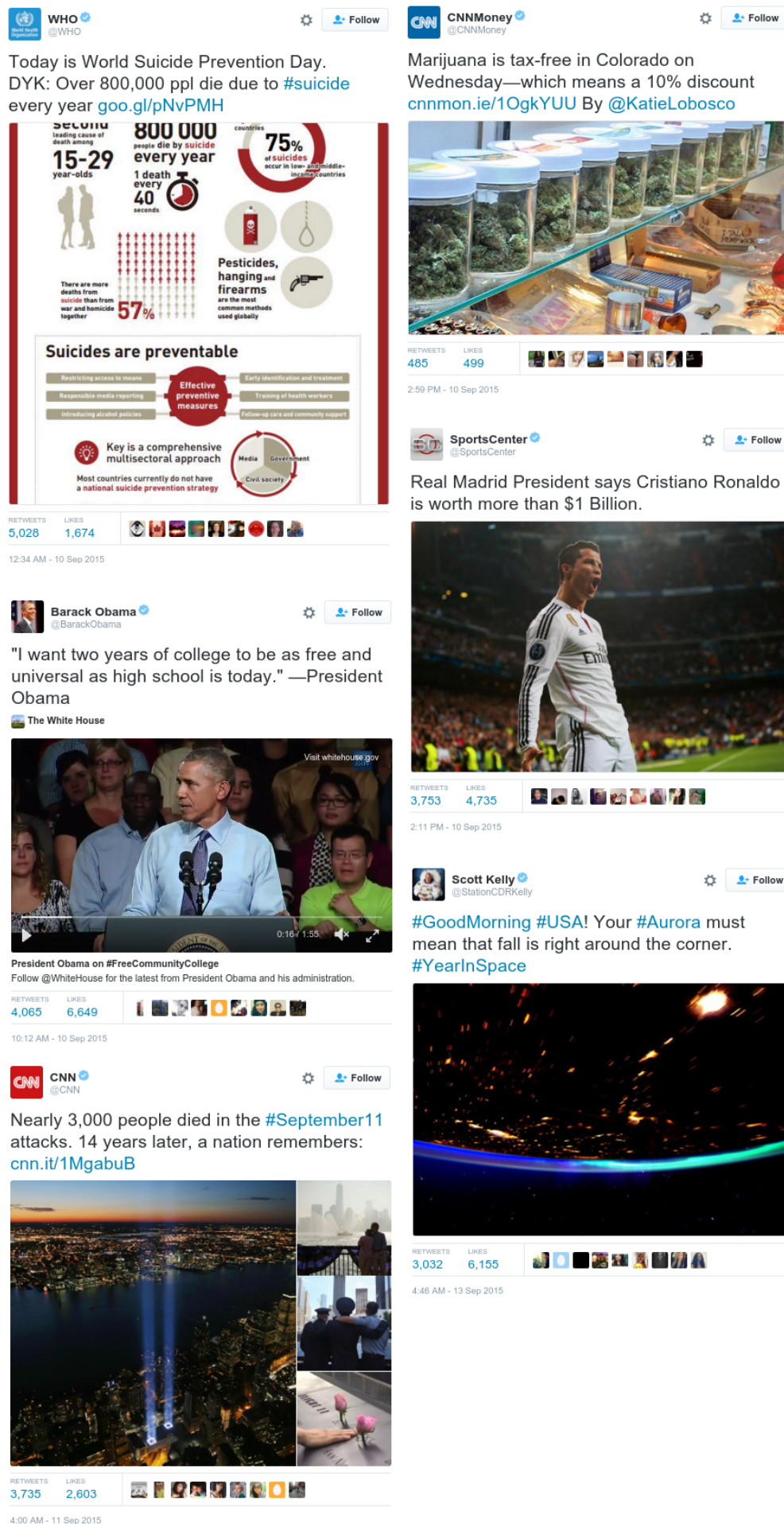


Figure 4.33: 1st to 6th most common saved Tweets. Order from left to right, and top to bottom.

hours. This was an obvious limitation for the “curious” search notifications which were fixed at the beginning of the study based on popular news gather before the start of the experiment. For instance, before the start of the study for the general category one popular topic was the commemoration about Marilyn Monroe death on the 5th August in 1962. However, very soon news about Marilyn Monroe in social media dropped to almost nothing during September. Although this research designed the search notifications to trigger relevant and recurrent news topics by carefully selecting keywords for each notification (e.g. commemoration of 9/11 for the general topic, champions league for the sports topic, popular hash-tags).

Rank	Twitter ID	Total Saves
1	641877384124305408	7
2	642095108775636992	4
3	642022723070201856	4
4	642082949815513089	4
5	642291495970697216	4
6	43027959432019968	4
7	642838975397269504	3
8	642300561245736960	3
9	643072390004518913	3
10	642054156505628672	3

Table 4.30: Top 10 most common saved Tweets

Table 4.31 presents that when the users of Newslab saved a social media news (the document), the average position of the document was 24th within the list of results. Also, the data shows that one user was so active in their exploration that saved a document in the 100th position of the ranked list. This means that the user had to scroll for quite sometime to get there. It seems that the scroll infinite feature made some participants unaware that they could search and explore other topics. Specially, as already highlighted, the data shows that the main search behaviour of users of the “cold” system was to browse until they found something interesting or got bored. There were also some people that saved the first result in the ranked list.

Mean document saved location	stddev	max	min
24.06	30.61	99	0

Table 4.31: Descriptive statistic for the position of documents saved during all search sessions.

4.5 Discussion

Self-reported measures as well as IIR metrics such as total dwelling time on task (TDT) and total estimated time on task were analysed and compared via non-parametric tests. Because of the

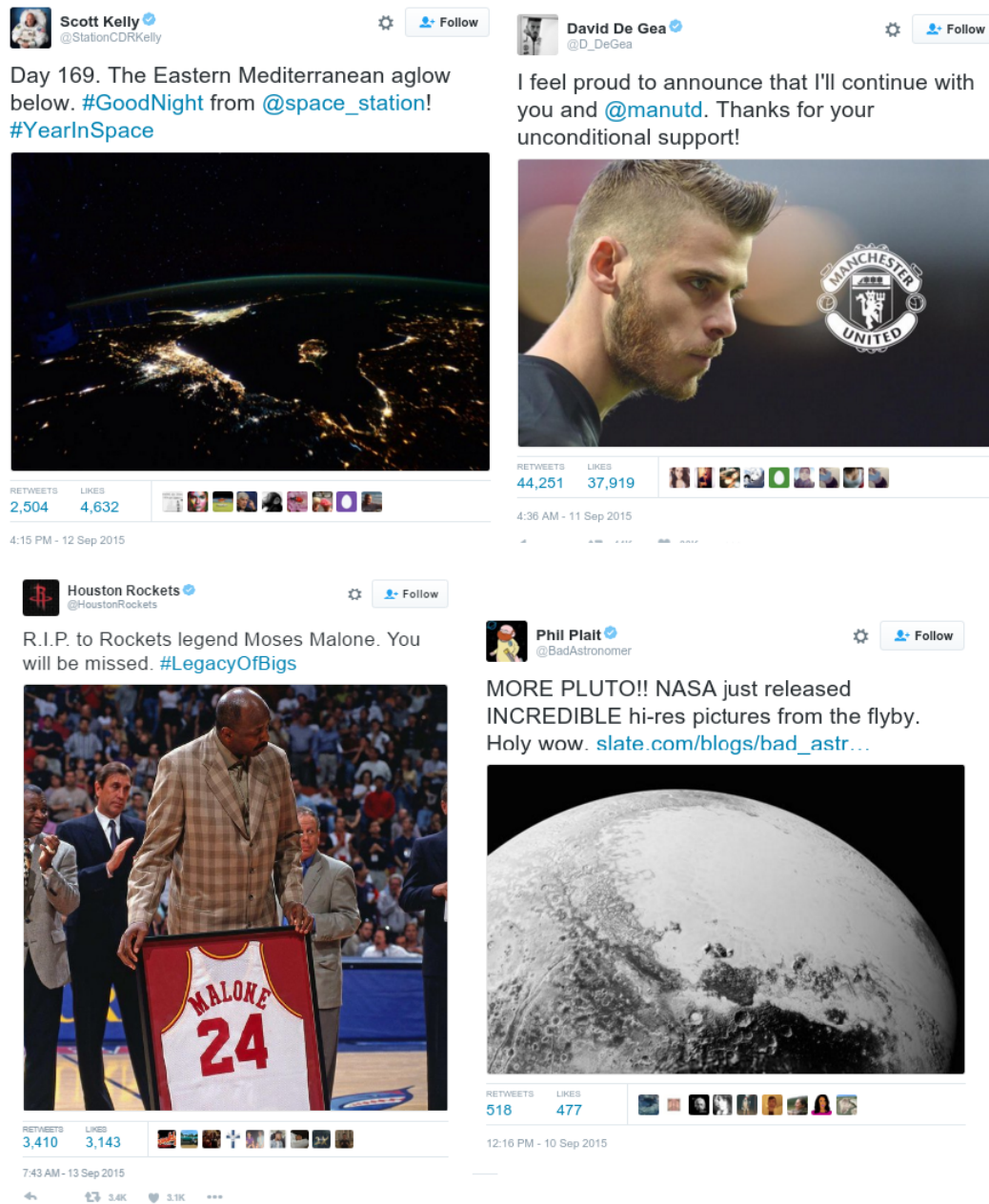


Figure 4.34: 7th to 10th most common saved Tweets. Order from left to right, and top to bottom.

small sample in the study, this research is cautious about making inferences and conclusions from the data. The results resemble many similarities in both psychometric ratings, search behaviour and qualitative comments about the search experience with the previous study described in chapter 3.

4.5.1 Does curiosity-driven search design increase users' engagement during casual-leisure search?

After validating the assumption that Task A and B did not influence strongly the search behavior and motivation of participants during the experiment. This research was able to reject the null hypothesis stating that participants exposed to the search notifications do not feel more novelty (NO) and surprise (UNEX) (proxies of curiosity) than those using 'cold' query-response paradigm.

Instead of creating NO and UNEX sub-scales as ad-hoc self-report instruments to measure UX without an established validity [28], NO psychometric scale was taken from O'Brien & Toms UES and UNEX(unexpected scale) was created by combining Pointers and Contrast serendipitous sub-dimensions from Björneborn [107]. Pointers and Contrast dimensions were merged due to the wording of each question and their relation with uncertainty and surprise. This psychometric scale measure how 'eye tempting' the content and interaction were during the search journey. Not other psychometric sub-scale had such a significant statistically difference between subjects and across tasks when manipulating system assignment variable ('Cold' vs 'Curious'). Thus, this research has enough evidence to state that the curiosity feeling evoked by the search notifications was an important positive factor in the perception of participants about their information interaction.

After analysing curiosity as personality trait, following Litman & Jimerson curiosity trait definition as a feeling of either "deprivation"(CFD) or "interest"(CFI) [19], this research did not find enough evidence that curiosity as a personality trait had a strong influence in the information seeking behaviour perceived during the experiment. Therefore, the curiosity feeling was not the result of a particular predisposition among the participants population. On the contrary, the comments and metrics suggests that the reason for experimenting curiosity was better explained by the independent variable of system assignment. This analysis brings another argument in favour of designing information interaction systems based on curiosity arousal principles such as novelty, partial exposure and surprise because it allows this research to claim that the curiosity construct can be evoked, independently of user's personality trait (CFD or CFI), by manipulating particular features in the IIR-system even when they are not explicitly

related to the topic content or personal interest.

During the study the other independent variable was personal interest (PI). The main hypothesis was that similar to O'Brien newsreaders study [118] and McCay-Peet et al.[121] where PI should predict and facilitate users' engagement and exploration. In the previous section, the results show that PI was a good predictor of engagement as defined by O'Brien et al. [117] and TDT.

Specially, in task B (the second). The results of PI in task B are interesting for leisure search research and the role of emotions in IIR because although both task were perceived in similar way, the result seem to suggest that after searching in NPI task A and perhaps being bored, participants personal interest (PI) had a more important effect when they were searching and evaluating the UX. In other words, the leisure search behaviour (or the desire to have fun) was reinforced in task B by previous exposure to a none personal interesting topic (NPI) in task A.

However, using Wilcoxon signed-rank nonparametric test, Novelty and UNEX sub-scales accepted the hypothesis of equal distribution for PI and NPI tasks because PI was not as good predictor of curiosity as system assignment. These findings support the general hypothesis that mapping curiosity arousal principles to the design of IIR-systems influence user's engagement (RQ1) [26].

4.5.2 Does curiosity-driven IIR-system increase users exploration during casual-leisure search?

An important analysis from the self reported data shows that personal interest (PI) and system assignment independent variable ('Curious' vs 'Cold') had an statistically significant effect in the total dwelling time (TDT) within subjects in this user study. Thus, a randomly chosen total time measurement from the 'curious' group was higher than a randomly chosen from the 'cold' group for both PI and NPI tasks. This result rejects the null hypothesis of equal distribution of TDT between 'curious' and 'cold' system assumed at the beginning of this study. Thus, curiosity arousal principles mapped into the search notifications increased the probability of higher TDT values during leisure news search on social media.

In addition, the nonparametric test gave system assignment independent variable a lower p value than PI variable for their prediction or effect in TDT. This result suggests that system assignment variable influence more the search behaviour than PI variable. The previous study, reported in chapter 3, show a similar relation between topic relevance and interactivity. In both studies, comments and psychometric metrics seem to suggest that due to the information interaction design the search experience itself can become more important than the topical

relevancy of individual results. In both studies, participants kept interacting, clicking, scrolling, and browsing because they were moved to do so by their curiosity or their “worry” for the unknown, and not due to their personal interest.

Although the interaction logs did not assess participants’ curiosity, they show a large amount of evidence that the “curious” search notifications in the IIR-system moved participants beyond “normal” search behaviour. The data describe that enabling “surprising” search notifications affected the search session time, the number of queries and the diversity of the queries. For instance, the data logs show that the mean number of queries for the search sessions with “Curious” IIR-system was almost three times more than the sessions with the “Cold” IIR-system (Figure 4.22). Other interesting pattern in data was the “Cold” query-response had a total of 34 unique queries and the “Curious” had a total of 233 unique queries (See Figure 4.23).

Why did users sometimes click but stop exploring after a few minutes like most of the “Cold” search sessions? Although different user’s traits should not be disregarded as a possible source of variation, other possibility is the personal interest (PI) variable which was manipulated before presenting one of the IIR-systems. Therefore, some searchers could have clicked the search notifications, felt some degree curiosity and engaged for a while but after getting information not relevant to their personal interest they quit exploring. Therefore, in those cases, the evoked curiosity produced by the external stimuli (i.e. search notifications) was not enough to trigger the intrinsic curiosity because they were assigned to a topic with little personal interest. For example, a female user who disliked a lot sport news and clicked in some notifications for a while but gave up after some minutes of exploring because not a single news was interesting for her. However, this is always better than other users who after scrolling down for some seconds left the web application because there was not a curious stimuli, nor they were interested.

In comparison to O’Brien (2011) [118] news reading study based on work task scenarios where most participants did not click in the link recommendation (23%), this user study shows that almost 38% of the search sessions with search notifications had more than one click event, and this action modified participants’ search behaviour. This data pattern in data logs was evident by the responses to the psychometric questionnaires as well. For instance, the descriptive statistic over the two questions: ‘Did you notice the Search Notifications or Dialogs?’ and ‘Did you click in the Search Notifications or Dialogs?’. 29 (87%) participants saw the notifications and 21 (71%) of those participants clicked on them.

The difference in search behaviour between O’Brien (2011) news reading study and Newslabels study (done as part of this PhD work) could either be due to the type of search scenario or the way the information interaction was designed. For this research, the most likely reason

for this difference in the search behaviour was the way in which the communication flow between the system and the user (e.i. the information interaction with push search notifications) and the affective state evoked by the search notifications (e.i. curiosity, or surprised) because Elswiler et al. [16] and Colbert & Boodoo [120] reported that in an experimental setting even if participants are instructed to browse and search an unlimited amount of time, they may still control their time in order to complete their task. A recent investigation presented by Gade & Hall (2016) [129], related to the field of book searching, suggested that there was not a significant difference between search behaviour during goal-oriented and casual leisure tasks across all participants. They detected an interesting difference in search behaviour between participants with a particular language skill (e.g. native English and non-native English speakers). However, further research should be done to compare the search behaviour between different search scenarios (e.g. task vs leisure), the type of search vertical (e.g. book search vs patent search vs social media news search) and the document type (e.g. tweet without photo, tweet with photo, tweet with gif animation, tweet with a video). These results consider in this section defend the hypothesis that mapping curiosity arousal principles to the design of IIR-systems influence exploration (RQ2).

