Pontifícia Universidade Católica
do Rio de Janeiro

**Dalai dos Santos Ribeiro**

# Exploring ontology-based information through the progressive disclosure of visual answer to related queries.

PUC-Rio - Certificação Digital Nº 1721764/CA

**Dissertação de Mestrado**

Dissertation presented to the Programa de Pós–graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática.

Advisor: Profª Simone Diniz Junqueira Barbosa

Rio de Janeiro
October 2019

Pontifícia Universidade Católica
do Rio de Janeiro

**Dalai dos Santos Ribeiro**

# Exploring ontology-based information through the progressive disclosure of visual answer to related queries.

Dissertation presented to the Programa de Pós–graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Informática. Approved by the Examination Committee.

**Profª Simone Diniz Junqueira Barbosa**
Advisor
Departamento de Informática – PUC-Rio

**Prof. Hélio Côrtes Vieira Lopes**
Departamento de Informática – PUC-Rio

**Prof. Bruno Feijó**
Departamento de Informática – PUC-Rio

Rio de Janeiro, October 10th, 2019

PUC-Rio - Certificação Digital Nº 1721764/CA

All rights reserved.

**Dalai dos Santos Ribeiro**

Bachelor in Computer Science (2017) at the Federal University of Maranhão (UFMA) with a sandwich year at Brunel University of London. Former member of PETComp, worked for Telemídia at UFMA from 2015 to 2017. Since 2017 he works at IDEIAS and DasLab at PUC-Rio.

PUC-Rio - Certificação Digital Nº 1721764/CA

Bibliographic data

Ribeiro, Dalai dos Santos

Exploring ontology-based information through the progressive disclosure of visual answer to related queries. / Dalai dos Santos Ribeiro; advisor: Simone Diniz Junqueira Barbosa. – Rio de Janeiro: PUC-Rio, Departamento de Informática, 2019.

v., 170 f: il. color. ; 30 cm

Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui bibliografia

1. Search Results Page;. 2. Progressive Disclosure;. 3. Interface Design;. 4. Exploratory Search;. 5. HCI. I. Barbosa, Simone Diniz Junqueira. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

## Acknowledgments

Taking into consideration the colossal hurdle of writing the acknowledgements, I hereby invoke the most powerful instrument known to humankind, the jewel of HCI research community, the answer to all the wickedness, the first and last resort: a **questionnaire**.

To the readers, some necessary remarks: your answers will be anonymized; there are no right answers (except there are); you may have more than one answer per question; you may have no answer per question; you may have the delusional hope that this will be the last questionnaire I'll make you answer; you'll soon discover it is not the case.

1) My parents top the acknowledgements because: a) making them proud is the reason I wake up and work every single day. b) they continuously supported (and sponsored) me during all my years as a student. c) they gave birth to me and raised me. d) they taught me to value education above greed. e) mom told I had to. f) all of the above.

2) Professor Simone DJ Barbosa: a) kept the author off the streets and fed for two years by providing him opportunities to work. b) is a brilliant advisor capable of offering excellent insight even for clumsy work, thus saving the dissertation. c) has a killer dubstep set and is available for clubs, wedding ceremonies, gender reveal parties, and bar mitzvah. d) tells more jokes than would be expected by the author of that scary contract. e) ignored her scientific instinct and, against all evidence, still believed in the author. f) did not censor this section.

3) Alan, Binho, Brandão and Julha (aka Júlia): a) are all part of a cult and became secretly married in a ritual. b) provided a roof and sofa when the author first arrived in Rio. c) emotionally supported the author numerous times during his master's. d) all cooked their speciality food to the author (except you Ruberth, I still waiting for those falafels) making him a happier (and fatter).

4) The author forced _____, _____, _____, and _____ to listen to many preliminary ideas, practice talks, terrible jokes and random rants. He also repeatedly subjected _____, _____, _____, and _____ to his research experiments questionnaires and mood swings. a) Rodrigo, Pedro, Sergio and João Victor. b) Micaele, Jack, Tamyres and Danilo. c) Daniel, Julia, Brandão and Romeu. d) Vinicius, Rômulo, Logn and Raul. e) Lauro, Luisa, Alysson and Guilherme f) Raphael, Michelangelo, Donatello and Leonardo.

PUC-Rio - Certificação Digital Nº 1721764/CA

5) Who kept the server and experiment (and thus the author) running? a) Raul's prayers b) javascript and cornstarch. c) Vinicius and Rômulo d) b and c.