4.5.3 Why did Focused Attention (FA) sub-scale was not different between system assignment groups as Ambiecities?

Although there was a similar interaction between the pop-up markers in Ambiecities and the notifications in Newslables, the role of perceptual bandwidth in a map-based information interaction (e.g. the possibility of drag the map in different location, zoom in and zoom out) could help individuals to feel more cognitive absorption (e.i Focused Attention (FA) in the O'Brien et al. [26] user's engagement framework).

Sundar (2015) suggest that "modality interactivity" (i.e. the variety of tools available on an interfaces such as sliders, drags, zoom, or mouse-overs) leads to greater absorption than message interactivity. For example, "when individuals operate a slider feature [...], the process requires users to adjust their motor response to drag their mouse from left to right, and perceptually code the visual changes according to their mouse movement[...]. During this process, their bandwidth for information uptake is expanded compared with the situation where they passively receive stimuli from media, because they can experience the dynamism of content by directly interacting with it" [127].

In a similar way, Ambiecities participant were more appreciative of the different ways of searching and exploring information enabled by the 'modality interactivity' when they evaluate

their search experience. In some of their comments, the participants establish a relation between the interactivity of medium and their cognitive absorption during the search experience (e.g. “Absorbing and made the time go quickly with the map interface”; “Yes, it was great being able to visit different cities and regions following the twitter traffic”).

Newslabels GUI did not offer such interactive tools that could expand the perceptual bandwidth of an individual while searching news. Perhaps, if the graphical time filters would have been designed as sliders and not buttons, individuals could have experience more cognitive absorption by adjusting the motor response to their dragging action and code the search results according to their mouse movement. So, the lack of modality interactivity in the information interaction could explain lower Focused Attention (FA) values in Newslabels than in Ambiecities.

Newslabels GUI did not offer such interactive tools that could expand the perceptual bandwidth of an individual while searching news. Perhaps, if the graphical time filters would have been designed as sliders and not buttons, individuals could have experience more cognitive absorption by adjusting the motor response to their dragging action and code the search results according to their mouse movement. So, the lack of modality interactivity in the information interaction could explain lower Focused Attention (FA) values in Newslabels than in Ambiecities.

4.6 Limitations

The work presented in this chapter described the efforts of research to design a methodology that overcomes the many difficulties involved in conducting casual-search in action. Although this research believes that the methodology represents a solid approach to studying casual-leisure search, some limitations remain.

The major limitation of the news search study is the size of the sample was small and that the study population consisted mainly of students. Although UK students may not be representative of all social media searcher, this research argues that the results are generalisable because all of them were active and regular users of the different social networks.

Another limitation of the work was the number of variables presented in the study. Other researchers have acknowledged the difficulties in performing IIR evaluations and one of the main difficulties is controlling the variables present in experimental designs[28]. In the news search study this research attempted to control the variables as much as possible, but as this research wanted to examine the effects of curiosity in an environment where user’s would feel free to perform the task without an observer and in the “wild”(from their own desktop computer), it was not possible to control all of the variables. For example, although the task descriptions

were the same, the participants were open to interpret the task in their own way, which may have meant looking for different information from another participant to solve the same task. Other limitation, it is the rate of change of content from social media streams. This meant that a top news in the application could change within minutes and that this research could not control or guarantee that all the participants had the same content. The only guarantee was the topic of the news.

Other limitation is that the only behavioural data recorded were total time on task, total queries, number of clicks and number of saved documents. For future studies other metrics such as mouse clicks, mouse movements, and scrolling events should be used. Future research should use physiological metrics (i.e., heart rate, electrodermal activity, electrocmytogram, eye movements, magnetic resonance imaging) because the information provided by them can increase the understanding of what it really means to be “Curious”. This PhD work investigated curiosity using behavioural data and psychometric questionnaires as proxy to evaluate and identify the existence and the degree of user’s curiosity.

In a similar way to Ambicites, total dwelling time on task results can be an indicator of both positive and negative experience. So, the behavioural data should always be analyzed in the correct context (i.e. the search scenario, the system, the user) and with other subjective information such as questionnaires or open questions.

It would have been interesting to gather search data of each participant over a long period of time, and understand if the effect of ‘surprising’ search notifications drops or increases over many search sessions with the same IIR-system. This could have been done by enabling sign in functionality into Newslab’s reading application and designing an “intelligent” search suggestions’ recommender, not a fixed recommender.

4.7 Implications

After the IIR experiment and the analysis of results, the implications for IIR are presented as follow.

4.7.1 On-line News Search on Leisure Scenarios

O’Brien (2011) demonstrated that Novelty, and Aesthetics were key factors in maintaining user engagement with on-line news content in none leisure scenarios [118] because of the capacity of the content to trigger curiosity (e.g. “headlines or content that evoked curiosity because they were outside the norm”) and the probability of noticing articles that participants could relate

to it personally whether or not the topic was new.

During this study of social media news search in leisure scenarios, curiosity was also evoked by both content and interactivity features such as the ‘curious notifications’. Participants comments and a between subject analysis of the results show that the system assignment variable (‘Curious’ vs ‘Cold’) affected in both task the Novelty and UNEX (Surprised) subscales suggesting that the system with push search notifications evoked participants curiosity even when the news were not from personally interesting topics. Thus, participants in the news leisure search scenario interacting with the ‘curious’ version were enticed to explore out of curiosity and performed an information searching behaviour similar to “Exploring for the experience” or “Needless Browsing” identified by Elswailer et al. [16] than those interacting with ‘cold’ system (i.e. look-up search, then moving on).

Drawing from the results, news search applications should be design to satisfy both the utilitarian and hedonic needs of their users. No just by thinking in how the content or a headline could trigger the users’ curiosity but designing an interactive two way communication channel that invites users to explore for more news. A communication channel that could make searchers “worry” about what they do not know and help them to do something about it [15]. An interactive two way communication channel as described in this experiment would also probably affect in a positive way the economic revenue of a news site by increasing the click rates, and dwelling times.

Comparing O’Brien (2011) [118] results and Newslabels study, this research finds an argument to highlight the power of mapping curiosity arousal principles from psychology in IIR-systems used for on-line news. As shown in the result section, the way in which the communication flows between the system and the user; and how this medium could be manipulated to change and encourage user’s exploration could be even as important as the content or headline of the story for the overall evaluation of the UX. For instance, twenty one (71%) of Newslabels participant clicked search notifications whereas seven (23%) participants clicked or used the recommended links in the study O’Brien (2011) [118]. The data logs also show that 38% (69) of all 182 search sessions with the push search notifications had more than one click event. This descriptive statistic must be interpreted by taking in consideration that several search sessions (or session’s ids) could be related to one participant, as already explained in the result section.

4.7.2 Affective IIR-systems

Humans interact emotionally with our machines. Thus, better emotional designs allow us to feel better and perhaps used them in better way. The study presented in this chapter shows how

evoking curiosity through ‘eye tempting’ and interactive notifications can motivate individuals to search and explore for more time in a leisure scenario than they planned.

Although leisure scenarios have not other purpose than their enjoyment, it is possible to imagine how other more serious search scenarios like learning could benefit from a design that triggers curiosity or “worry” for what is unknown.

In the physical world, emotions guide many of our actions and decisions. As shown both in Ambiecities and Newslables studies, human emotion could affect positively our decision making process when searching on-line. For instance, individuals could be moved to search and explore more because they are feeling curious due to some “eye tempting” stimulus than because of a conscious rational information need. As already mentioned, curiosity is closely related to emotional constructs such as fear, pleasure, boredom, and anxiety. As shown in both studies, understanding users’ curiosity and properly infusing the curiosity stimuli into the human-computer interaction process can potentially help IIR-systems to achieve persuasive goals.

Newslables notifications were completely random and prefix. In other words, the curious IIR-system presented was very simple and fixed curiosity companion. Moreover, most of the users perceived the notifications as inspiring, provocative, and compelling. Even some users’ thought that the notifications were adapted to their actions or their preference.

Future work should focus in defining practical metrics and models to estimate the curiosity of humans while searching and exploring based on psychological theory of curiosity and IIR research. Other future research should be the creation of artificial curious search companion that based in a short and long-term relationships could produce inquisitive recommendations and may incorporate information of the partner into its cognitive development.

4.8 Conclusions

‘Curious’ system information interaction encourage participants to be engage and keep interacting. Novelty, and UNEX sub-scale ratings suggest that the search notifications evoked participants curiosity in either PI and NPI tasks. Within PI and NPI tasks, participants exposed to curious system had a higher total dwelling time (TDT) than those using the cold system. Interaction logs showed that the curious group of sessions had more unique queries, more queries, more TDT and more used of time-filters. Therefore, the Newslables study supports the principal hypothesis of this thesis which is that mapping curiosity arousal principal to the design of IIR-systems increase user’s engagement (RQ1) and exploration (RQ2)

Personal interest (PI) was a good predictor of general user's engagement and TDT. Specially, in the task B. However, PI did not explain well Novelty and UNEX sub-scales as the independent variable of system assignment ('Cold' vs 'Curious'). Thus, showing that the perceived curiosity feeling was explain by the appearance of "eye tempting" notifications, and not personal interest of participants in a given topic.

Also, this research did not find enough evidence that curiosity as a personality trait had a strong influence in the information seeking behaviour perceived during the experiment. Therefore, the feeling of curiosity is explained by an external stimuli, not a personality trait among participants.

With the Newslables study, this research have presented more pragmatic evidence of how existing IIR-systems could apply and benefit from applying curiosity arousal mechanism to the design of the search experience even when they are based on the query-response paradigm and the results are display on a rank list. The application architecture and the technologies could be replicated easily across different search applications because the whole IIR-system was designed with open source technologies and open standards such as HTML5 Web Sockets to enable this two way communication channel.

Both Ambicities and Newslables studies have highlighted the importance of designing a two way channel of communication between the system and searcher; and how this medium could be manipulated to encourage users to explore and discover the unknown. The empirical evidence gather during these two studies suggest this medium can even be as important as the quality of the content for the overall evaluation of the UX. For instance, "eye tempting" notification can evoked curiosity and 'inspired' individuals to look for new things and find unexpected interesting information (i.e. to seek serendipitous discovery) when they are getting no results or they are about to leave the website because they have already happy with their "normal" search experience.

Future research should investigate the cumulative effect of multiple exposures over a period of time to a curious IIR-system and find how does the user's performance and enjoyment varies when compare with other users who used classical IIR-systems. With the help of affective computing and artificial intelligence fields, future IIR research should also aim to design IIR-systems that could estimate based on the user's interaction the best moment and content to display a search recommendation in order to enable more dwelling time and user's engagement.

Chapter 5

Conclusion

“The important thing is not to stop questioning. Curiosity has its own reason for existing.”

Albert Einstein ¹

This research started with aim of reflecting on IIR beyond classical utilitarian view point and understanding search as more than a ‘cold’ query-response process in order to discover evidence of other important factors that could contribute to better search experiences [26].

Through detailed analysis of two user studies and the design of two different search systems based on curiosity theory [1, 18, 19, 20, 8, 21] , this research established a set of findings and generalizable search behaviours in favour of the usefulness of incorporating curiosity principles such novelty, partial exposure and surprise in the design of IIR systems and their UX. Specially mapping those principles into the interactivity of the system, and not relaying simply in the content presented by the system.

The details of the results and conclusions in Chapter 3 and 4 echoes the idea that ‘I liked what I got’ from my search results could also imply ‘I am not worried about what I did not’[15]. In both experiments, this research was able to observe, reproduce and explain common casual or fun search behaviours [16] by considering hedonistic motivations such as curiosity.

5.1 Analysis of Research Questions

This PhD work aimed to develop a set of research findings that contribute to, and extend, existing research in IIR. The outline plan sought to identify and collect evidence of the usefulness of mapping curiosity arousal principles into IIR systems both leisure and why not- serious search scenarios

¹ Old Man’s Advice to Youth: ‘Never Lose a Holy Curiosity’, LIFE Magazine, 2 May 1955, p. 64

RQ1 Does curiosity-driven search design increase users' engagement [26] during casual-leisure search?

RQ2 Does curiosity-driven IIR-system increase users exploration [27] during casual-leisure search?

RQ3 How could casual-leisure search behaviour [16] be studied in a formal experimental setting?

Each of these research questions is now formally addressed in turn and related to the substantive content of the analysis chapters.

5.1.1 Does curiosity-driven search design increase users' engagement during casual-leisure search?

In both studies, Ambicities and Newslables, this research evaluated participants' search engagement and user experience in relation with metrics such as novelty (NO), focused attention (FA), aesthetics appeal (AE) from O'Brien and Toms [26]. Pointers and contrast serendipitous sub-dimensions from Björneborn [107] (the UNEX scale) were used in the second experiment to further understand the multidimensional nature of the search experience. The UNEX subjective metric had the intention to assess how surprising, unexpected and "eye tempting" was the interactivity defined by the IIR-system.

Chapter 3 explain that Ambicities' participants acknowledged that they might have explored beyond what they planned because of their curiosity feeling. Focused attention (FA) sub-scale values suggested that participants were in a state of deeper state of interaction with Ambicities rather than with the standard Twitter search interface. Novelty (NO) sub-scale scores were the highest sub-scale score for the Ambicities group. This high values were empirical evidence that participants were driven by curiosity, and strongly entice to keep exploring.

Chapter 4 show as well that Novelty psychometric sub-scale (NO) had a significant statistically difference between subjects and across tasks when manipulating system assignment variable ('Cold' vs 'Curious'). Thus, this research gather enough evidence to state that the feeling of curiosity evoked by the search notifications was an important positive factor in the perception of participants about their information interaction. The surprising, unexpected and "eye tempting" UNEX psychometric score posits that participants exposed to the search notifications system felt more curiosity than those using the 'cold' query-response paradigm.

Having the premise of low quality content extracted from real-time social media media streams and their opinions, this research also observed that Ambicities' participants felt cu-

rious specially not because of the content, but due to the push interactivity enable in both applications, created by using web sockets and design with curiosity arousal principles such as novelty, surprise and partial exposure. This observation became more evident in the second user study where the Personal Interest variable (PI) was manipulated. Although personal interest was a good predictor of general user's engagement and total dwelling time. Personal interest did not explain Novelty(NO) and UNEX sub-scales as well as the independent variable of system assignment ('Cold' vs 'Curious'). Thus, showing that the perceived curiosity feeling was explain by the appearance of "eye tempting" notifications, and not personal interest of participants in a given topic, nor a perfect "content match" that solved the participant's information needs.

These results contrast with previous researchers where personal interest (PI) was a good predictor of user's engagement in work scenarios [117]. These findings support the general idea that mapping curiosity arousal principles to the design of IIR-systems influence some user's engagement scales more than personal interest. In other words, system assignment was better predictor of curiosity (or its collative variables) than PI [26].

Ambicities and Newslables participants' comments about the search experience were very similar when evaluating their search experience with the "curious" system. When reflecting in either the incoming pop-ups in a map or the search notification, most participants stated that they were 'distracted', 'inspired' or 'moved' to look for new things, to stumble-upon interesting information, to keep exploring and clicking.

In Chapter 4, after analysing curiosity as personality trait, following Litman & Jimerson curiosity trait definition as a feeling of either "deprivation"(CFD) or "interest"(CFI) [19], this research did not find enough evidence that curiosity as a personality trait had a strong influence in the information seeking behaviour perceived during the experiment. Therefore, the curiosity feeling was not the result of a particular predisposition among the participants population. On the contrary, the comments and metrics suggests that the reason for experimenting curiosity was better explained by the independent variable of system assignment. This analysis brings another argument in favour of designing information interaction systems based on curiosity arousal principles such as novelty, partial exposure and surprise because it allows this research to claim that the curiosity construct can be evoked, independently of user's personality trait (CFD or CFI), by manipulating particular features in the IIR-system even when they are not explicitly related to the topic content or personal interest. Therefore, this research gather enough evidence from both experimental settings to show that users' engagement during casual-leisure search increased when participants searched with the curious system because of the curiosity design features rather than the participants curiosity trait or the topic of the search.

5.1.2 Does curiosity-driven IIR-system increase users exploration during casual-leisure search?

All the experiments followed an user-centred approach in a naturalistic scenario with mixed-methods perspective between IIR behavioural data and UX subjective self-reported variables[28]. This research used Total Dwelling Time (TDT), Guessed Total Dwelling Time (GTDT), and Absolute Error Mean Total Dwelling Time. In the second experiment, this research employed other interaction metrics such as number of queries issued, number of documents viewed, query length, number of saved documents, and number of unique terms in the query.

During the Ambiecities experiment, participants had higher recorded total dwelling time than those using the standard Twitter interface. This suggests the existence of one or more interactive features in Ambiecities design that triggered curiosity and made the experience engaging, and enticing for participants even though Twitter search interface had probably higher topically relevant content.

According to the results in chapter 3, the Ambiecities group had higher absolute error between perceived total time on task and recorded total time. Unlike in past research [113, 24], FA sub-scale values were on average higher and closer to the UES sub-scales mean value for the group interacting curiosity-driven IIR-system. FA sub-scale scores, TDT, and GTDT suggest that participants were in a state of deep interaction with Ambiecites and participants subjective responses showed that they might have explored beyond what they planned.

Therefore, Ambiecities' participants were enticed to explore out of curiosity and performed an information searching behaviour similar to "Exploring for the experience" or "Needless Browsing" identified by Elweiler et al. [16] than those interacting with classical query-response paradigm (i.e. look-up search, then moving on).

For the Newslables experiment, the recorded behavioural data (from the interaction logs and Surveygizmo tool) demonstrates that personal interest (PI) and system assignment independent variable ('Curious' vs 'Cold') had an statistically significant effect in the total dwelling time (TDT) within subjects in this user study. Thus, a randomly chosen total time measurement from the 'curious' group was higher than a randomly chosen from the 'cold' group for both PI and NPI tasks. This result rejects the null hypothesis of equal distribution of TDT between 'curious' and 'cold' system assumed at the beginning of this study. Thus, curiosity arousal principles mapped into the search notifications increased the probability of higher TDT values during leisure news search on social media.

In addition, the nonparametric test gave system assignment independent variable a lower p value than PI variable for their prediction or effect in TDT. This result suggests that system

assignment variable influence more the search behaviour than PI variable. The previous study, reported in chapter 3, show a similar relation between topic relevance and interactivity. In both studies, comments and psychometric metrics seem to suggest that due to the information interaction design the search experience itself can become more important than the topical relevancy of individual results. In both studies, participants kept interacting, clicking, scrolling, and browsing because they were moved to do so by their curiosity or their “worry” for the unknown, and not due to their personal interest.

Although the interaction logs did not assess participants’ curiosity, they show a large amount of evidence that the “curious” search notifications in the IIR-system moved participants beyond “normal” search behaviour. The data describe that enabling “surprising” search notifications affected the search session time, the number of queries and the diversity of the queries. For instance, the data logs show that the mean number of queries for the search sessions with “Curious” IIR-system was almost three times more than the sessions with the “Cold” IIR-system (Figure 4.22). Other interesting pattern in data was the “Cold” query-response had a total of 34 unique queries and the “Curious” had a total of 233 unique queries (See Figure 4.23).

The search applications based on curiosity arousal principles were able to move a number of participants to explore for more time than what they were planning to do so, by making them a bit “worried about what they didn’t know” or “making them curious” for the next piece of information [15].

This finding opens a door of new research and validates the importance of this PhD work because it suggests that it is not the user’s fault to stop exploring after having a ‘normal’ search experience [15], it is the fault of the IIR-system which is not able to trigger or evoke the affective state needed to keep exploring in a leisure scenario. Both IIR-systems used during this PhD work have successfully triggered such an affective state in most of participants by mapping curiosity arousal principles to the search experience. Thus, this affective state moved participants interacting with the curiosity driven search system to keep exploring and discovering more information.

5.1.3 How could casual-leisure search behaviour be studied in a formal experimental setting?

Study fun or ‘casual’ search behaviour in an experimental setting is not simple [16]. Drawing from Colbert & Boodoo’s conclusion regarding the limitations of an experimental set up [120], this research decided that an open-ended leisure search scenario was not enough to generate “casual-leisure” behaviour, or engaged participants.

Instead of only relaying in a “hedonistically motivated task”, this research sought to design and develop a search experience driven by an hedonic motivation, curiosity. The assumption was to let users interact with the ‘curious’ IIR-system and observe how naturally they were able to enjoy and how much they were distracted or entrained having in mind their current simulated scenario.

The changed in the approach meant that some participants showed casual leisure behaviours such as “Exploring for the experience”, “Needles Browsing” during the user studies [130]. In both studies participants comments and the interaction data show that the design of the IIR-system moved them to search, browse and explore because they were inspired or compelled to do it out of curiosity rather than an specific information need. For example, in Ambicities one participant mentioned:

“I felt like I was clicking on tweet’s markers because I could not help it, not because I thought I would find something interesting. I click because I was feeling noseey. After 20 clicks I realized I did not care about them but I still wanted to read them. The tweet was there; I had to look at it”. (Taken from section 3.6.4)

In Newslables another participant said:

“Popup boxes had me clicking on things and learning about new topics. However, this is the same reason I would never actually use this site, too much time-consuming. You feel compelled to keep clicking on popup boxes”. (Taken from section 4.4.7)

Thus, this research was able to reproduce search behaviours in an experimental setting similar to the ones described by Elswailer et al. where the hedonistic need was more important than any information need [16].

This methodological idea of moving the hedonistic perceived value from the search scenario description and protocol planning to the interaction with IIR-system built around curiosity (i.e. hedonistic need) was an important innovation presented to the IIR field of casual-leisure search in action and the key to the success of both experiments [130].

5.1.4 What do the findings mean respect to the search experience design?

Often curiosity is reflected as personality trait that influence search and exploratory behaviours. For instance, a recent user study of Bailey & Kelly (2016) [131] determine that curiosity was an important affective state to determine search expertise and search outcomes in none leisure scenarios. However, the importance of curiosity does not end there. The curiosity concept

could also be understood as an affective state evoked by an external stimuli that are novel (e.g., unexpected changes or violated expectations), conflicting (i.e., arousing two or more incompatible responses), uncertain (i.e., leading to outcomes that one is not sure about), and complex (e.g., presenting variety and diversity) [18].

Chapter 3 and chapter 4 explain how curiosity arousal principles were introduced to the search experience of both a map-based search application and a news query-response application. In both studies, the search applications based on curiosity arousal principles were able to moved a number of participants to explore for more time than what they were planning to do so, by making them a bit “worried about what they didnt know”[15] or “making them curious” for the next piece of information. Also the design of this curiosity driven search experience affected how participants’ evaluated their engagement and UX in relation with metrics such as novelty, focused attention, aesthetics from O’Brien and Toms [26], and pointers and contrast serendipitous sub-dimensions from Björneborn [107] (the UNEX scale). After the data analysis, these were the metrics with more statistically difference.

During the second study, this research did not find any statistically strong relationship between curiosity as personality trait, as described by Litman et al. [1, 18, 19, 20, 8, 21], and both the psychometric ratings of UX and the total dwelling time. Therefore, the fact that some psychometric measurements such as Novelty(NO) , pointers and contrast serendipitous sub-dimensions (UNEX), and total dwelling time in the task were statistically influenced by the type of IIR-system adds up to the argument of designing IIR-systems with curiosity arousal principles because the results suggest that it was the interaction with the ‘curious’ IIR-system, not the personality of the participant, what affected mostly the search behaviour, and the final outcome.

A successful form of search experience should mimic the way in which humans conduct face-to-face conversations because it is not just the content that it is important, but the way in which content might be presented and perceived is vital for effective human computer interaction or interactive information retrieval (IIR) in different scenarios. Search should be more than a findability problem [14].

This research has looked to the most powerful factor behind information seeking, and exploration, curiosity. There are a lot of research in the field of psychology in relation to human and animal curiosity mechanisms, and their effect in learning. In the recent years, computer science field such as robotics and artificial intelligence had started to used this concept to design better learning algorithms and robots [111].

This research applied curiosity arousal theory in the information interaction design of two

social media search applications in order to determine if taking curiosity arousal principles into account in the design of IIR systems has a positive effect in the user's engagement and exploration behaviours in a leisure scenario from both utilitarian and hedonic perspectives of the search experience.

The empirical data gathered during the two user studies (Ambicites and Newslab) show that mapping these curiosity principles in the interactivity and content presentation of the search applications evoked statistically more user's curiosity and exploration across different leisure search scenarios.

Although this PhD work did not measure each particular features of curiosity theory mapped in the design of the applications in isolation and their inter-relation between them. This research was able to observe that the sum of the different curiosity arousal mechanisms employed with careful design produce in most participants a more exploratory search pattern and moved them to look beyond what they were planning to do so. Therefore, IIR-systems will benefit from incorporating curiosity arousal principles such as novelty, surprise and partial exposure in order to provide a better search experience.

Also in both applications, this research used HTML5 Web Sockets and created both push and pull information interaction. From the described user studies in chapter 3 and 4, the results suggest that IIR-systems should start use protocols of communication like WebSockets or similar ones to create UX where the stream of communication follows both directions, pull and push. Although there are several technical challenges related to the use of these protocols in web applications, these protocols create a better communication medium. Otherwise, the capacity of the system to 'entice' or 'pick' interest is reduced; then the whole UX relies only in the elusive content (the destination) and disregards interactivity, a big part of the communication medium between the system and the user.

With push interactivity and curiosity arousal principles in a search application, the searcher utilitarian success is not guaranteed, but the experience of a more vivid information journey could motivate a user to explore in depth and go beyond the first search page of results. For instance, "eye tempting" notification can evoke curiosity and 'inspired' individuals to look for new things and find unexpected interesting information (i.e. to seek serendipitous discovery) when they are getting no results or they are about to leave the website because they have already been happy with their "normal" search experience.

In general, the PhD work proposes that IIR field must start paying more attention on how to support affective states as well cognitive states such as curiosity in the design of IIR-systems. As explained in chapter 2, IIR field, biased its origins in LIS, has concentrated mainly

in the cognitive state of searchers and the findability problem. Thus, this research suggest that IIR research should forget, at least for while, content or information objects to create a more effective interaction and communication. Although this might sound as an exaggeration, IIR should put more effort to enhance the medium of communication between a searcher and IIR-system rather than look for a perfect match of information needs, specially in leisure and none leisure scenarios such as learning where affective states could potentially drive a more engaging search experience.

5.2 Limitations

The user studies had an important weakness regarding the sample size and in terms of the recruited subjects (i.e. most of them students).

This PhD work did not assess other personality factors beside curiosity as a trait [1, 18, 19, 20, 8, 21]. For instance, some other personality traits could be very important in order to be in ‘flow’, or reach a high level of focused attention [23].

Most of the evidence presented by this research is based on psychometric measurements, which are a reflection of the perception of participants and their opinions. Search interaction data was also used as proxy to identify and estimate the degree of curiosity. However, there could be other metrics to understand and investigate better the whole concept of curiosity.

Regarding the total time on task results and the interpretation of this metric, this research has shown that total time can be an indicator of both positive and negative experience. So, the behavioural data should always be analyzed in the correct context (i.e. the search scenario, the system, the user) and with other subjective information such as questionnaires or open questions.

This PhD work has failed to isolate specific curiosity arousal principles in the search experience, measure them and understand their inter-relationship. Future research could define a smaller granularity measurements to test specific search features within an IIR-system and measure how each feature affect the psychometric evaluation. For instance regarding the search notifications, future work could measure the frequency of the notifications, the duration in the screen, the location where they appear on the screen, the colors, their shape, and evaluate how do small variations evoke or hinder the needed affective state to explore in depth and make valuable discoveries.

This research did not to evaluate the implementation or mapping of each curiosity principles in the design of the search experience because the main focused for this doctoral work was to

evaluate the complete search experience and to test whether all the small features sum up and influence the searchers behaviour or the search outcome. This research elaborated the mapping of curiosity arousal principles by doing an extensive literature review in UX, IIR and psychological fields; then when the concepts had been acquired, with creative thinking and without any formal process both applications were designed. Future research should continue work in how this creative process should be structure and the mapping of curiosity arousal principles operationalize.

5.3 Contributions

1. **First Contribution:** *An argument for, and demonstration of, the use of curiosity arousal principles in the information interaction design of search applications to improve user's engagement and exploration.*

Based on two empirical studies in different leisure contexts, this research has been able to argue that mapping novelty, uncertainty and partial exposure concepts from Psychology into the design of the search experience evoked user's curiosity and moved them to explore more than what they "normally" do with standard query-response information interaction paradigm.

In the first study, this research compare psychometric rating of UX scales and total time in a search task between two applications: a map based curious search application and Twitter search functionality. The results showed that user's exposed to the map-based interface had higher total dwelling time (TDT), novelty (NO), and focused attention (FA) than the Twitter group [26]. Thus, rejecting the null hypothesis of equal distribution or no statistically significant effect of curiosity arousal principles in TDT, NO and FA. This results were presented partially in the Information Interaction in Context Symposium (August, 2014) [130]. Chapter 3 explains in detail the methodology and discusses the results.

For the second study, this research created a news search application based on social media with two possible information interaction modes: 'cold' query-response and 'curious' query-response by allowing pushing randomly appealing search notifications. The results showed that search notification designed with the same collative curiosity variables (i.e. novelty, uncertainty and partial exposure) evoked user's curiosity. Therefore, participants exposed to the search notifications had statistically higher values for total dwelling time on task within subjects, NO, and UNEX sub-scales between subjects than those inter-

acting with the standard query-response [26, 107]. Again, providing evidence to reject the null hypothesis assumed at the begging of this research of not statistically significant effect of curiosity arousal principles in the search experience.

In the second study, this PhD recorded an interaction log that demonstrated that participants who click the search notifications were more likely to elaborate more queries and, explore for more time than those who did not click. The logs also show that the queries, and therefore, the search sessions of participants exploring with the IIR-system with search notifications were more diverse than the ones made by users of the baseline system. Chapter 4 describes the experimental setup, the interaction information design, the results and their implications for IIR and leisure search field. The outline contribution was created from RQ1 and RQ2.

In both studies the psychometric scales and TDT measurements were supported and explained by the subjective comments of participants.

2. Second Contribution: *Two experimental setups to study real leisure search behaviour in action.*

According to Elsweiler et al. [16] user's with a strong hedonistic motivation follow search behaviours that break traditional IIR models after gathering user's self-reported search experience on social media and studying people's information seeking behaviours when using television. During these search information behaviour session, users might find topically irrelevant (e.g. low content) information but they may keep exploring because the system satisfies their current leisure need, or they might keep exploring because every result they get has content so curious and funny that they want to keep exploring even beyond what they have already plan.

However to study this kind of behaviour in a experimental setting is not easy. For example, Colbert & Boodoo [120] reported that, in experimental environment, even if participants are instructed to browse and search an unlimited amount of time, they may still control their time in order to complete their task and make difficult to establish statistical difference in the user's engagement between two websites. Therefore, this research concluded that an open-ended leisure search scenario was not enough to generate "causal-leisure" behaviour.

As mentioned in the methodology section, Wilson & Elsweiler [31] proposed the creation of an experimental setting where "users are provided with hedonistically motivated task" by making a faux-study (e.g. in a laboratory setting, users are told there is an unforeseen

delay or problem, and instructed to use a computer while they wait. Then, after some time when they seem to be bored, then the study continues). But there has not reported any recent research in casual-leisure search applying this methodology.

Instead of designing the experiment around the creation of a “truly” hedonistic motivation, this research created the casual-leisure experiment around two assumptions: 1) search has some hedonic and utilitarian aspects related both to content and interaction that can trigger positive or negative emotions. Those ‘casual’ emotions could be the drivers of user experience that reflects what people normally do in order to get enjoyment; 2) the search experience can be design to generate or encourage the existence of hedonistic need such as curiosity. This change in the methodological approach was vital to the success and viability of this experiments (RQ3).

Then this research hypothesized that participants exploring with an IIR-system built mapping concepts of curiosity arousal mechanism from psychological theory into the search experience could cause casual-leisure search behaviours such “Exploring for the experience”, “Needless Browsing” in a experimental setting. According to the literature review conducted by this research, there has not been presented in IIR a study of casual-leisure search in action before the paper presented in the Information Interaction in Context Symposium (August, 2014) [130].

This methodological shift meant moving the hedonistic perceived value from the search scenario description and protocol planning to the interaction with IIR-system built around an hedonistic need (i.e. curiosity). Following the aforementioned assumption of UX that any product has pragmatic and hedonistic aspects [119]. Therefore, if the IIR-system highlighted the hedonic aspects of information searching, may be users would find the needed motivation to keep exploring for more time.

For this research, there were two ways of evoking curiosity in IIR-system: content and interactivity. For both experiments, these research choose to focus in the interactivity. On the other hand, the content was similar in the sense that the source of information was Twitter. As a result, various participants using Ambicities, and Newslabls showed similar search behaviours as the ones reported by Elsweller et al. [16] where users explore out of curiosity, and for more time than they had previously estimated. This finding opened another methodological door to study casual-leisure search in action, and it is one of the novel contributions of this PhD work. Although the remaining challenge would be to define a set of recommendation to make the experimental setting reproducible in other

IIR search scenarios such as learning.