6) Regarding his time at IDEIAS/DaSLab, the author is most thankful for: a) Professor Hélio Lopes's optimism and generosity towards the author. b) The playlist DASLAB - Greatest Hits available on Spotify. c) The support and knowledge of all their coworkers. d) The countless hours of work that Alysson spent to bring JARVIS to life. e) The unmatched resilience from Rodrigo, Jessica and Iago, whose fate required from them the excruciating task of working for the author.

7) About Adriana, Eriko, Julha, Marilia, Mayara, Murilo, Nando, Rafa, Susy, Victor, it is correct to say that: a) their WhatsApp messages often distracted the author of what he should be doing. b) their support is one of the greatest gifts anyone can have and what kept the author from hurting himself. c) Rafa is a bit mad she did not rank higher because did not notice the list is in alphabetical order. d) They give the best advices. e) they always know how to make the author smile after a hard day.

8) The author is also eternally grateful for: a) the members of the jury for comments and insights. b) the members of the jury for availability and comprehension with the late submission of this dissertation. c) Prof. Hélio Lopes for all support and work opportunities. d) Prof. Carlos Salles for the advice and for continuously believing in me. e) Prof. Alexandre Oliveira, for helping me in the first steps in the world of academic research. f) Prof. Nayna Patel for introducing me to HCI research, thus helping me define the career path that led me here. g) Prof Rob Macredie for the support and career advice I keep coming back every time I need to make a decision. h) The department staff (shout out to Regina!) for all the help and attention. i) Aunt Angela coffees for all the times they kept me (somehow) working. j) All the sofas, floors, and tables that I slept on at PUC-Rio. k) All the professors at PUC-Rio for the knowledge shared. l) All the people on the internet who had the same tech problems I had and came back to write a solution (especially to the anonymous user uzuMaki789, you did God's work).

9) During this journey, the author was sponsored by the following: a) BTG Pactual. b) Petrobras. c) CAPES. d) Mom.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

PUC-Rio - Certificação Digital Nº 1721764/CA

## Abstract

Ribeiro, Dalai dos Santos; Barbosa, Simone Diniz Junqueira (Advisor). **Exploring ontology-based information through the progressive disclosure of visual answer to related queries.**. Rio de Janeiro, 2019. 170p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Web search has become the predominant method for people to fulfill their information needs. Although widespread, the traditional model for search result pages is only satisfactory if the user knows quite precisely how to phrase their query to match their intended information. We propose a new model for search page results, which goes beyond providing a navigable list of visualization search results, by implicitly generating related queries to expand the search space and progressively disclosing the corresponding results.

PUC-Rio - Certificação Digital Nº 1721764/CA

## Keywords

Search Results Page; Progressive Disclosure; Interface Design; Exploratory Search; HCI.

## Resumo

Ribeiro, Dalai dos Santos; Barbosa, Simone Diniz Junqueira. **Explorando informações baseadas em ontologia através da revelação progressiva de respostas visuais para consultas relacionadas.**. Rio de Janeiro, 2019. 170p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A busca na Web se tornou o método predominante para as pessoas suprirem suas necessidades de informação. Embora seja difundido, o modelo tradicional de páginas de resultados de pesquisa só é satisfatório se o usuário souber, com bastante precisão, como elaborar sua consulta para corresponder à busca das informações desejada. Propomos um novo modelo para páginas de resultados de pesquisa, que vai além de fornecer uma lista navegável de resultados em forma de visualizações, através da geração implícita de consultas relacionadas para expandir o espaço de busca, revelando progressivamente os resultados correspondentes.

PUC-Rio - Certificação Digital Nº 1721764/CA

## Palavras-chave

Páginas de Resultados de Busca; Descoberta progressiva; Design de Interface; Busca Exploratória; IHC;

# Table of contents


1 Introduction **27**

2 Background Research **29**
2.1 Iterative Model of Information-Seeking Behavior 29
2.2 Exploratory Search 31

3 Related Work **32**
3.1 Traditional search systems 32
3.2 Flexible User Interfaces for Searching Data Visualizations 36
3.2.1 Datatone 36
3.2.2 Eviza 37
3.2.3 Articulate 38

4 Proposed Solution: Progressive Disclosure of Related Search Results **40**
4.1 Research Context 41
4.2 The Related Question Mechanism 43