5.4 Future Work

This previous section summarised the main contributions of this doctoral work. The ramifications of this work are notable and warrant further investigation. Many avenues emerged for the research described to be taken further. This section describes the main opportunities for progressing this PhD work.

5.4.1 Understanding the Role of Curiosity in IIR

As discussed in chapter 2, curiosity is an old, yet critical, concept in the psychology of motivation. Various exploratory or information-seeking behaviours have been deemed as curiosity. For instance, animals' orienting response (i.e., their immediate response to changes in their environment, such as change in illumination or unusual sound) is considered "perceptual curiosity", whereas humans' desire for information and knowledge is categorized as "epistemic curiosity" [1, 18, 36].

This PhD work shows how the concept of curiosity could be operationalized in the design of IIR-system to support more exploration and user's engagement. Although exploration and user's engagement are itself important concepts in IIR and for human well-being, these are just behavioural proxies of the existence of this affective state or hedonistic motivation in organism whose final goal is the acquisition of knowledge and the maximization of learning[36]. Therefore, perhaps the greatest goal of design IIR-systems' based on curiosity arousal principles would be to enhance learning, not simply the modification of the search behaviour.

Curiosity is a motivational factor, which affects memory. Research in education and psychology shows that by incorporating curiosity into ways of teaching and learning, the memory of students can be enhanced and education can be improved. For example, Kang et al. (2009) [132] scanned subjects with fMRI while they were reading trivia questions. The subjects level of curiosity when reading questions was correlated with activity in the brain where previous research involved anticipated reward or encoding prediction error. These discoveries led to a behavioral study showing that subjects spent more finite resources (e.g waiting time) to discover answers when they were more curious. The fMRI also showed that curiosity increases activity in memory areas when subjects guess incorrectly, which suggests that curiosity may enhance memory for surprising new information. This estimation about memory enhancement was confirmed in a behavioral study where higher curiosity in the initial session was correlated with

better recall of surprising answers 10 days later. In a more recent paper, Gruber et al. (2014) [128] found association between the mechanisms supporting extrinsic reward motivation and intrinsic curiosity, and highlight the importance of stimulating curiosity to create more effective learning experiences. This research demonstrated that curiosity not only modulates memory circuits in the human brain, but enhances the discovery of incidental information. In other words, this suggests that within a learning scenario if an extrinsic mechanism is used to trigger curiosity simultaneously even when unrelated to the goals of the scenario, the mechanism has the potential of boosting the subjects intrinsic motivation and improve the subject's learning process.

Therefore, future research in IIR should measure the possible effect on learning and memory retention of subjects when designing curiosity driven search experiences based novelty, uncertainty, partial exposure and surprise [1, 18, 36]. How could unrelated curiosity arousal mechanism, such search notifications in IIR-system, improve the intrinsic curiosity motivation in the searching process?

Despite the importance of information seeking for animals and humans, our understanding of its mechanisms is in its infancy. Future research in IIR should seek to understand the role of curiosity together with other fields such as psychology and neurology.

5.4.2 Affective IIR-systems. Designing, Implementing and Evaluating Curiosity Driven Search Tools

This thesis has increased the understanding the research community has about how curiosity could increment users' engagement and exploratory behaviour. However, the implementation of the curiosity arousal principles in IIR-systems has been limited to the design of the information interaction and the user's perception.

For instance, in the second experiment, this research added the concept of companion agent who would recommend search notifications to searchers with the objective of evoking participants curiosity. However, the implementation of this companion agent did not use any of the existing artificial curiosity models build in robotics or virtual learning environments, neither track user's interaction to predict the right time to give a recommendation [111]. The complete behaviour of the agent was random within certain range of settings (making the agent really 'dumb'). Therefore, a next step on this research should be making real "curious" companions (at least one more intelligent), not only relying in the human perception and UX design to trigger curiosity after a search notifications is display when the user did not expected.

This models could employ users' interactions such as mouse movements, scrolling and queries. Then, the presentation or push notification mechanism could personalize and make context-aware to send notifications with the right content (at least the popular ones or the ones the person his more likely to enjoy) and in the best 'arousal' moment of the search exploration. Perhaps when users are disengaging from the search experience or leaving the website [28].

Further, more work needs to be done in understanding how IIR-systems should be design to allow us to feel better and perhaps used them in better way. An increased understanding of how to this mapping affects searches behaviour will provide information architects, and UX designers with a better idea of the tools they should be creating. The doctoral work presented in this thesis shows how evoking curiosity through 'eye tempting' and interactive notifications can motivate individuals to search and explore for more time in a leisure scenario than they planned. Although leisure scenarios have not other purpose than their enjoyment, it is possible to imagine how other more serious search scenarios like learning could benefit from a design that triggers curiosity or "worry" for what is unknown.

5.4.3 IIR evaluation in leisure scenarios

The work presented in this thesis concerning the facilitation of casual-leisure search study and evaluation represents an important contribution to the field of IIR that will enable more IIR casual-leisure studies to be performed in the future. However, there is still a lot of work that could be done to improve on the experimental setup. Again, the work in this thesis focused purely on evaluating the search experience and behaviour for information objects coming from social media within open-ended simulated scenarios. In order to broaden the usefulness of the setup strategy work needs to be done to create taxonomies or frameworks for different kinds of information objects and simulated search scenarios.

Further work should focus in creating valid leisure scenarios by controlling curiosity arousal mechanisms and measuring their effect in other aspects of the information seeking process. For instance, in a similar posteriori research to this doctoral work, Law et al (2016) [133] examine the potential for curiosity as a new type of intrinsic motivational driver to incentivize crowd worker. Their main idea was to embed curiosity-inducing designs in crowdsourcing task interfaces to improve worker engagement and performance. Then they conducted a set of experiments on Amazons Mechanical Turk and their results demonstrate that curiosity interventions improve worker retention without degrading performance. As IIR-systems should also support learning within an information seeking process, one possible interesting future experiment should evaluate how curiosity driven UX or other possible designs for leisure scenarios could increase

memory retention both short term and long term.

Another avenue of research should seek to define practical metrics and models to estimate or predict the curiosity of humans while searching and exploring based on psychological theory of curiosity, IIR and Computer Science research. For example, Menk and Sebasti   [134] have focused on predicting curiosity by using information extracted from users’ profiles on Facebook and a set of features that can be used to describe their degree of curiosity. Also Wu et al. summarized various approaches that could be used to create computational curiosity models and systems [111]

5.5 Conclusion

This chapter has concluded this thesis by summarising the contributions made and proposing opportunities to further the research. The findings of this doctoral work have fundamental implications for the design of IIR-systems and their evaluation in leisure scenarios. This work has shown IIR-systems can be improved by considering curiosity as an affective state that could be evoked in the design of the UX. Further, the casual leisure experimental setup propose in the thesis provides a means through which researchers can assess ‘casual’ leisure search “in action” and the role of curiosity in the information seeking process.

Bibliography

- [1] Jordan Litman. Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & Emotion*, 19(6):793–814, 2005.
- [2] T.D. Wilson. Models in information behaviour research. *Journal of Documentation*, 55(3):249–270, 1999.
- [3] Carol Collier Kuhlthau. Seeking meaning. a process approach to library and information services. pages XVII, 247, 2004.
- [4] Kalervo Järvelin and T. D. Wilson. On conceptual models for information seeking and retrieval research. *Inf. Res.*, 9(1), 2003.
- [5] Peter Ingwersen. Interactive information seeking, behaviour and retrieval. edited by ian ruthven and diane kelly. london: Facet publishing, 2011, 296 pp. \$ 89.95 (paperback). (isbn: 978-1-85604-707-4). *JASIST*, 63(10):2122–2125, 2012.
- [6] Ayse Goker, Hans Myrhaug, and Ralf Bierig. Context and information retrieval. *Information Retrieval: Searching in the 21st Century*, pages 131–57, 2009.
- [7] Kazutoshi Umemoto, Takehiro Yamamoto, and Katsumi Tanaka. Scentbar: A query suggestion interface visualizing the amount of missed relevant information for intrinsically diverse search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '16, pages 405–414, New York, NY, USA, 2016. ACM.
- [8] Jordan A. Litman. Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, 44(7):1585 – 1595, 2008.
- [9] Jordan A Litman. Relationships between measures of i-and d-type curiosity, ambiguity tolerance, and need for closure: An initial test of the wanting-liking model of information-seeking. *Personality and Individual Differences*, 48(4):397–402, 2010.

- [10] Gary Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41–46, April 2006.
- [11] Ryen W White and Resa A Roth. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1):1–98, 2009.
- [12] Robert S Taylor. The process of asking questions. *American documentation*, 13(4):391–396, 1962.
- [13] Jaime Teevan, Christine Alvarado, Mark S. Ackerman, and David R. Karger. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 415–422, New York, NY, USA, 2004. ACM.
- [14] Tony Russell-Rose and Tyler Tate. *Designing the Search Experience: The Information Architecture of Discovery*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2013.
- [15] Karen Sparck Jones. *Retrieving information or answering questions?* British Library, 1990.
- [16] David Elsweiler, Max L Wilson, and Brian Kirkegaard Lunn. Understanding casual-leisure information behaviour. *Library and Information Science*, 1:211–241, 2011.
- [17] Steve Jobs. You’ve got to find what you love, Commencement address to Stanford University USA (Online transcript) Stanford Report., 2005. [Online; accessed 12/12/2014].
- [18] D.E. Berlyne. *Conflict, Arousal and Curiosity*. McGraw-Hill, 1960.
- [19] Tiffany L. Jimerson and Jordan A. Litman. The measurement of curiosity as a feeling of deprivation. *Journal of personality assessment*, 82(2):147–157, 2004.
- [20] Jordan A. Litman and Mark V. Pezzo. Dimensionality of interpersonal curiosity. *Personality and Individual Differences*, 43(6):1448–1459, 2007.
- [21] Todd B Kashdan, Paul Rose, and Frank D Fincham. Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment*, 82(3):291–305, 2004.
- [22] Merriam-Webster Online. Merriam-Webster Online Dictionary, 2009.

- [23] Mihaly Csikszentmihalyi. Flow and the psychology of discovery and invention. *Harper-Perennial, New York*, 1997.
- [24] Heather L. O'Brien and Elaine G. Toms. Examining the generalizability of the user engagement scale (ues) in exploratory search. *Inf. Process. Manage.*, 49(5):1092–1107, September 2013.
- [25] Tefko Saracevic. Evaluation of evaluation in information retrieval. In *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '95, pages 138–146, New York, NY, USA, 1995. ACM.
- [26] Heather L. O'Brien and Elaine G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6):938–955, April 2008.
- [27] Jon Pearce, Shanton Chang, Basil Alzougool, Gregor Kennedy, Mary Ainley, and Susan Rodrigues. Search or explore: do you know what you're looking for? In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, pages 253–256. ACM, 2011.
- [28] Diane Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1-2):1–224, 2009.
- [29] Pia Borlund. The iir evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3), 2003.
- [30] Bookshelf - adaptive software development: A collaborative approach to managing complex systems extreme programming explained: Embrace change, software process quality: Management and control. 2000.
- [31] Max L Wilson and David Elswailer. Casual-leisure searching: the exploratory search scenarios that break our current models. *HCIR 2010*, page 28, 2010.
- [32] Maristella Agosti, Norbert Fuhr, Elaine Toms, and Pertti Vakkari. Evaluation Methodologies in Information Retrieval (Dagstuhl Seminar 13441). *Dagstuhl Reports*, 3(10):92–126, 2014.
- [33] Marc Hassenzahl. User experience and experience design. *The Encyclopedia of Human-Computer Interaction*, 2nd Ed., 2013.
- [34] Thomas D. Wilson. Human information behavior. *InformingSciJ*, 3:49–56, 2000.

- [35] Marti A. Hearst. *Search User Interfaces*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.
- [36] George Loewenstein. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1):75, 1994.
- [37] Robert P Collins, Jordan A Litman, and Charles D Spielberger. The measurement of perceptual curiosity. *Personality and individual differences*, 36(5):1127–1141, 2004.
- [38] W. Wundt. *Grundzüge der physiologischen Psychologie*. Engelmann, 1874.
- [39] Charles D Spielberger and Laura M Starr. Curiosity and exploratory behavior. *Motivation: Theory and research*, pages 221–243, 1994.
- [40] Kent C Berridge, Terry E Robinson, and J Wayne Aldridge. Dissecting components of reward:liking, wanting, and learning. *Current opinion in pharmacology*, 9(1):65–73, 2009.
- [41] Jordan A. Litman, Robert P. Collins, and Charles D. Spielberger. The nature and measurement of sensory curiosity. *Personality and Individual Differences*, 39(6):1123–1133, 2005.
- [42] Brenda Dervin and Michael Nilan. Information needs and uses. *Annual review of information science and technology*, 21:3–33, 1986.
- [43] Karen E Fisher, Sanda Erdelez, and Lynne McKechnie. *Theories of information behavior*. Information Today, Inc., 2005.
- [44] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. pages 2–4, 2005.
- [45] Robert H MacArthur and Eric R Pianka. On optimal use of a patchy environment. *American Naturalist*, pages 603–609, 1966.
- [46] E. L. Charnov. Optimal foraging, the Marginal Value Theorem. *Theoretical Population Biology*, 9(2):129–136, April 1976.
- [47] Pamela Effrein Sandstrom. An optimal foraging approach to information seeking and use. *The library quarterly*, pages 414–449, 1994.
- [48] Peter Pirolli and Stuart Card. Information foraging in information access environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 51–58, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

- [49] Peter Pirolli and Stuart Card. Information foraging. *Psychological review*, 106(4):643, 1999.
- [50] K. Bystrom and Kalervo Järvelin. Task complexity affects information seeking and use. *Inf. Process. Manage.*, 31(2):191–213, 1995.
- [51] Katriina Byström and Preben Hansen. Work tasks as units for analysis in information seeking and retrieval studies. In *Proceedings of the 4th International Conference on Conceptions of Library & Information Science (CoLIS4)*, pages 239–252, 2002.
- [52] Reijo Savolainen. Everyday life information seeking: Approaching information seeking in the context of way of life. *Library & information science research*, 17(3):259–294, 1995.
- [53] Reijo Savolainen. *Everyday information practices: a social phenomenological perspective*. Scarecrow Press, 2008.
- [54] Robert Stebbins. Serious leisure and work. *Sociology Compass*, 3(5):764–774, 2009.
- [55] Robert A Stebbins. *Serious leisure: A perspective for our time*, volume 95. Transaction Publishers, 2007.
- [56] Robert A Stebbins. Serious leisure. *Society*, 38(4):53–57, 2001.
- [57] Jenna Hartel. The serious leisure frontier in library and information science: hobby domains. *Knowledge organization*, 30(3-4):228–238, 2003.
- [58] Catherine Sheldrick Ross. Finding without seeking: The information encounter in the context of reading for pleasure. *Inf. Process. Manage.*, 35(6):783–799, November 1999.
- [59] David Elswailer, Stefan Mandl, and Brian Kirkegaard Lunn. Understanding casual-leisure information needs: A diary study in the context of television viewing. In *Proceedings of the Third Symposium on Information Interaction in Context, IiiX '10*, pages 25–34, New York, NY, USA, 2010. ACM.
- [60] Peter Ingwersen and Kalervo Järvelin. *The Turn: Integration of Information Seeking and Retrieval in Context (The Information Retrieval Series)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
- [61] T.B. Hahn and C.P. Bourne. *A History of Online Information Services, 1963-1976*. 2003.
- [62] G. Salton. *The SMART Retrieval SystemÔExperiments in Automatic Document Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1971.