5 Evaluation Study **46**
5.1 Alternative Search User Interface Models 46
5.2 Experimental Design 47

6 Results **50**
6.1 Study participants 50
6.2 NASA Task Load Index Results 51
6.2.1 Mental Demand (NASA TLX) 53
6.2.2 Physical Demand (NASA TLX) 53
6.2.3 Temporal Demand (NASA TLX) 54
6.2.4 Performance (NASA TLX) 54
6.2.5 Effort (NASA TLX) 55
6.2.6 Frustration (NASA TLX) 56
6.3 Technology Acceptance Model (TAM) Results 56
6.3.1 I find the search model X easy to use 59
6.3.2 Using the search model X gives me greater control over my search 59
6.3.3 The search model X enables me to accomplish tasks more quickly 60
6.3.4 Using the search model X enhances my effectiveness on making a search 60
6.3.5 Using the search model X makes it easier to do my job 61
6.3.6 Using the search model X makes it easier to do searches 61
6.3.7 Learning how to use the search model X system is easy for me 62
6.3.8 Learning to operate the search model X is easy for me 63
6.3.9 Interacting with the search model X is often frustrating 63
6.3.10 The search model X is rigid and inflexible to interact with 64
6.3.11 It is easy for me to remember how to perform tasks using the search model X 64
6.3.12 Interacting with the search model X requires a lot of mental effort 65


PUC-Rio - Certificação Digital Nº 1721764/CA

6.3.13 My interaction with the search model X is clear and understandable 65
6.3.14 I find it takes a lot of effort to become skillful at using the search model X 66
6.3.15 Overall, I find the search model X easy to use 66
6.3.16 It is easy to become skillful at using the search model X 67
6.3.17 I feel confident finding information in the search model X system 67
6.3.18 I have the necessary skills for using the search model X system 68
6.3.19 I have no difficulty accessing and using the search model X 68
6.4 Interviews 69
6.4.1 Model 69
6.4.2 Design 70
6.4.3 Configuration 71
6.4.4 Implementation 71
6.5 Efficiency 73
6.5.1 Was there a significant difference between the number of searches across models, considering all users? 73
6.5.1.1 Was there a significant difference between the number of searches across models, considering users with little knowledge of the models? 74
6.5.1.2 Was there a significant difference between the number of searches across models, considering users with NO knowledge of the models? 75
6.6 Users' Overall Preferences 75
6.7 Study Limitations 77
6.8 Discussion 77

7 Conclusions **79**
7.1 Contributions 79
7.2 Future work 79
7.2.1 Additional Studies 80
7.2.2 Progressive Query Expansion 80
7.2.3 User Feedback - Implicit or Explicit 80

A Experiment Tasks **85**

B Informed Consent Form **87**

C Questionnaire **88**

D Questionnaire Results **93**
D.1 Hypothesis tests - All users 93
D.1.1 TLX 93
D.1.1.1 TLX: 6 Mental Demand - How mentally demanding was the task? - All users 93
D.1.1.2 TLX: 7 Physical Demand - How physically demanding was the task? - All users 94
D.1.1.3 TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users 95
D.1.1.4 TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users 96

PUC-Rio - Certificação Digital Nº 1721764/CA

D.1.1.5 TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users 97
D.1.1.6 TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users 98
D.1.2 TAM 99
D.1.2.1 TAM: 12 I find the search model X easy to use - All users 99
D.1.2.2 TAM: 13 Using the search model X gives me greater control over my search - All users 100
D.1.2.3 TAM: 14 The search model X enables me to accomplish tasks more quickly - All users 101
D.1.2.4 TAM: 15 Using the search model X enhances my effectiveness on the job - All users 102
D.1.2.5 TAM: 16 Using the search model X makes it easier to do my job - All users 103
D.1.2.6 TAM: 17 Using the search model X makes it easier to do searches - All users 104
D.1.2.7 TAM: 18 Learning how to use the search model X system is easy for me - All users 105
D.1.2.8 TAM: 19 I find the search model X complicated to use - All users 106
D.1.2.9 TAM: 20 Learning to operate the search model X is easy for me - All users 107
D.1.2.10 TAM: 21 Interacting with the search model X is often frustrating - All users 108
D.1.2.11 TAM: 22 The search model X is rigid and inflexible to interact with - All users 109
D.1.2.12 TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users 110
D.1.2.13 TAM: 24 Interacting with the search model X requires a lot of mental effort - All users 111
D.1.2.14 TAM: 25 My interaction with the search model X is clear and understandable - All users 112
D.1.2.15 TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users 113
D.1.2.16 TAM: 27 Overall, I find the search model X easy to use - All users 114
D.1.2.17 TAM: 28 It is easy to become skillful at using the search model X - All users 115
D.1.2.18 TAM: 29 I feel confident finding information in the search model X system - All users 116
D.1.2.19 TAM: 30 I have the necessary skills for using the search model X system - All users 117
D.1.2.20 TAM: 31 I have no difficulty accessing and using the search model X - All users 118
D.2 Hypothesis tests - Users with little knowledge of model J3 119
D.2.1 TLX 119
D.2.1.1 TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models 119
D.2.1.2 TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models 120