- [63] Robert N Oddy. Information retrieval through man-machine dialogue. *Journal of documentation*, 33(1):1–14, 1977.
- [64] Nicholas J Belkin, Robert N Oddy, and Helen M Brooks. Ask for information retrieval: Part i. background and theory. *Journal of documentation*, 38(2):61–71, 1982.
- [65] Nicholas J Belkin, Thomas Seeger, and Gernot Wersig. Distributed expert problem treatment as a model for information system analysis and design. *Journal of Information Science*, 5(5):153–167, 1982.
- [66] Nicholas J Belkin, Colleen Cool, Adelheit Stein, and Ulrich Thiel. Cases, scripts, and information-seeking strategies: On the design of interactive information retrieval systems. *Expert systems with applications*, 9(3):379–395, 1995.
- [67] Tefko Saracevic. Modeling interaction in information retrieval (ir): A review and proposal. In *Proceedings of the ASIS annual meeting*, volume 33, pages 3–9. ERIC, 1996.
- [68] Tefko Saracevic. The stratified model of information retrieval interaction: Extension and applications. In *PROCEEDINGS OF THE ANNUAL MEETING-AMERICAN SOCIETY FOR INFORMATION SCIENCE*, volume 34, pages 313–327. LEARNED INFORMATION (EUROPE) LTD, 1997.
- [69] Gary Marchionini. *Information Seeking in Electronic Environments*. Cambridge University Press, New York, NY, USA, 1995.
- [70] Donald A. Norman. *The Design of Everyday Things*. Basic Books, New York, reprint paperback edition, 2002.
- [71] Marcia J Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5):407–424, 1989.
- [72] Marcia J. Bates. Where should the person stop and the information search interface start? *Inf. Process. Manage.*, 26(5):575–591, October 1990.
- [73] Timothy G. Armstrong, Alistair Moffat, William Webber, and Justin Zobel. Improvements that don’t add up: Ad-hoc retrieval results since 1998. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM ’09, pages 601–610, New York, NY, USA, 2009. ACM.
- [74] Tefko Saracevic. Relevance: A review of the literature and a framework for thinking on the notion in information science. part iii: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):2126–2144, 2007.

- [75] Linda Schamber, Michael Eisenberg, and Michael S. Nilan. A re-examination of relevance: Toward a dynamic, situational definition. *Inf. Process. Manage.*, 26(6):755–776, November 1990.
- [76] Linda Schamber and Judy Bateman. User criteria in relevance evaluation: Toward development of a measurement scale. In *Proceedings of the ASIS Annual Meeting*, volume 33, pages 218–25. ERIC, 1996.
- [77] Tefko Saracevic. Relevance: A review of the literature and a framework for thinking on the notion in information science. part ii: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):1915–1933, 2007.
- [78] Katriina Byström and Preben Hansen. Conceptual framework for tasks in information studies: Book reviews. *J. Am. Soc. Inf. Sci. Technol.*, 56(10):1050–1061, August 2005.
- [79] Don Norman. Emotion & design: Attractive things work better. *interactions*, 9(4):36–42, July 2002.
- [80] Hannu Korhonen, Juha Arrasvuori, and Kaisa Väänänen-Vainio-Mattila. Analysing user experience of personal mobile products through contextual factors. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*, MUM '10, pages 11:1–11:10, New York, NY, USA, 2010. ACM.
- [81] Peter Morville and Jeffery Callender. *Search Patterns: Design for Discovery*. O'Reilly Media, Inc., 1st edition, 2010.
- [82] Paul André, m.c. schraefel, Jaime Teevan, and Susan T. Dumais. Discovery is never by chance: Designing for (un)serendipity. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, C&C '09, pages 305–314, New York, NY, USA, 2009. ACM.
- [83] Sanda Erdelez and Kevin Rioux. Sharing information encountered for others on the web. *New Rev. Inf. Behav. Res.*, 1(January):219–233, January 2000.
- [84] Russell Beale. Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing. *Int. J. Hum.-Comput. Stud.*, 65(5):421–433, May 2007.
- [85] Sudheendra Hangal, Abhinay Nagpal, and Monica Lam. Effective browsing and serendipitous discovery with an experience-infused browser. In *Proceedings of the 2012 ACM*

- International Conference on Intelligent User Interfaces*, IUI '12, pages 149–158, New York, NY, USA, 2012. ACM.
- [86] Alexander Kohn. Fortune or failure: Missed opportunities and chance discoveries in science. 1989.
- [87] Stephann Makri and Ann Blandford. Coming across information serendipitously—part 2: A classification framework. *Journal of Documentation*, 68(5):706–724, 2012.
- [88] Lucy A. Suchman. *Plans and Situated Actions: The Problem of Human-machine Communication*. Cambridge University Press, New York, NY, USA, 1987.
- [89] Bill N. Schilit, Norman I. Adams, and Roy Want. Context-aware computing applications. pages 85–90, December 1994.
- [90] David R. Morse, Stephen Armstrong, and Anind K. Dey. The what, who, where, when, why and how of context-awareness. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '00, pages 371–371, New York, NY, USA, 2000. ACM.
- [91] Albrecht Schmidt, Michael Beigl, and Hans w. Gellersen. There is more to context than location. *Computers and Graphics*, 23:893–901, 1998.
- [92] Anind K. Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5(1):4–7, January 2001.
- [93] Max Wilson and Ebook Central. *Search-User Interface Design*. Morgan Claypool Publishers, San Rafael, 2011.
- [94]
- [95] Anastasios Tombros and Mark Sanderson. Advantages of query biased summaries in information retrieval. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '98, pages 2–10, New York, NY, USA, 1998. ACM.
- [96] Offer Drori and Nir Alon. Using document classification for displaying search results lists. *Journal of Information Science*, 29(2):97–106, 2003.
- [97] Pernilla Qvarfordt, Gene Golovchinsky, Tony Dunnigan, and Elena Agapie. Looking ahead: Query preview in exploratory search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13, pages 243–252, New York, NY, USA, 2013. ACM.

- [98] Marti A. Hearst. Clustering versus faceted categories for information exploration. *Commun. ACM*, 49(4):59–61, April 2006.
- [99] M. Dork, Nathalie H. Riche, G. Ramos, and S. Dumais. Pivotpaths: Strolling through faceted information spaces. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2709–2718, 2012.
- [100] Nicholas J. Belkin, Charles L.A. Clarke, Ning Gao, Jaap Kamps, and Jussi Karlgren. Report on the sigir workshop on ”entertain me”: Supporting complex search tasks. *SIGIR Forum*, 45(2):51–59, January 2012.
- [101] Morgan Harvey, Max Wilson, and Karen Church. Workshop on searching for fun 2014. In *Proceedings of the 5th Information Interaction in Context Symposium, IliX ’14*, pages 6–6, New York, NY, USA, 2014. ACM.
- [102] Yashar Moshfeghi and Joemon M. Jose. On cognition, emotion, and interaction aspects of search tasks with different search intentions. In *Proceedings of the 22Nd International Conference on World Wide Web, WWW ’13*, pages 931–942, Republic and Canton of Geneva, Switzerland, 2013. International World Wide Web Conferences Steering Committee.
- [103] Theresa Dirndorfer Anderson. Kickstarting creativity: supporting the productive faces of uncertainty in information practice. 2010.
- [104] Jaime Teevan, Daniel Ramage, and Meredith Ringel Morris. #twittersearch: A comparison of microblog search and web search. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM ’11*, pages 35–44, New York, NY, USA, 2011. ACM.
- [105] Yuheng Hu, Shelly D Farnham, and Andrés Monroy-Hernández. Whoo. ly: Facilitating information seeking for hyperlocal communities using social media. In *Proceedings of the 2013 ACM annual conference on Human factors in computing systems*, pages 3481–3490. ACM, 2013.
- [106] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. The bohemian bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’12*, pages 1461–1470, New York, NY, USA, 2012. ACM.

- [107] Lennart Björneborn. Serendipity dimensions and users' information behaviour in the physical library interface. *Information Research*, 13(1), 2008.
- [108] Younghwa Lee and Kenneth A. Kozar. Designing usable online stores: A landscape preference perspective. *Information & Management*, 46(1):31–41, 2009.
- [109] Dong-Mo Koo and Young-Yun Choi. Knowledge search and people with high epistemic curiosity. *Computers in Human Behavior*, 26(1):12–22, 2010.
- [110] Rob Tieben, Tilde Bekker, and Ben Schouten. Curiosity and interaction: Making people curious through interactive systems. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction*, BCS-HCI '11, pages 361–370, Swinton, UK, 2011. British Computer Society.
- [111] Qiong Wu, Chunyan Miao, and Bo An. Modeling curiosity for virtual learning companions. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 1401–1402, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.
- [112] Eric van Dijk and Marcel Zeelenberg. When curiosity killed regret: Avoiding or seeking the unknown in decision-making under uncertainty. *Journal of Experimental Social Psychology*, 43(4):656 – 662, 2007.
- [113] Heather Lynn O'Brien. The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interact. Comput.*, 22(5):344–352, September 2010.
- [114] Tony Russell-Rose, Joe Lamantia, and Stephann Makri. Defining and applying a language for discovery. In Andreas Nrnberger, Sebastian Stober, Birger Larsen, and Marcin Detyniecki, editors, *Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation*, Lecture Notes in Computer Science, pages 3–28. Springer International Publishing, 2014.
- [115] Robert F. DeVellis. *Scale development: theory and applications*, volume 26. Sage, Thousand Oaks, Calif; London, 2003.
- [116] Heather L. O'Brien and Elaine G. Toms. Is there a universal instrument for measuring interactive information retrieval?: The case of the user engagement scale. In *Proceedings of the Third Symposium on Information Interaction in Context*, IiX '10, pages 335–340, New York, NY, USA, 2010. ACM.

- [117] Heather L. O'Brien and Mahria Lebow. Mixedmethods approach to measuring user experience in online news interactions. *Journal of the American Society for Information Science and Technology*, 64(8):1543–1556, 2013.
- [118] Heather L O'Brien. Exploring user engagement in online news interactions. *Proceedings of the American Society for Information Science and Technology*, 48(1):1–10, 2011.
- [119] Marc Hassenzahl. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction*, 19(4):319–349, 2004.
- [120] Martin Colbert and Angela Boodoo. Does 'letting go of the words' increase engagement: A traffic study. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 655–667, New York, NY, USA, 2011. ACM.
- [121] Lori McCay-Peet, Mounia Lalmas, and Vidhya Navalpakkam. On saliency, affect and focused attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 541–550, New York, NY, USA, 2012. ACM.
- [122] Jonathan Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4):73–93, 1992.
- [123] Yuping Liu and L. J. Shrum. What is interactivity and is it always such a good thing? implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising*, 31(4):53–64, 2002.
- [124] Sally J. McMillan and Jang-Sun Hwang. Measures of perceived interactivity: An exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. *Journal of Advertising*, 31(3):29–42, 2002.
- [125] RFC Ietf. 2616, hypertext transfer protocolhttp/1.1. *URL* <http://www.rfc.net/rfc2616.html>, 1999.
- [126] Ian Fette and Alexey Melnikov. The websocket protocol. *URL* <https://tools.ietf.org/html/rfc6455>, 2011.
- [127] Jeeyun Oh and S. S. Sundar. How does interactivity persuade? an experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *Journal of Communication*, 65(2):213–236, 2015.
- [128] Matthias J Gruber, Bernard D Gelman, and Charan Ranganath. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, 84(2):486–496, 2014.

- [129] Maria Gäde and Mark M. Hall. The role of language skills in interactive social book search sessions. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, CHIIR '16, pages 257–260, New York, NY, USA, 2016. ACM.
- [130] Juan D. Millan-Cifuentes, Ayse Göker, Hans Myrhaug, and Andrew MacFarlane. Curiosity driven search: When is relevance irrelevant? In *Proceedings of the 5th Information Interaction in Context Symposium*, IiX '14, pages 279–282, New York, NY, USA, 2014. ACM.
- [131] Earl Bailey and Diane Kelly. Developing a measure of search expertise. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, CHIIR '16, pages 237–240, New York, NY, USA, 2016. ACM.
- [132] Min Jeong Kang, Ming Hsu, Ian M Krajbich, George Loewenstein, Samuel M McClure, Joseph Tao-yi Wang, and Colin F Camerer. The wick in the candle of learning epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, 20(8):963–973, 2009.
- [133] Edith Law, Ming Yin, Joslin Goh, Kevin Chen, Michael A. Terry, and Krzysztof Z. Gajos. Curiosity killed the cat, but makes crowdwork better. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 4098–4110, New York, NY, USA, 2016. ACM.
- [134] Alan Menk and Laura Sebastiá. Predicting the human curiosity from users' profiles on facebook. In *Proceedings of the 4th Spanish Conference on Information Retrieval*, CERI '16, pages 13:1–13:8, New York, NY, USA, 2016. ACM.

Appendix A

Publications

Curiosity driven search: When is relevance irrelevant?

Juan D. Millan-Cifuentes
City University London, UK
AmbieSense Ltd
juan.millan-
cifuentes.1@city.ac.uk

Ayse Göker
Robert Gordon University
AmbieSense Ltd
Aberdeen, UK
a.s.goker@rgu.ac.uk

Hans Myrhaug
AmbieSense Ltd
Aberdeen, UK
hans@ambiesense.com

Andrew MacFarlane
City University London, UK
A.Macfarlane-1@city.ac.uk

ABSTRACT

Classical information search behaviour models based on work-task scenarios fail to explain common leisure search scenarios motivated by a hedonistic need rather than a defined information need. This paper presents work into such unstructured search driven by curiosity. In order to explore this hedonistic catalyst, a social media search application was designed in which the search experience is triggered by the user's spatio-temporal context during their exploration rather than query-response based information retrieval. We report a study with real users and a simulated casual-leisure search task where results indicated that relevance is not relevant for some searches.

Keywords

Curiosity, Casual-Leisure, Social Media, Context, Search Experience

1. INTRODUCTION

For many years digital search has been understood just as a findability problem and has been related to the search box paradigm due to its origins in library systems. Classical Information Search Behaviour (ISB) models had focused their attention on the destination (the results or the goal) of the search with metrics such as precision or recall instead of the search journey [12] (e.g engagement, flow, telepresence, etc).

Social media services have shown that ISB models based on work-task scenarios do not explain common information leisure search behaviours[5]. For example, an individual who searches their social media universe for hours after a long day at work may do so out of curiosity, to relax or for other hedonistic¹ motivation, rather than because of a clear in-

formation need. Therefore, models such as Elsweller et al's [5] have shown the importance of re-thinking information searching behaviour theory for casual-leisure search.

Query-response information retrieval (IR) systems have been used quite successfully in work-based scenarios, although many have highlighted the importance of moving beyond this paradigm when the work-task is exploratory or involves learning [17]. But if a casual search-leisure user has an undefined information need or a hedonistic motivation, why should the search experience be based on an interaction model that assumes the existence of this information need and the user's capacity to communicate it? [14, 15].

This paper explores casual-leisure behaviour in social media by eliciting curiosity driven search sessions. Instead of following the query-response paradigm for the casual leisure searcher, we propose the design of a search experience triggered by the user's context. Instead of focusing on the interaction design in the communication of the information need, the intention is to model and produce an environment around the hedonistic need or a set of "transient information needs to keep the session going" [5] by allowing the users to communicate their own context or "desired context". For example, a user searches on Twitter to find out what is happening in London when what he really wants is to relax while he is on the bus going home.

The goal of the paper is to propose a design to improve the search experience for a casual searcher. The paper contains two main novel contributions. Firstly, it describes a simulated "casual-leisure" search scenario such as "wasting or killing time" or "exploring for the experience" in the laboratory using the user's context to elicit curiosity from social media content. Secondly, it was found that for some participants relevance was not relevant (e.g. users find irrelevant information but are happy with their search experience) in an initial comparison between a user's context triggered search and a classical query-response based search experience from questionnaires and recorded search sessions.

2. RELATED WORK

Information Search Behaviour. Although search is commonly understood as just a findability problem and restricted to a query-response paradigm, research in Human Computer Interaction (HCI) and IR has highlighted the need intended to be derogatory.

¹The word hedonistic is used relating to pleasure and is not

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IIIX '14 August 26 - 29 2014, Regensburg, Germany

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

Copyright 2014 ACM 978-1-4503-2976-7/14/08 \$15.00.

<http://dx.doi.org/10.1145/2637002.2637042> ...\$15.00.

to address exploratory search work-task scenarios [17, 12].

From Taylor’s [14] “visceral needs” to “compromised need” ISB researchers have acknowledged the vital role of the user in the searching process. They understand all ISB as a consequence of an information need that demands to be satisfied. However, social media (e.g Facebook, Twitter, Instagram, etc), mobile devices and other pervasive technologies have made information accessible to people in leisure scenarios and revealed ISB motivated by hedonistic need rather than informational. Elswiler and Wilson [5, 18] have demonstrated how classical ISB focused on work-tasks scenarios fall short in explaining common casual-leisure search behaviours because they were created in library and information science. Their findings suggest that IR systems for leisure search scenarios should be designed to address hedonistic search intentions. This work builds on their model and seeks to study leisure search behaviours in action motivated by a leisure need (e.g curiosity, fun, relax, etc).

Curiosity. Human emotions and behaviour is strongly affected by curiosity. In a recent model, Kashdan et al. [9] highlight exploration and absorption (e.g losing track of time) as important factors to assess curiosity. Using the absorption concept, they highlight overlapping characteristics between curiosity and “flow theory” [3], the idea of an “optimal level of experience” in which a person is fully immersed by perceiving the correct balance between the challenge and the skills to face it.