PUC-Rio - Certificação Digital Nº 1721764/CA

D.2.1.3 TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models 121
D.2.1.4 TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models 122
D.2.1.5 TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models 123
D.2.1.6 TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models 124
D.2.2 TAM 125
D.2.2.1 TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models 125
D.2.2.2 TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models 126
D.2.2.3 TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models 127
D.2.2.4 TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models 128
D.2.2.5 TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models 129
D.2.2.6 TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models 130
D.2.2.7 TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models 131
D.2.2.8 TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models 132
D.2.2.9 TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models 133
D.2.2.10 TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models 134
D.2.2.11 TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models 135
D.2.2.12 TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models 136
D.2.2.13 TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models 137
D.2.2.14 TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models 138
D.2.2.15 TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models 139
D.2.2.16 TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models 140
D.2.2.17 TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models 141

PUC-Rio - Certificação Digital Nº 1721764/CA

D.2.2.18 TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models 142
D.2.2.19 TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models 143
D.2.2.20 TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models 144
D.3 Hypothesis tests - Users with no knowledge of model J3 145
D.3.1 TLX 145
D.3.1.1 TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models 145
D.3.1.2 TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models 146
D.3.1.3 TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models 147
D.3.1.4 TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models 148
D.3.1.5 TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models 149
D.3.1.6 TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models 150
D.3.2 TAM 151
D.3.2.1 TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models 151
D.3.2.2 TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models 152
D.3.2.3 TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models 153
D.3.2.4 TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models 154
D.3.2.5 TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models 155
D.3.2.6 TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models 156
D.3.2.7 TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models 157
D.3.2.8 TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models 158
D.3.2.9 TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models 159
D.3.2.10 TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models 160
D.3.2.11 TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models 161
D.3.2.12 TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models 162

PUC-Rio - Certificação Digital Nº 1721764/CA

D.3.2.13 TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models 163
D.3.2.14 TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models 164
D.3.2.15 TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models 165
D.3.2.16 TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models 166
D.3.2.17 TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models 167
D.3.2.18 TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models 168
D.3.2.19 TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models 169
D.3.2.20 TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models 170

PUC-Rio - Certificação Digital Nº 1721764/CA

# List of figures


Figure 2.1 Information-seeking process (Marchionini, 1997) 30

Figure 3.1 Resnick et al. (2001) main interface 33
Figure 3.2 Chen and Dumais (2000) main interface 34
Figure 3.3 Drori (2003) main interface 35
Figure 3.4 Dumais et al. (2001) main user interface 36
Figure 3.5 Datatone's main user interface 37
Figure 3.6 Eviza's main interface 38
Figure 3.7 Articulate's main user interface 39

Figure 4.1 Simplified ontology of Calvanese et al. (2017) 41
Figure 4.2 JARVIS architecture 42
Figure 4.3 Simplified ontology of Calvanese et al. (2017) with annotations 43
Figure 4.4 Relation between a question in natural language and the corresponding elements in an ontology. 44
Figure 4.5 JARVIS search user interface 45

Figure 5.1 User interface of the Traditional SUI (J1) 47
Figure 5.2 User interface of the Related-links SUI (J2) 47

Figure 6.1 NASA Task Load Index results 52
Figure 6.2 NASA Task Load Index - Mental Demand Results 53
Figure 6.3 NASA Task Load Index - Physical Demand Results 53
Figure 6.4 NASA Task Load Index - Temporal Demand Results 54
Figure 6.5 NASA Task Load Index - Performance Results 55
Figure 6.6 NASA Task Load Index - Effort Results 55
Figure 6.7 NASA Task Load Index - Frustration Results 56
Figure 6.8 TAM results (part 1) 57
Figure 6.9 TAM results (part 2) 58
Figure 6.10 The search model X is easy to use - Results 59
Figure 6.11 The search model X gives me greater control over my search - Results 59
Figure 6.12 The search model X enables me to accomplish tasks more quickly - Results 60
Figure 6.13 Using the search model X enhances enhances my effectiveness on making a search - Results 61
Figure 6.14 Technology Assessment Model - Makes it easier to do my job Results 61
Figure 6.15 Using the search model X makes it easier to do searches - Results 62
Figure 6.16 Learning how to use the search model X system is easy for me - Results 62
Figure 6.17 Learning to operate the search model X is easy for me - Results 63