According to Berlyne [1], curiosity is triggered by novelty, complexity, uncertainty and conflict. These interrelated variables have been used to create different kinds of stimulus (services and products) such as video games, fashion magazines, etc. Recently, interactive system designers of public spaces have developed a curiosity model around these principles to design playful interactive systems [16]. Instead of focusing on the content, they analysed how the interaction by itself could affect the user’s curiosity. Inspired by their findings, it can be deduced that curiosity could be based not just on the content (e.g topically relevant results), but also in the way the search experience is designed. This research aims to find out if a curiosity driven search experience could encourage people to explore and display casual-leisure information behaviour.

Contextual Search. Context is a powerful variable to understand and modify human behaviour. A comprehensive definition for context in computer science is provided by Dey [4] as “any information that can be used to characterise the situation of an entity”.

Moreover, this definition of context seems too abstract. Therefore, literature on context frameworks was reviewed and adopted one due to its extensiveness and logical division of the context. It consists of: Task, Spatio-temporal, Personal, Social, and Environment categories [6, 7].

In order to improve the search results, researchers in IR have used context to understand and predict search intent [6]. However, most of them have used the context concept within the IR system, but not when designing the search experience. We used the above mentioned framework to model the search experience.

Recently there has been a lot of interest in using contextual features around a social media content (e.g location and time). For example, Whoo.ly [8] is a web application that connects people with their hyper-local communities using event detection algorithms over Twitter data. Their search

experience is not driven by contextual features following a similar layout of typical curated news media aggregators.

3. METHODOLOGY

In order to study casual-leisure search behaviour, the evaluation and experiment followed a user-centred approach both in the laboratory and naturalistic scenarios [10].

User Study Setup. A simulated task scenario was presented with a loosely-defined information need in order to generate an information environment for participants and gather their judgments [2]. The participants were asked to explore “what things are happening in their city or other parts of world” while they wait for their friends in the simulated scenario. Twitter was chosen because previous researchers have highlighted microblogging as an important scenario in casual leisure search behaviour [5, 18].

The participants were given user engagement questionnaires to evaluate the search experience[11]. Regarding the interaction data, the length of the search session during the simulated task was recorded.

Some parts of the user study were conducted in a controlled usability laboratory, but most were done on-line e.g. announcements via social media and email using Survey-Gizmo². Thus allowing a more naturalistic study setting without an observer.

At the beginning of the study, the participants filled in a pre-questionnaire. Then they were invited to take part in the simulated scenario using Ambicities or Twitter. The applications were randomly assigned to each participant similar to Hu et al. [8]. During the simulated leisure search session, there was no minimum or maximum time for the task. The simulated search sessions lasted an average of 12.06 minutes. After the session, they were presented with a post-questionnaire where they evaluated the search experience.

Participants. There were 28 participants, 5 in the laboratory and 23 who joined the on-line study. The sample comprised 19 males and 9 females. Most reported daily use of social media information and familiarity with popular social networks sites. The participants answered that when they choose a leisure activity 96% use Internet and 77% follow Word of Mouth.

Designing Ambicities. In order to increase the “casual search behaviour” specially for the study, it was assumed that the search experience should be “session focused rather than result focused” [5]. It was necessary to make the search journey itself more important than the destination during the interactive user experience and elicit curiosity around the user’s context. Therefore Ambicities³ web application was built around the “transient information need” [5]: what is happening around a location according to Twitter?. The application uses Web Sockets, and Geo-location. There were two views: Map and List as shown in Figure 1.

The aim of the system was to engage people during the search session, invite them to explore for a longer period and to experience ‘flow’ [3] rather than an IR-system which goal is to retrieve topically relevant documents as quickly as possible. Previous research found a strong relationship between context and the users’ motivation in casual leisure scenarios [5]. Spatio-temporal context features like Now, Recent, Near Me, Near a particular Location (e.g clicking a location but-

²<http://www.surveygizmo.com>

³<http://www.ambicities.com/main/>

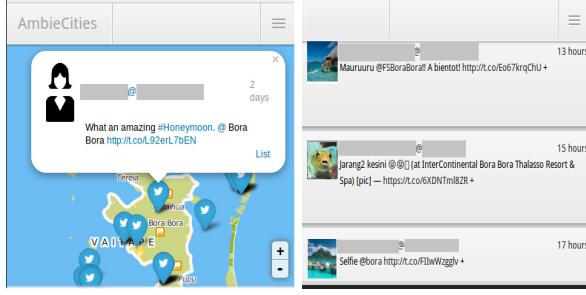


Figure 1: Ambiecities Mobile Map and List View

ton, dragging the map to a particular location or typing the name of a particular location) were used and enabled users to choose their “desired context” or “current context”. The content was not controlled because the aim was to investigate the importance of the experience and how contextual factors could be matched with the user’s hedonistic need.

4. RESULTS

Participants using Ambiecities spent more time on average exploring than those using Twitter even when they felt the content was not relevant, because they were driven by curiosity. Consider a participant Caroline⁴. She was trying to discover “fun” things using Ambiecities, but it seemed nothing was happening where she was looking. Most of the tweets were about “people just talking” she mentioned. She kept looking and searching for more tweets even in other locations for almost 12 minutes. After using Ambiecities, she said: “I felt like I was clicking on tweets because I could not help it, not because I thought I would find something interesting. I click because I was feeling nosey. After 20 clicks I realized I did not care about them but I still wanted to read them. The tweet was there; I had to look at it”.

She also said about her experience: “A good way to waste time [...]” and when discussing the time spent she answered: “I felt exploratory and just waiting for something to come up, clicking around and seeing new tweets pop up”.

This scenario illustrates the effect of eliciting curiosity using context triggered search for social media rather than query-response search. Designing the search experience with spatio-temporal context features (e.g Now, Recent, Near Me, Around a Location of Interest) encouraged hedonistic motivation instead of a pure informational by inviting people to keep exploring and waiting for more tweets whether relevant, non-relevant, interesting, etc.

In the post questionnaire, participants had similar scores to the item “The search experience was fun” both using Twitter ($\bar{x} = 3.8$, $\sigma = 1.02$) or Ambiecities ($\bar{x} = 3.4$, $\sigma = 0.92$). But when they were asked “which of the following features made the information display by the App enjoyable for you?”: 21% participants using Ambiecities selected Topical Relevance in comparison with 71% using Twitter. This interesting finding highlights how the experience (e.g. the interaction using spatio-temporal context) was more important than retrieving relevant results. Figure 2 depicts the answers to this question. In a way, the search journey

⁴All names and identifying details reported have been changed. Minor changes to the transcripts have been made for readability.

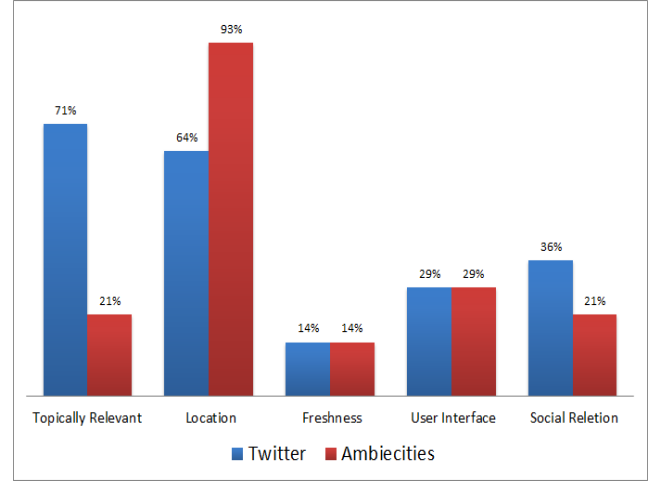


Figure 2: Which of the following features made the information displayed by the App enjoyable for you?

Table 1: Time Spent Searching in Minutes

Application	n	Σ Time	$\bar{x}$	σ
Ambiecities	14	233.46m	16.67m	13.59m
Twitter	14	110.42m	8.49m	7.36m

for some became more important than the outcome of the process in terms of relevance and usefulness (e.g finding a relevant tweet) as reported in [5].

When using Ambiecities, Caroline and other users’ behaviour were very similar to real casual search “wasting time scenarios” where behaviours like “exploring for the experience” and “needless browsing” were identified [5]. Four participants used Ambiecities in the on-line survey for more than 29 minutes whilst the longest Twitter search session was less than 17 minutes. Table 1 summarizes the time spent in the simulated scenario for all participants.

Some participants experience telepresence [13] and absorption during the simulated search scenario. In the post-questionnaire, 71% of Ambiecities ($\bar{x} = 3.8$, $\sigma = 1.2$) and 50% of Twitter ($\bar{x} = 3.2$, $\sigma = 1.3$) participants answered ‘Agree’ or ‘Strongly Agree’ to the item “The time I spent searching just slipped away”. For example after finishing the task, Peter wrote on the text area where the users should put their results (e.g places or events, relevant tweets, etc): “Absorbing and made the time go quickly with the map interface”.

Instead of submitting his results, he decided to put how the application made him feel and described his search journey. Peter later described, why he felt this way when he answered the question, “while using Ambiecities, did you enjoy your exploration?”: “Yes, it was great being able to visit different cities and regions following the twitter traffic”.

At the end of the session, after navigating around his locality, his place of birth and another city he had visited in South America, he asked the interviewer: “Do I close the app?, Everyone will know where I have been”.

Spatio-temporal triggered social media content made some people focus on “being there” or being in some other spatio-temporal context rather than “being here” [13]. So the participants who experienced telepresence did not evaluate their

interaction process in terms of finding relevant or useful tweets. Instead, they evaluated their search experience based on how much they forgot their immediate surroundings (similar to escapism [5]).

The participants' ISB using Ambiecities usually followed an initial navigation to known places around their current city. Then they navigated to places where they had some emotional relationship or interest, like "my hometown", "my home country", other cities they have visited, where friends live, or where they wish to travel.

In contrast, most participants using Twitter followed look-up search with shorter search sessions. Others just used what people or organizations in their network were talking about. For example, Karen, who followed Londonist⁵, without triggering any search looked at her Twitter timeline when performing the simulated scenario. She found a tweet from Londonist and went to the official site. Later she said: "I need to check Londonist more often, there are cool things". She described orienteering behaviour similar to Teevan et al. [15] by using her social context (e.g. friends, organizations she follow, etc.).

5. CONCLUSION AND FUTURE WORK

This paper presents initial evidence from ongoing research on how, by eliciting curiosity using spatio-temporal context in simulated social media search scenarios, participants were more likely to display "casual searching behaviours" and not typical look-up behaviour. Some participants were compelled to explore and stay longer to enjoy the experience rather than to find relevant information.

Designing the search experience encouraging curiosity and telepresence factors using the user's context could have a positive effect in other serious leisure search scenarios such as planning holidays, or choosing a new fiction book to read. Many participants in the user study mentioned that they were willing to participate in a longitudinal user study with an enhanced version of Ambiecities and that a search scenario for tourism should be targeted using both curated content and live information from social media. We also plan to report a full user engagement evaluation similar to O'Brien et al. [11] and correlate the data with the recorded interaction history.

Finally, additional dimensions of user's context (e.g. Personal, Social Context) are being considered both to create the user experience and the IR model to see if the users will spend even longer exploring out of curiosity [6]. Understanding leisure search behaviour would help to go beyond the query-response paradigm and design search experiences, where the search journey itself becomes more important than reaching the destination.

6. ACKNOWLEDGEMENT

To David Corney and David Haynes for their feedback.

7. REFERENCES

- [1] D. Berlyne. *Conflict, Arousal and Curiosity*. McGraw-Hill, 1960.
- [2] P. Borlund. The iir evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3), 2003.
- [3] M. Csikszentmihalyi. Flow and the psychology of discovery and invention. *HarperPerennial, New York*, 1997.
- [4] A. K. Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5(1):4–7, Jan. 2001.
- [5] D. Elswiler, M. L. Wilson, and B. K. Lunn. Understanding casual-leisure information behaviour. *Library and Information Science*, 1:211–241, 2011.
- [6] A. Goker, H. Myrhaug, and R. Bierig. Context and information retrieval. *Information Retrieval: Searching in the 21st Century*, pages 131–57, 2009.
- [7] A. Göker and H. I. Myrhaug. User context and personalisation. In *ECCBR Workshop on Case Based Reasoning and Personalisation*, Aberdeen, 2002.
- [8] Y. Hu, S. D. Farnham, and A. Monroy-Hernández. Whoo. ly: Facilitating information seeking for hyperlocal communities using social media. In *Proceedings of the 2013 ACM annual conference on Human factors in computing systems*, pages 3481–3490. ACM, 2013.
- [9] T. B. Kashdan, P. Rose, and F. D. Fincham. Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of personality assessment*, 82(3):291–305, 2004.
- [10] D. Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1-2):1–224, 2009.
- [11] H. L. O'Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6):938–955, Apr. 2008.
- [12] T. Russell-Rose and T. Tate. *Designing the Search Experience: The Information Architecture of Discovery*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2013.
- [13] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992.
- [14] R. S. Taylor. The process of asking questions. *American documentation*, 13(4):391–396, 1962.
- [15] J. Teevan, C. Alvarado, M. S. Ackerman, and D. R. Karger. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 415–422, New York, NY, USA, 2004. ACM.
- [16] R. Tieben, T. Bekker, and B. Schouten. Curiosity and interaction: Making people curious through interactive systems. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction*, BCS-HCI '11, pages 361–370, Swinton, UK, 2011. British Computer Society.
- [17] R. W. White and R. A. Roth. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1):1–98, 2009.
- [18] M. L. Wilson and D. Elswiler. Casual-leisure searching: the exploratory search scenarios that break our current models. *HCIR 2010*, page 28, 2010.

⁵<http://londonist.com/>

Designing autotelic searching experience for casual-leisure by using the user's context

Juan D. Millan Cifuentes
2nd year PhD
City University
Northampton Square
London EC1V 0HB, UK
juan.millan-
cifuentes.1@city.ac.uk

Ayşe Göker
Robert Gordon University
Garthdee House
Aberdeen AB10 7AQ, UK
a.s.goker@rgu.ac.uk

Andrew MacFarlane
City University
Northampton Square
London EC1V 0HB, UK
A.Macfarlane-
1@city.ac.uk

ABSTRACT

Which is more important: the journey or the destination? Classical Interactive Information Retrieval (IIR) based on work-task scenarios usually puts the emphasis on the destination of the search (the results) with metrics such as precision and recall rather than the search journey. But social media, mobile devices and other pervasive technologies have made information accessible to people in leisure scenarios and open up casual-leisure search behaviours motivated by hedonistic need such as having fun, or relaxing instead of a well-defined information need. During search sessions users might find irrelevant information but they may keep exploring because the IR system satisfies their current leisure need. This research aims to understand better casual-leisure search behaviour and design new IR systems to support autotelic search experiences.

1. MOTIVATION

“Happiness is the absence of the striving for happiness” Chuang Tzu (Taoism, 389-286 BC)

Everyday we find people, places and things that have changed our lives without looking for them [5]. Often the experience related to the finding is as important as the content itself [16]. What if in order to discover “unexpected useful” [11] information we need to stop looking and following the query-response paradigm? For long time digital IIR has been approached as a findability problem and related to a magic search box. Classical IIR models have focused their attention on the destination (the results or the goal) of the search with metrics such as precision and recall instead of the search experience itself [17] (e.g engagement, flow, telepresence, etc).

Ubiquitous social media services have shown that classical information behaviour models based on work-task scenarios fall short of common information leisure search behaviours. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IIIX '14 August 26 - 29 2014, Regensburg, Germany

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

Copyright 2014 ACM 978-1-4503-2976-7/14/08 \$15.00.

<http://dx.doi.org/10.1145/2637002.2637061> ...\$15.00.

For example, an individual might search social media for hours because he is bored rather than because he wants to fill a gap in his knowledge. Elswiler et al. [4] have highlighted the need to reconsider information behaviour theory for casual-leisure search and the design of IR-systems for leisure scenarios.

Query-response based information retrieval systems have been apparently successful in work-based scenarios, although many have highlighted the importance of moving beyond this paradigm when the work-task is exploratory or involves learning [12, 22]. Search engines are limited not just because of their ranking algorithms or technology stack, but by the interaction paradigm and the assumptions they have made about the searcher [19, 20].

This research aims to explore casual-leisure behaviour in social media by eliciting curiosity driven search sessions and how to design autotelic¹ search experiences (e.g. users find irrelevant information but are happy with their search experience, and then when they are fully absorbed they explore in depth and may discover unexpectedly useful information).