PUC-Rio - Certificação Digital Nº 1721764/CA

Figure 6.18 Interacting with the search model X is often frustrating - Results 63
Figure 6.19 The search model X is rigid and inflexible to interact with - Results 64
Figure 6.20 It is easy for me to remember how to perform tasks using the search model X - Results 64
Figure 6.21 Interacting with the search model X requires a lot of mental effort - Results 65
Figure 6.22 My interaction with the search model X is clear and understandable - Results 65
Figure 6.23 I find it takes a lot of effort to become skillful at using the search model X - Results 66
Figure 6.24 Overall, I find the search model X easy to use - Results 66
Figure 6.25 It is easy to become skillful at using the search model X - Results 67
Figure 6.26 I feel confident finding information in the search model X system - Results 67
Figure 6.27 I have the necessary skills for using the search model X system - Results 68
Figure 6.28 I have no difficulty accessing and using the search model X - Results 68
Figure 6.29 Bubble plot of number of searches using each model 73

Figure D.1 Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - All users 93
Figure D.2 Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - All users 94
Figure D.3 Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users 95
Figure D.4 Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users 96
Figure D.5 Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users 97
Figure D.6 Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users 98
Figure D.7 Response to survey - TAM: 12 I find the search model X easy to use - All users 99
Figure D.8 Response to survey - TAM: 13 Using the search model X gives me greater control over my search - All users 100
Figure D.9 Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - All users 101
Figure D.10 Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - All users 102
Figure D.11 Response to survey - TAM: 16 Using the search model X makes it easier to do my job - All users 103
Figure D.12 Response to survey - TAM: 17 Using the search model X makes it easier to do searches - All users 104
Figure D.13 Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - All users 105

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure D.14 Response to survey - TAM: 19 I find the search model X complicated to use - All users 106
Figure D.15 Response to survey - TAM: 20 Learning to operate the search model X is easy for me - All users 107
Figure D.16 Response to survey - TAM: 21 Interacting with the search model X is often frustrating - All users 108
Figure D.17 Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - All users 109
Figure D.18 Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users 110
Figure D.19 Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - All users 111
Figure D.20 Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - All users 112
Figure D.21 Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users 113
Figure D.22 Response to survey - TAM: 27 Overall, I find the search model X easy to use - All users 114
Figure D.23 Response to survey - TAM: 28 It is easy to become skillful at using the search model X - All users 115
Figure D.24 Response to survey - TAM: 29 I feel confident finding information in the search model X system - All users 116
Figure D.25 Response to survey - TAM: 30 I have the necessary skills for using the search model X system - All users 117
Figure D.26 Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - All users 118
Figure D.27 Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models 119
Figure D.28 Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models 120
Figure D.29 Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models 121
Figure D.30 Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models 122
Figure D.31 Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models 123
Figure D.32 Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models 124
Figure D.33 Response to survey - TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models 125

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure D.34 Response to survey - TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models 126
Figure D.35 Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models 127
Figure D.36 Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models 128
Figure D.37 Response to survey - TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models 129
Figure D.38 Response to survey - TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models 130
Figure D.39 Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models 131
Figure D.40 Response to survey - TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models 132
Figure D.41 Response to survey - TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models 133
Figure D.42 Response to survey - TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models 134
Figure D.43 Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models 135
Figure D.44 Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models 136
Figure D.45 Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models 137
Figure D.46 Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models 138
Figure D.47 Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models 139
Figure D.48 Response to survey - TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models 140
Figure D.49 Response to survey - TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models 141

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure D.50 Response to survey - TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models 142
Figure D.51 Response to survey - TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models 143
Figure D.52 Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models 144
Figure D.53 Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models 145
Figure D.54 Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models 146
Figure D.55 Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models 147
Figure D.56 Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models 148
Figure D.57 Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models 149
Figure D.58 Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models 150
Figure D.59 Response to survey - TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models 151
Figure D.60 Response to survey - TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models 152
Figure D.61 Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models 153
Figure D.62 Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models 154
Figure D.63 Response to survey - TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models 155
Figure D.64 Response to survey - TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models 156
Figure D.65 Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models 157

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure D.66 Response to survey - TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models 158
Figure D.67 Response to survey - TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models 159
Figure D.68 Response to survey - TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models 160
Figure D.69 Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models 161
Figure D.70 Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models 162
Figure D.71 Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models 163
Figure D.72 Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models 164
Figure D.73 Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models 165
Figure D.74 Response to survey - TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models 166
Figure D.75 Response to survey - TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models 167
Figure D.76 Response to survey - TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models 168
Figure D.77 Response to survey - TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models 169
Figure D.78 Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models 170

PUC-Rio - Certificação Digital Nº 1721764/CA

# List of tables

Table 5.1 Experiment groups 48

Table 6.1 Characterization of study participants 51
Table 6.2 Compilation of Interview Results 72
Table 6.3 Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering all users? 73
Table 6.4 Conover-Iman post-test results with Bonferroni correction - Was there a significant difference between the number of searches across models, considering all users? 74
Table 6.5 Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering users with little knowledge of the models? 74
Table 6.6 Conover-Iman post-test results with Bonferroni correction - Is there a significant difference between the number of searches across models, considering users with little knowledge of the models? 75
Table 6.7 Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering users with NO knowledge of the models? 75
Table 6.8 Study participants' preferences 76

Table D.1 Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - All users 93
Table D.2 Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - All users 94
Table D.3 Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users 95
Table D.4 Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users 96
Table D.5 Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users 97
Table D.6 Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users 98
Table D.7 Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - All users 99

PUC-Rio - Certificação Digital Nº 1721764/CA

Table D.8 Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - All users 100
Table D.9 Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - All users 101
Table D.10 Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - All users 102
Table D.11 Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - All users 103
Table D.12 Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - All users 104
Table D.13 Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - All users 105
Table D.14 Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - All users 106
Table D.15 Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - All users 107
Table D.16 Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - All users 108
Table D.17 Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - All users 109
Table D.18 Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users 110
Table D.19 Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - All users 111
Table D.20 Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - All users 112
Table D.21 Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users 113
Table D.22 Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - All users 114
Table D.23 Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - All users 115

PUC-Rio - Certificação Digital Nº 1721764/CA

Table D.24 Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - All users 116
Table D.25 Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - All users 117
Table D.26 Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - All users 118
Table D.27 Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models 119
Table D.28 Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models 120
Table D.29 Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models 121
Table D.30 Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models 122
Table D.31 Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models 123
Table D.32 Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models 124
Table D.33 Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models 125
Table D.34 Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models 126
Table D.35 Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models 127
Table D.36 Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models 128
Table D.37 Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models 129
Table D.38 Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models 130

PUC-Rio - Certificação Digital Nº 1721764/CA

Table D.39 Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models 131
Table D.40 Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models 132
Table D.41 Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models 133
Table D.42 Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models 134
Table D.43 Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models 135
Table D.44 Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models 136
Table D.45 Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models 137
Table D.46 Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models 138
Table D.47 Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models 139
Table D.48 Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models 140
Table D.49 Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models 141
Table D.50 Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models 142
Table D.51 Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models 143
Table D.52 Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models 144
Table D.53 Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models 145

PUC-Rio - Certificação Digital Nº 1721764/CA

Table D.54 Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models 146
Table D.55 Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models 147
Table D.56 Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models 148
Table D.57 Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models 149
Table D.58 Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models 150
Table D.59 Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models 151
Table D.60 Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models 152
Table D.61 Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models 153
Table D.62 Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models 154
Table D.63 Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models 155
Table D.64 Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models 156
Table D.65 Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models 157
Table D.66 Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models 158
Table D.67 Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models 159
Table D.68 Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models 160

PUC-Rio - Certificação Digital Nº 1721764/CA

Table D.69 Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models 161
Table D.70 Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models 162
Table D.71 Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models 163
Table D.72 Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models 164
Table D.73 Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models 165
Table D.74 Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models 166
Table D.75 Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models 167
Table D.76 Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models 168
Table D.77 Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models 169
Table D.78 Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models 170

PUC-Rio - Certificação Digital Nº 1721764/CA

PUC-Rio - Certificação Digital Nº 1721764/CA

*"Shoot for the moon; if you miss you will die in outer space, which is cool."*

**Yekaterina Petrovna Zamolodchikova**

# 1
# Introduction

"In a little more than a decade, the Web has become the default global repository for information" (Wilson et al., 2010). Search has remarkably contributed to this result and it has become ubiquitously associated with the Web itself, to the point of becoming a default tool in any modern browser and one of the most popular activities online, already in 2008 (Fallows, 2008). In 2018, Google searches amounted to 5 billion a day.[1]

As stated by Wilson et al. (2010), "Web search, as provided by Google, Microsoft, Yahoo, etc., allows users to find the information they need via the simplest of interaction paradigms": the user types in keywords or a natural language query and obtains a related ranked result list. If the results do not fulfill the user's information needs, he/she may create a new query to obtain new results, making the information seeking process naturally iterative (Chowdhury and Chowdhury, 2002).

Joho and Jose (2006) claim that "the main purpose of search engines is to help people find information that is useful or relevant for completing a task". However, formulating a good query can be cognitively challenging for users (Belkin et al., 1982), so queries are often approximations of a user's underlying need (Thompson, 2002). The design of search user interfaces can positively contribute towards finding relevant information.

Traditionally, a user submits a search query through a search dialog box and, in response to the query, a search engine delivers one or more search result pages (SRPs) to the user. SRPs often consist of multiple pages of items that are related to the search query submitted by the user. Most of the initial search results are closely related to the query but, as the user navigates to later results, they are increasingly less related to it.