2. RESEARCH QUESTIONS

- Is it possible to design autotelic search experiences in order to study casual leisure search in action?
- Does curiosity-driven search experience elicit casual-leisure search behaviours using context through social media?
- How does the search journey become more important than getting topically relevant results? (e.g engagement, telepresence, serendipity)

3. METHODOLOGY

In order to assess casual-leisure search behaviour in action, the evaluation has followed the user-centred approach both in the laboratory and naturalistic scenarios[9].

Microblogging content from Twitter was used as the main data source and for the use case scenarios because previous casual leisure behaviour researchers have mentioned microblogging as a very important scenario of casual leisure search behaviour [4, 23].

¹Autotelic: having a purpose in and not apart from itself [13]. The word comes from Greek roots: auto (self), and telos (goal) [3].

Therefore Ambiecities² an HTML5 web application was built using Geo-location and WebSockets (push and pull interactions) to enable curiosity driven search sessions based on the user’s context and data from Twitter. An initial user study with 28 participants was conducted with using an on-line survey and a simulated search scenario.

Curiosity was chosen as the hedonistic need to generate casual-leisure search scenarios. Researchers in Psychology have associated curiosity with emotion and human behaviour. For example, Wundt [24] explained that curiosity could model a non-linear relationship between a given stimulus and the hedonic value based on the idea of an “optimal level of stimulation”. Later, Loewenstein [10] described curiosity as consequence of an unpleasant sensation called “gap in our knowledge” which is decreased by exploratory behaviour. Similar researches suggest a close relation between curiosity and other human emotions such as fear, fun, boredom, anxiety, and humor [7].

In a recent model, Kashdan [8] noticed two important factors to assess curiosity: exploration and absorption (e.g. losing track of the time). Taking the absorption concept, he explains overlapping features between curiosity and Csikszentmihalyi’s [3] ‘flow’ theory, the idea of an “optimal level of experience” in which a person is fully immersed by perceiving the correct balance between the challenge and the skills to face it.

According to Berlyne [1] curiosity is triggered by novelty, complexity, uncertainty and conflict. Tieben et al. [21] developed a curiosity model around these principles to design playful interactive systems. Instead of focusing on the content, they analysed how the interaction by itself could affect the user’s curiosity. Their findings suggest that it is possible to elicit curiosity based not just on the content (e.g. topical relevant results), but in the way the search experience is designed.

Also, context is a powerful variable to understand and modify human behaviour. However, most research has used the context concept within the IR system, but not when designing the search experience. Ambiecities was built around the idea of “transient information need”: what is happening around a location according to Twitter? and using the User Context Framework [6] to model the search experience. Previous research found a strong relationship between context and the users motivation in casual leisure scenarios [4]

The aim of the interaction design in Ambiecities was to facilitate autotelic search experiences and invite users to explore for a longer period rather than an IR-system, whose goal is to retrieve topical relevant documents as quickly as possible. As described by Csikszentmihalyi “an autotelic activity is one we do for its own sake because to experience it is the main goal [...] rather than in order to achieve some later external goal” [3]. So, instead of following typically discrete IIR approaches (i.e. submitting a query, then getting a static rank list of results, etc), the aim was to create a continuous and highly interactive search experience where a user would see markers pop up around the map they were exploring at the time. The most recent tweet is represented with a red marker, and all other other tweets are represented with a blue marker. Also to encourage the exploration of other places and the feeling of telepresence each time the application is refreshed, the user would start from a new

Table 1: Time Spent Searching

Application	n	Σ Time	$\bar{x}$	σ
Ambiecities	14	233.46m	16.67m	13.59m
Twitter	14	110.42m	8.49m	7.36m

location chosen at random from a set of pre-recommended locations (e.g. Paris, Caracas, Bora Bora, Cancun, etc). Spatio-temporal context features in the interaction design were intended to encourage users’ curiosity and telepresence (e.g. Now, Recent, Near Me, Near a particular Location). There was not any control over the content because the main goal was to see how important the experience could be to satisfy the user’s hedonistic need.

A simulated scenario was presented without an information need to generate an information environment for participants and collect their judgements [2]. Each participant had to interpret this simulated task situation: *‘You went to the best attraction in your city using public transport. Some good friends are visiting the city for their holidays and you are going to meet them at the station. You have been waiting them for almost half an hour. They send you a message saying they have got lost, but they will be coming as fast as they can. Feeling bored, you happen to see a big public touch screen with an app displaying what people are talking about. In order to change your mood, you would like to explore what is happening in your city or other parts of world while you wait for your friends’* [14].

For the qualitative measures in the user study, O’Brien et al. [15] user engagement questionnaire were used to evaluate the search experience. Regarding the interaction data, the length of the search session was recorded.

Users started the study by accepting a consent form and filling in a questionnaire. Then they were invited to take part in the simulated scenario using Ambiecities or Twitter. The applications were randomly assigned to each participant. During the simulated leisure search session, there was no minimum or maximum time for the task. The simulated search sessions lasted an average of 12 minutes. After the session, users were presented with a post-questionnaire where they evaluated the search experience [14, 15].

The initial findings of the the user study were that sometimes topical relevance was irrelevant to the simulated casual-leisure scenario while participants were using Ambiecities (e.g. users find irrelevant information but are happy with their search experience). For example, one participant mentioned that just to see the tweets pop up made her keep clicking and waiting for more results even though some of them were topically irrelevant. Search sessions using Ambiecities were longer. Table 1 summarizes the time spent in the simulated scenario for all participants.

Also some participants experience telepresence [18] and absorption during the simulated search scenario. In the post-questionnaire, 71% Ambiecities ($\bar{x}=3.8$, $\sigma=1.2$) and 50% Twitter ($\bar{x}=3.2$, $\sigma=1.3$) participants answered ‘Agree’ or ‘Strongly Agree’ to the item “The time I spent searching just slipped away”.

On the contrary, most participants using Twitter followed look-up searches with shorter search sessions and orienteering behaviour similar to that reported by Teevan et al. [20] using their microblogging social context (e.g. friends, organizations they follow, etc.)

²<http://www.ambiecities.com/main/>

The user study had limitations regarding the sample size and it was biased in terms of the recruited subjects (i.e. most of them students). Also the pre-questionnaires assessed common familiarity with social media but not personality factors which seem to be very important in order to be in ‘flow’ [3]. For some of the five laboratory subjects, the presence of an observer had a negative effect on their curiosity-driven searches and they tended to feel guilt or apologize when they were engaged.

4. FUTURE PLAN

Understanding leisure search behaviour would help to go beyond the query-response paradigm and design an autotelic search experience, when the search journey itself is more important than reaching the destination. Designing the search experience encouraging curiosity and telepresence factors using the user’s context could have a positive effect in other search scenarios such as planning holidays.

Future work includes more user studies using psychological methodologies to assess the relationship between cognitive styles, personality, curiosity, ‘flow’ factors (e.g. telepresence, absorption) and searching behaviour (e.g. dwelling time, clicks, locations visited, relevance, engagement).

Many participants in the user study mentioned that they would be willing to participate in a longitudinal user study using an enhanced version of Ambicities targeting a search scenario for tourism using both general content (e.g. Wikipedia, blogs or other tourist information) and live information from social media (e.g. photos, videos, tweets).

5. ACKNOWLEDGEMENT

Thanks to Hans Myrhaug, David Corney and David Haynes for their advice and comments. This work was partially supported by AmbieSense Ltd.

6. REFERENCES

- [1] D. Berlyne. *Conflict, Arousal and Curiosity*. McGraw-Hill, 1960.
- [2] P. Borlund. The iir evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3), 2003.
- [3] M. Csikszentmihalyi. Flow and the psychology of discovery and invention. *HarperPerennial, New York*, 1997.
- [4] D. Elsweiler, M. L. Wilson, and B. K. Lunn. Understanding casual-leisure information behaviour. *Library and Information Science*, 1:211–241, 2011.
- [5] S. Erdelez. Information encountering: a conceptual framework for accidental information discovery. In *Proceedings of an international conference on Information seeking in context*, pages 412–421. Taylor Graham Publishing, 1997.
- [6] A. Goker, H. Myrhaug, and R. Bierig. Context and information retrieval. *Information Retrieval: Searching in the 21st Century*, pages 131–57, 2009.
- [7] T. Kashdan. *Curious?: Discover the Missing Ingredient to a Fulfilling Life*. HarperCollins, 2009.
- [8] T. B. Kashdan, P. Rose, and F. D. Fincham. Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment*, 82(3):291–305, 2004.
- [9] D. Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1-2):1–224, 2009.
- [10] G. Loewenstein. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1):75, 1994.
- [11] S. Makri and A. Blandford. Coming across information serendipitously—part 2: A classification framework. *Journal of Documentation*, 68(5):706–724, 2012.
- [12] G. Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41–46, Apr. 2006.
- [13] Merriam-Webster Online. Merriam-Webster Online Dictionary, 2009.
- [14] Y. Moshfeghi and J. M. Jose. On cognition, emotion, and interaction aspects of search tasks with different search intentions. In *Proceedings of the 22Nd International Conference on World Wide Web, WWW ’13*, pages 931–942, Republic and Canton of Geneva, Switzerland, 2013. International World Wide Web Conferences Steering Committee.
- [15] H. L. O’Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6):938–955, Apr. 2008.
- [16] C. S. Ross. Finding without seeking: The information encounter in the context of reading for pleasure. *Inf. Process. Manage.*, 35(6):783–799, Nov. 1999.
- [17] T. Russell-Rose and T. Tate. *Designing the Search Experience: The Information Architecture of Discovery*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2013.
- [18] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4):73–93, 1992.
- [19] R. S. Taylor. The process of asking questions. *American documentation*, 13(4):391–396, 1962.
- [20] J. Teevan, C. Alvarado, M. S. Ackerman, and D. R. Karger. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’04*, pages 415–422, New York, NY, USA, 2004. ACM.
- [21] R. Tieben, T. Bekker, and B. Schouten. Curiosity and interaction: Making people curious through interactive systems. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction, BCS-HCI ’11*, pages 361–370, Swinton, UK, 2011. British Computer Society.
- [22] R. W. White and R. A. Roth. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1):1–98, 2009.
- [23] M. L. Wilson and D. Elsweiler. Casual-leisure searching: the exploratory search scenarios that break our current models. *HCIR 2010*, page 28, 2010.
- [24] W. Wundt. *Grundzüge der physiologischen Psychologie*. Engelmann, 1874.

Context modelling for serendipitous discoveries in exploratory mobile search over social media

Juan David Millan Cifuentes
City University London
School of Informatics
Juan.Millan-Cifuentes.1@city.ac.uk

Ubiquitous information seeking systems have changed the way urban citizens understand and explore their world. At one time our five senses were the only way we could comprehend our surroundings, but now smartphones and other devices with access to social media have become bodily extensions which allow humans to expand their social interaction and sense what is “really happening”. This paper proposes research into serendipitous discoveries and casual mobile explorations enabled by context modelling over social media. This research will use the “surrounding context” concept as lens to present information and create aggregated views of what is happening with multiple visualization techniques like real-time ambient tag cloud and heat maps. The evaluation will follow a user-centred approach both in laboratory and naturalistic scenarios focussing on leisure.

Social Media, Serendipity, Context Modelling, Exploratory Mobile Search, Casual Search

1. INTRODUCTION

Information seeking systems such as search engines often perform well when users have a clear idea of their information need where the user search activity is limited to lookup or verify (Marchionini, 2006). But search is not just a findability problem, especially when users are seeking insightful knowledge as they explore (they want something, but they might not know how or where to look for it) or they don't know what they need (they may start with little or no purpose in mind and change it with another after a successful new discovery) (Spencer, 2006). However if this is a priority concern over web search design experience for a traditional desktop user, how much more so when the primary user is mobile, strongly affected by contextual factors (such as location, time, task or social group) and the search is motivated by strong casual desire instead of a knowledge gap (Wilson et al., 2010).

Moreover, humans are social creatures even when they search for information. Their casual desires could be motivated, satisfied and understood within a social context. This may explain why so many people visit social media sites. They want to share what is happening to them whilst at the same time sense what is trending around the world or an urban location where they plan to go. For example some people when travelling rather than to follow a guide book might prefer to go where “*the locals go*”

(Balduini et al., 2012). But commercial search engines fail to respond to complex contextual queries such as; which is the most popular pub nearby after the last performance in the theatre? or what is the most popular restaurant for university students to have lunch around the university campus? According to recent studies many people now prefer to ask this type of question in their Social Network Sites (SNS) because they may receive more personalized and trustworthy answers (Morris et al, 2010).

In this regard, the proposed research explores the role of context modelling and social media to enable serendipitous discoveries and exploration for mobile users performing leisure searches. Both in academia and industry, there has been much research on the social dimension of search using geo-related content especially from Twitter and Facebook streams to satisfy the information needs of mobile users (Church et al., 2010; Balduini et al., 2012; Oussalah et al., 2013). However, they have not focused on how to model the end user's context for social seeking information beyond the physical location (Göker & Myrhaug, 2008; Emmanouilidis et al., 2013). Nor have they considered how to design for serendipity and exploration when a mobile user undertakes local searches motivated by spontaneous curiosity or casual leisure (Wilson et al., 2010). This research will propose an user-centred evaluation both in the

laboratory and naturalistic scenarios using a prototype application which will crawl real time streams of data from social media and create aggregated views of what is happening based in the user's "surrounding context" when engaging in casual mobile search.

2. RELATED WORK

The following section focus on four related areas to underline this work: serendipity, context modelling, social media and exploratory mobile search.

2.1 Serendipity

This property "*may indeed be the holy grail of the search experience*" (Russell-Rose & Tate, 2013: 77) and plenty of information seeking research has seen its value, but serendipity is a difficult property to study formally from an engineering perspective because it is hard to define and capture. For this research serendipity goes beyond the accidental discovery and explores "*what it means to have a prepared mind and an infrastructure to support discovery*" (André et al., 2009).

Common design search patterns such as "See Also" panels, recent searches and viewed items are used to support serendipitous exploration on commercial sites like Amazon and eBay. In the academic field, recent systems have been designed to promote serendipitous discoveries over web search (Erdelez & Rioux, 2000). These have used standard information retrieval techniques to model interest based on previous digital experience such as email and chat archives, they then use this information to highlight terms or pages (Beale, 2007; Hangal et al., 2012). The aid provided by these systems enhances the mind capability to recall previous experience in the moment of searching and assumes that increasing the mind's ability of recall might fill the knowledge gap to explore and discover.

In this proposal, serendipity would be assumed as not just a personal act, but rather a social process where unexpected insights from other people's behaviours may help the individual make valuable decisions they would otherwise have been unable to make. Every invention or discovery has been understood in a social context even when chance has played a vital role. The opposite is also true and is well documented, there have been cases where due to some social context, individuals have lacked the capacity to understand the value of other people or their own discoveries even when they are not "accidental" (Kohn, 1989). To exemplify the importance of social context, think of a traveller who is looking for new places of interest, and then instead of following the plans of written tourist brochures or guides, chooses to

unexpectedly change their destination based on key recommendations given by another traveller or local person (Makri & Blandford, 2012). Certainly, social media as a ubiquitous communication channel could improve this kind of interaction.

2.2 Context Modelling

The idea of context may seem obvious when we reflect on the way we understand conversations or information, but it could be quite a challenge to apply it practically. Beside context as term has been misused. It has become for information retrieval and computer science domains almost a buzz word to pique the interest of an incautious reader.

There are several definitions. For Schilit (1994) context-aware computing defines context as "where you are, who you are with, and what resources are nearby". On the other hand, Morse (2000) context is explained as "implicit situational information"; Schmidt (1999) goes further and interprets context as "interrelated conditions in which something occurs" pointing out possible relations between context features. A more comprehensive definition of context is provided by Dey (2001) when he defined it as "any information that can be used to characterise the situation of an entity". All these definitions, although different assume context is associated with the user's surrounding rather than the user model that focuses on their inner states.

However these definitions don't give practical ways in which context could be used for an information seeking application. Therefore User Context Model frameworks like Myrhaug & Goker (2002) are needed due to its extensiveness and logical division of the context universe in five categories (Task, Spatiotemporal, Personal, Social, and Environmental). This framework has been adopted for the purpose of modelling the user's context whilst performing casual mobile searches for information using social media. The main contribution regarding context modelling will be an empirical study in the relationship between context features, exploration and the user perception of an "unexpected insightful discovery".

2.3 Social Media

Media content creation, edition and distribution has been transformed by SNS and smart mobile devices like mobile phones, tablets and other ubiquitous sensors. World events are documented and captured in real-time by individuals and organizations (e.g. sport events, earthquakes, hurricanes, stock markets, etc.) (Teevan et al., 2011). Media content is generated at astonishing rates by heterogeneous multilingual social mass of people who just want to communicate their thoughts, opinions and values.