As the users may need to navigate through many SRPs (Chen and Dumais, 2000), from their point of view, the traditional model is only satisfactory if they know quite precisely how to phrase their query to match their intended search for information.

Let us consider as an example a user named Jack who wants to know the movie *genre* that generated the highest box office in 2018, but who formulates

[1]https://hostingfacts.com/internet-facts-stats/

PUC-Rio - Certificação Digital Nº 1721764/CA

the following query: "Which *movies* had the highest gross revenue in 2018?". In this case, the search results would likely contain a list of individual highest grossing movies, with links to details on each movie. Jack might then think he would need to inspect the movies one by one to try to figure out to which genre most of them belong, a very tedious task. When inspecting a movie, he may see that there is genre information associated to each movie and, realizing this is the term he should include in the query, he might reformulate the query to "Which *movie genre* had the highest gross revenue in 2018?", which then brings the intended information in the search results. This scenario has a successful ending, but in many other situations the user cannot figure out the specific query formulation needed to find the intended information.

Our work focuses on searches whose results can be represented as data visualizations. At first glance, this may seem similar to "image search" mechanisms. However, as data visualizations usually represent underlying structured data, the known relations between the data points and data sets can be used as a resource to expand the search. In our work, we assume that the data are described by an ontology (Gruber, 1993), such as those we can find in linked-open data (LOD)[2], e.g., DBPedia[3].

In this dissertation we propose a new model for search user interfaces, focusing on the search results page. Our proposal goes beyond providing a navigable list of visualization search results. It assumes an API for implicitly generating related queries to expand the search space, and progressively discloses the corresponding results. Our hypothesis is that such mechanism can improve search results, especially in situations where users cannot figure out how to formulate the precise query to yield the intended results.

This document is organized as follows. Section 2 reports the background research of our work. Section 3 presents the related work on search user interfaces, focusing on search result pages when searching for visualizations. Section 4 describes JARVIS, our proposed solution. Section 5 describes how we planned to evaluate our proposal with users. Section 6 describes and discusses the results of the study we conducted to evaluate our proposal. Finally, Section 7 concludes the dissertation, describing its contributions and pointing to future work.

[2]https://lod-cloud.net/
[3]https://wiki.dbpedia.org/

PUC-Rio - Certificação Digital Nº 1721764/CA

# 2
# Background Research

This section introduces the research foundations upon which this work is based, namely information-seeking behavior. We describe two information-seeking models: Iterative (Section 2.1) and Exploratory (Section 2.2).

Wilson (1999) defines information-seeking behavior as a set of activities that people engage in when identifying their information need, searching for it through an information resource, and using the results to satisfy that need.

Although terms such as information seeking, information retrieval, and information search behavior are often used interchangeably, in our work we will use the complementary definitions of information retrieval and information seeking by Wilson et al. (2010) and Marchionini (1989). For Wilson et al. (2010), information retrieval is a "paradigm where users enter a keyword into a system, which responds by returning the results that are most relevant to the keywords used", whereas information seeking is a broader and more complex term, encompassing "activities such as searching, browsing, and navigation". Marchionini (1989) adds that information seeking is a special case of problem-solving, which involves recognizing and interpreting the information problem to be solved and the associated planned search, and is influenced by experience, knowledge, and the information need.

Understanding human information-seeking processes is the foundation for the design of effective and usable search systems (Wilson, 1999). In the next sections, we describe existing models of information-seeking behavior.

PUC-Rio - Certificação Digital Nº 1721764/CA

## 2.1
## Iterative Model of Information-Seeking Behavior

Marchionini (1997) laid the foundation for the traditional information-seeking process and defines it as a set of sub-processes that are "both systematic and opportunistic". A detailed view of his vision for information seeking process is represented in Figure 2.1.

In this work we will use a simplified version of this process defined by Hearst (1999). This model, such as the one defined by Marchionini (1997), assumes the process is iterative and that the user information need does not change. The model comprises the following sequence of steps (Hearst, 1999):