Streams of media content are extremely important as sources of information. But most of their true value is still undiscovered due to complex technical data mining challenges such as inconsistent quality (e.g. photomontage, stemming text with misspelling or social slang), lack of data (e.g. few occurrence of the same information) and the dynamic nature (e.g. newer information makes older sources irrelevant, millions of messages every minute force index scalability).

On the other hand, social media search design experience continues to ignore the user's context in the collection and delivery of the content. In other words, social media search is context free. Meanwhile, the user's perception is limited, whereas the total information is almost unending.

Paradoxically, more than ever before media content posted online has spatio-temporal and social context metadata. For example every photo and message upload in the SNS has a creation date and it is created by someone connected with a social graph of friends or followers. In addition, plenty of media content is loaded with geo-information due to inclusion of GPS and other location systems (Derczynski et al., 2013).

As previously mentioned, this research endeavours to use the "surrounding context" concept as lens to filter and present information from social media and create aggregated views of what is happening (Göker & Myrhaug, 2002).

2.4 Exploratory Mobile Search

When looking for leisure activities a mobile user's information search goals could be:

- Show me some cool local venue I don't already know near my location
- Waiting for my girlfriend at the station. Searching for the latest information.

These goals highlights that mobile search is not just a findability problem. The search of mobile users is affected by strong contextual factors and triggered by the hunger for casual information or simple curiosity rather than a fixed idea of what they want or an explicit need (Wilson et al., 2010). Therefore it would involve more than lookup search mode because sometimes users might not even know where to start the search (Marchionini, 2006).

When users lack "*knowing what they need to know*" in order to find or discover they need to build up an overall view of the information landscape without restriction and discover new patterns which will broaden their perception and help them to articulate their information needs (Spencer, 2006).

Nonetheless most research into exploratory search has been carried out for desktop, not for mobile users (Church et al., 2010) and with specialized

experts looking for new areas of research with good search skills rather than common users who may be trapped into a predetermined cognitive style (Marchionini, 2006; Kang et al., 2010). However mobile touchable devices are perfect search devices because their interfaces are not keyboard oriented, the mobile users must tap buttons, swiping or squeezing the screen with their fingers (Neumann & Schmeier, 2012). This characteristic of mobile devices may facilitate users to be engaging in learn and investigative search mode with other visualization structures. For example the "*more like this*" design pattern for tablets transforms any search task into an enjoyable browsing experience inviting the user to explore and learn (Nudelman, 2010).

This research will explore context modelling and multiple visualizations as "*infrastructure to support discovery*" when the user's search are motivated by curiosity and casual leisure desire of experiencing new things (e.g. what is "really happening" around me that I don't already know?). It also will be more concerned with increasing recall than precision using multiple techniques of visualization such as a novel contextual *ambient cloud tags* (Baldauf et al., 2009) and heat maps instead of fixed menu layouts. Although ambient cloud tags (basically tag clouds of geo-related data filter by the users location) have been used before with different SNS like Flickr and Wikipedia based on the tags and geographic information of an article or photo, none have used real-time streams from social media to build them using different contextual features proposed by this research (Joshi et al., 2010; Baldauf & Simon, 2010).

3. MOTIVATION AND RESEARCH QUESTIONS

The following scenario illustrates the goal of the research and also highlights the current problems with information services when users are in exploratory mode.

John is PhD Student at City University London. This is his third year and he is writing the results of his investigation. After he finishes writing, no one is in the office, it's already dark. He needs to relax, and have fun in a good place. He wants a beer. Before he used different applications to discover new places, but he finds most of them recommend the same places he has already been to and although the beer is good some places are not as cool as he would like even when they appear highly ranked. He wishes to broaden his knowledge of what is happening around him in order to make a good decision.

Now he opens the prototype application using his tablet. An ambient tag cloud is generated using social media information filtered by his surrounding context. The word "King" appears in big bold letters

and it is growing in size. The tag clouds dynamically change as new information arrives from social media. After observing for a few seconds, he clicks “King” and then he receives a group of context related tweets from people going to the King’s Head Theatre 5 blocks away from the university. He keeps reading the tweets, and he realizes that a lot of people seem interested. Even people he follows have commented about the performance. He also notices there are still some tickets because one of the resellers just sent a message using the social network. Instead of going to a Pub, he decides to change his destination and runs in order to arrive for the performance.

As this hypothetical scenario highlights, sometimes human beings want to discover new things and escape from the routine, often they don’t know where to start and they find themselves exploring with no specific purpose. This research will aim to answer the following research questions or topics.

Understand the role of serendipity, exploratory mobile search and user’s context model in social media. What is the role of user context modelling in the serendipitous discovery and exploration of social media information? Which context features trigger more exploration or serendipity?

Effectiveness of social media content for serendipitous discoveries and exploration using context modelling in casual mobile search. Does context relevant social media affect serendipitous discovery and exploration while mobile users undertake casual search? Do contextual based visualizations such as ambient cloud tags and heat maps enable exploration and serendipitous discoveries in a casual mobile search over social media? Which context feature and visualization produces more engagement and satisfaction in a casual mobile search over social media?

4. DESIGN ARCHITECTURE

To the best of the author’s knowledge, the only system in the academic field similar to this research has been the BOTTARI system (2012). The system merges Twitter streams and Points of Interest (POIs) to recommend restaurants and other venues based on temporally weighted opinions. However it doesn’t focus on the exploratory search behaviour and limits the user’s context model to spatial features such as location. Furthermore, it doesn’t explore the occurrence of serendipity. A recent beta commercial application called *Spindle*, which proposes a discovery social engine based on social media information from different social networks (Olanoff D., 2012).

In order to elaborate, an interesting empirical study regarding context features, exploration and serendipity, this research proposes a system to support contextual casual search over social media. The architecture proposes three main components based on current real time big data technologies (e.g. Storm¹) and HTML5. The following figure depicts the high level architecture.

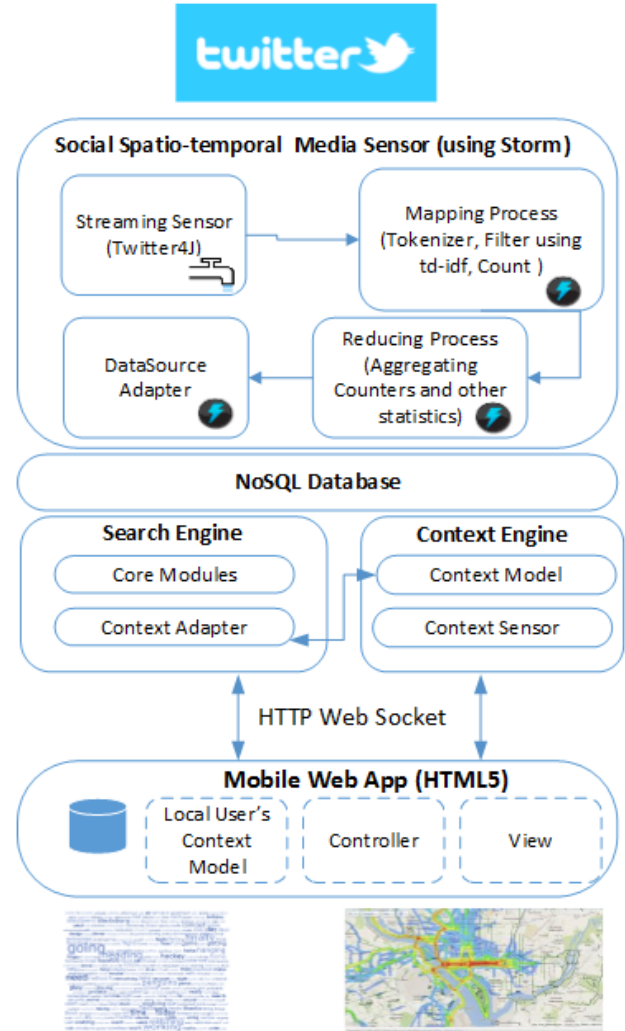


Figure 1: Prototype High-level Architecture Design to support exploratory mobile search and context modelling over spatiotemporal social media.

The first component is the Social Spatio-Temporal Sensor which will process social data (e.g. Twitter, Wikipedia, etc.), stored and indexed them based on content and spatiotemporal context. The tweets without geo-location would be clustered around the most relevant POIs existing in the database. The second component is the Context Engine based on Myrhaug & Goker (2002). The user’s context model would include for the social media domain features such as time, location, preference (*hashtags*) and social relationships from SNS. The third component represents any search engine and the adapter to

¹ Twitter storm. Available <https://github.com/nathanmarz/storm>.

the Context Engine. The fourth and final component is the mobile web application which is design around HTML5 standards (e.g. Geo-location, Websockets, etc.).

5. EVALUATION AND EXPERIMENT

In order to answer the previous research questions, the evaluation and experiment will follow an user-centred approach both in laboratory and naturalistic scenarios (Kelly, 2009).

For laboratory, *simulated work task* scenarios would be created in order to generate an informative environment for participants and help them with their judgments (Borlund, 2003). But accordingly to (Wilson et al., 2010), the simulated scenarios of casual search should also provide a “*hedonistically motivated tasks*” background. The scenarios will be created from social media streams of data (e.g. Twitter Streaming API filter by location and name of places).

For the naturalistic scenarios, a group of participants would use the application proposed in the previous section in their daily life while their interaction is logged. Also they would create a self-report on the usage of the application.

After finishing both laboratory and naturalistic “experiments”, interviews will be carried out for some of the participants.

Regarding subjective measures in the evaluation, this research will look especially into O'Brien & Tom (2008) framework to analyse the effect of context features (e.g. location, time, social relationships from the SNS), and visualization over exploration engagement.

6. CONCLUSION

This work is an initial phase of PhD research where the main goal is to understand how context features could affect the casual mobile search and influence the “*unexpected insightful*” discovery of information using real-time social media streams of data and multiple visualization techniques. This study can lead us to a better understanding of serendipity and casual mobile search from IR perspective. This research also endeavours to highlight the importance of designing the search experience using multiple user's context features.

ACKNOWLEDGEMENT

Thanks to A. MacFarlane, A. Göker, H. Myrhaug and the anonymous reviews for their helpful comments.

REFERENCES

- André, P.; m. c. schraefel; Teevan, J. & Dumais, S. T. (2009), Discovery is never by chance: designing for (un)serendipity., in Nick Bryan-Kinns; Mark D. Gross; Hilary Johnson; Jack Ox & Ron Wakkary, ed., 'Creativity & Cognition' , ACM , pp. 305-314 .
- Baldauf, M., Fröhlich, P., Reichl, P. (2009). The Ambient Tag Cloud: A New Concept for Topic-driven Mobile Urban Exploration. Proceedings of the 3rd European Conference on Ambient Intelligence, Salzburg, Austria: Springer LNCS.
- Baldauf, M. & Simon, R. (2010), Getting context on the go: mobile urban exploration with ambient tag clouds., in Ross Purves; Paul Clough & Christopher B. Jones, ed., 'GIR' , ACM.
- Balduini, M., et al. (2012), BOTTARI: An augmented reality mobile application to deliver personalized and location-based recommendations by continuous analysis of social media streams , *J. Web Sem.* 16 , 33-41 .
- Beale, R. (2007), Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing , *Int. J. Hum.-Comput. Stud.* 65 (5) , 421-433 .
- Borlund, P. (2003). The IIR Evaluation Model: A Framework for Evaluation of Interactive Information Retrieval Systems. In *Information Research*, volume 8. number 3, paper no. 152.
- Church, K, et al. (2010), SocialSearchBrowser: a novel mobile search and information discovery tool., in Charles Rich; Qiang Yang; Marc Cavazza & Michelle X. Zhou, ed., 'IUI' , ACM, pp. 101-110.
- Derczynski, L., Yang, B., & Jensen, C. S. (2013). Towards Context-Aware Search and Analysis on Social Media Data. In *Proceedings of the 16th Conference on Extending Database Technology (EDBT 2013)*.
- Dey, A. K. (2001). Understanding and Using using multiple user's context features and Context. *Personal and Ubiquitous Computing*, 5(1):4–7.
- Erdelez, S. & Rioux, K. (2000), Sharing information encountered for others on the Web, *New Review of Information Behaviour Research* 1(1), 219-233.
- Emmanouilidis, C.; Koutsiamanis, R.-A. & Tasidou, A. (2013), Mobile guides: Taxonomy of architectures, context awareness, technologies and applications, *J. Network and Computer Applications* 36 (1) , 103-125.
- Göker, A. & Myrhaug, H. I. (2002), User Context and Personalisation, in 'ECCBR Workshop on

- Case Based Reasoning and Personalisation, Aberdeen'.
- Göker, A. & Myrhaug, H. I. (2008), Evaluation of a mobile information system in context , *Inf. Process. Manage.* 44 (1) , 39-65.
- Hangal, S.; Nagpal, A. & Lam, M. S. (2012), Effective browsing and serendipitous discovery with an experience-infused browser., in Carlos Duarte; Luís Carriço; Joaquim A. Jorge; Sharon L. Oviatt & Daniel Gonçalves, ed., 'IUI' , ACM, pp. 149-158.
- Joshi, D., Luo, J., Yu, J., Lei, P., Gallagher, A (2010) Rich location-driven tag cloud suggestions based on public, community, CMM'10, pp. 21 - 26
- Kang, R.; Fu, W.-T. & Kannampallil, T. G. (2010), Exploiting knowledge-in-the-head and knowledge-in-the-social-web: effects of domain expertise on exploratory search in individual and social search environments., in Elizabeth D. Mynatt; Don Schoner; Geraldine Fitzpatrick; Scott E. Hudson; W. Keith Edwards & Tom Rodden, ed., 'CHI' , ACM, pp. 393-402 .
- Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1-2), 1-224.
- Kohn, A (1989). Fortune or failure: missed opportunities and chance discoveries. Cambridge, MA: Blackwell.
- Makri, S. & Blandford, A. (2012). Coming Across Information Serendipitously: Part 2: A Classification Framework. *Journal of Documentation*, 68(5), 706-724.
- Marchionini, G. (2006), Exploratory search: from finding to understanding, *Commun. ACM* 49 (4) , 41--46 .
- Morris M. R., et al. (2010). A Comparison of Information Seeking Using Search Engines and Social Networks. In *Proceedings of 4th International AAAI Conference on Weblogs and Social Media*, pp. 291--294.
- Morse, D. R., Armstrong, S., and Dey, A. K. (2000). The what, who, where, when, why and how of context-awareness. In *Conference on Human Factors in Computing Systems (CHI)*, pages 371--371, The Hague, The Netherlands. ACM Press.
- Neumann, G. & Schmeier, S. (2012), Exploratory Search on the Mobile Web., in Joaquim Filipe & Ana L. N. Fred, ed., 'ICAART (1)' , SciTePress, pp. 82-91 .
- Nudelman, G (2010), More Like This: A Design Pattern, [Online], Available: <http://www.uxmatt.com/mt/archives/2010/01/more-like-this-a-design-pattern.php> [12 May 2012].
- O'Brien, H.L., Toms, E.G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology* 59(6), 938-955.
- Olanoff, D. (2012). Latest Spindle For iOS Hits The Store: Discovery That's Great For Businesses And Consumers, Raises \$2.3M, [Online], Available: <http://techcrunch.com/2012/11/15/the-next-version-of-spindle-for-ios-hits-the-store-and-is-an-amazing-promo-tool-for-businesses/> [05 May 2012].
- Oussalah, M.; Bhat, F.; Challis, K. & Schnier, T. (2013), A software architecture for Twitter collection, search and geolocation services, *Knowl.-Based Syst.* 37, 105-120.
- Russell-Rose, T., Tate, T. (2013), *Designing the Search Experience*, 1st Edition, Morgan Kaufmann, Watltham.
- Schilit, B. N.; Adams, N. & Want, R. (1994), Context-Aware Computing Applications, in 'In Proceedings of the Workshop on Mobile Computing Systems and Applications' , pp. 85--90 .
- Schmidt, A., Beigl, M., and Gellersen, H.-W. (1999). There is More to Context than Location. *Computers & Graphics Journal*, 23(6):893--902
- Spencer, D. 2006. "Four Modes of Seeking Information and How to Design for Them". Boxes & Arrows: http://www.bboxesandarrows.com/view/four_modes_of_seeking_information_and_how_to_design_for_them [05 May 2012].
- Teevan, J., Ramage, D., & Morris, M. R. (2011, February). # TwitterSearch: a comparison of microblog search and web search. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM. pp. 35-44.
- Wilson, M. L. & Elsweiler, D. (2010) Casual-leisure Searching: the Exploratory Search scenarios that break our current models. In: 4th International Workshop on Human-Computer Interaction and Information Retrieval, pp. 28-31.