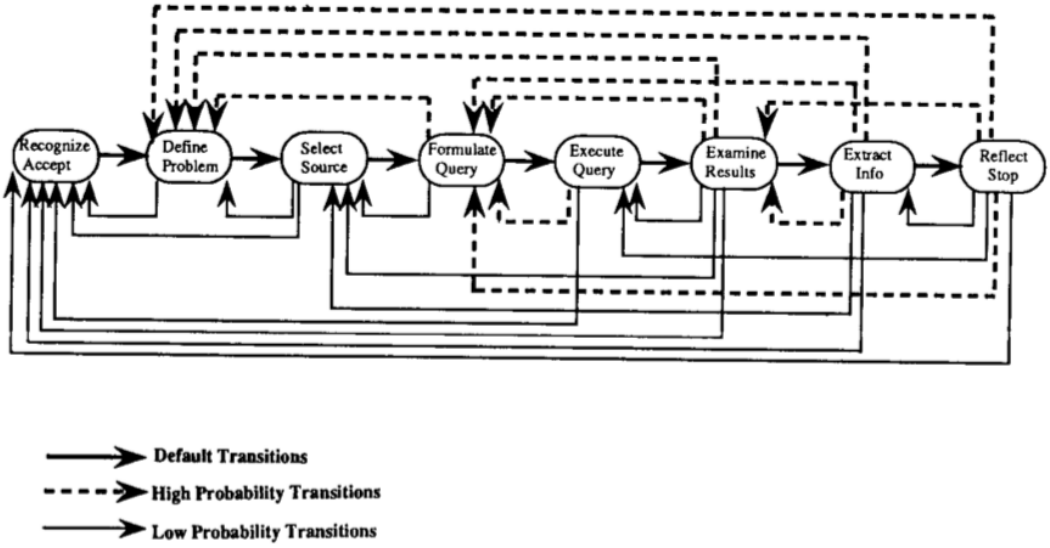



Figure 2.1: Information-seeking process (Marchionini, 1997)

1. "Recognize the information need.
2. Select the information repository to search.
3. Form a search query.
4. Send the query to the system.
5. Receive the results.
6. Evaluate and interpret the results.
7. Stop, if the information need is fulfilled, or
8. Reformulate the query and return to Step 4."

Although widespread, the iterative model of information seeking does not capture the richness of genuine information-seeking processes (Hearst, 2009), especially because users' information demands may change during the search process as a result of their interaction with the search system. The user can at the same time present a behavior that is both systematic and unsystematic, starting his/her search processes following the hierarchical approach presented by Hearst (2009), and then switching to a more dynamic behaviour that uses the initial result set as a starting point that informs further queries, as pointed out by Marchionini (1997). Marchionini (1997) also advocates that, because individual factors affect information-seeking interaction, there is a need for new models that better account for the dynamic nature of information seeking, *i.e.*,

PUC-Rio - Certificação Digital Nº 1721764/CA

models that can address the challenges of describing how users employ different search tactics and how they can make sense of the results.

## 2.2 Exploratory Search

The traditional method of information seeking is well supported by search engines, especially when the user has well-defined information needs. However, when the user lacks the knowledge or contextual awareness to formulate queries or navigate complex information spaces or the information, the search system should provide more support for a complex information seeking process, where the user is able to browse and explore the results in order to fulfill his/her needs (Wilson et al., 2010).

Exploratory Search research tackles this issue by studying information-seeking models that blend querying and browsing with a focus on learning and investigating, instead of information lookup (Marchionini, 2006). White et al. (2005) distinguish three typical situations in which exploratory search happens:

- The user has partial or no knowledge of the search target
- The search moves from certainty to uncertainty as the user is exposed to new information
- The user is actively seeking useful information and determining its structure

Heimonen (2012) argue that the exploratory search influences the development of search interfaces, and that using information about the search activity and the target documents can help reduce uncertainty. He also calls for supporting a variety of search strategies.

O'Day and Jeffries (1993) describe this incremental search behavior as a process of exploration through a series of related but distinct searches on a specific topic. They identify three distinct search modes:

- "Monitoring a well-known topic over time;
- Following a plan of information gathering;
- Exploring a topic in an undirected fashion."

This shows that even exploratory information seeking has structure and continuity, which could be supported by the search system.

PUC-Rio - Certificação Digital Nº 1721764/CA

# 3
# Related Work

Much work has been done on search systems. This chapter introduces the related work investigated in our research. Section 4.1 presents the research context where this work is inserted. Section 3.1 describes traditional search systems, and Section 3.2 describes flexible user interfaces for search for data visualizations.

## 3.1
## Traditional search systems

Much research on the design of search results pages considers traditional search systems that are focused on documents and webpages, such as commercial solutions like Google, Microsoft, Yahoo, etc. They are often presented as a list of results related to the query and keywords given by the user.

Resnick et al. (2001) designed a tabular interface to support quicker scanning of results in comparison with the list interface (Figure 3.1). The interface mapped the columns of the table onto the different elements of the abstract exhibited in a traditional list interface such as Google Search. The authors claim that their interface allowed a faster scanning of results in comparison with the list interface.

PUC-Rio - Certificação Digital Nº 1721764/CA

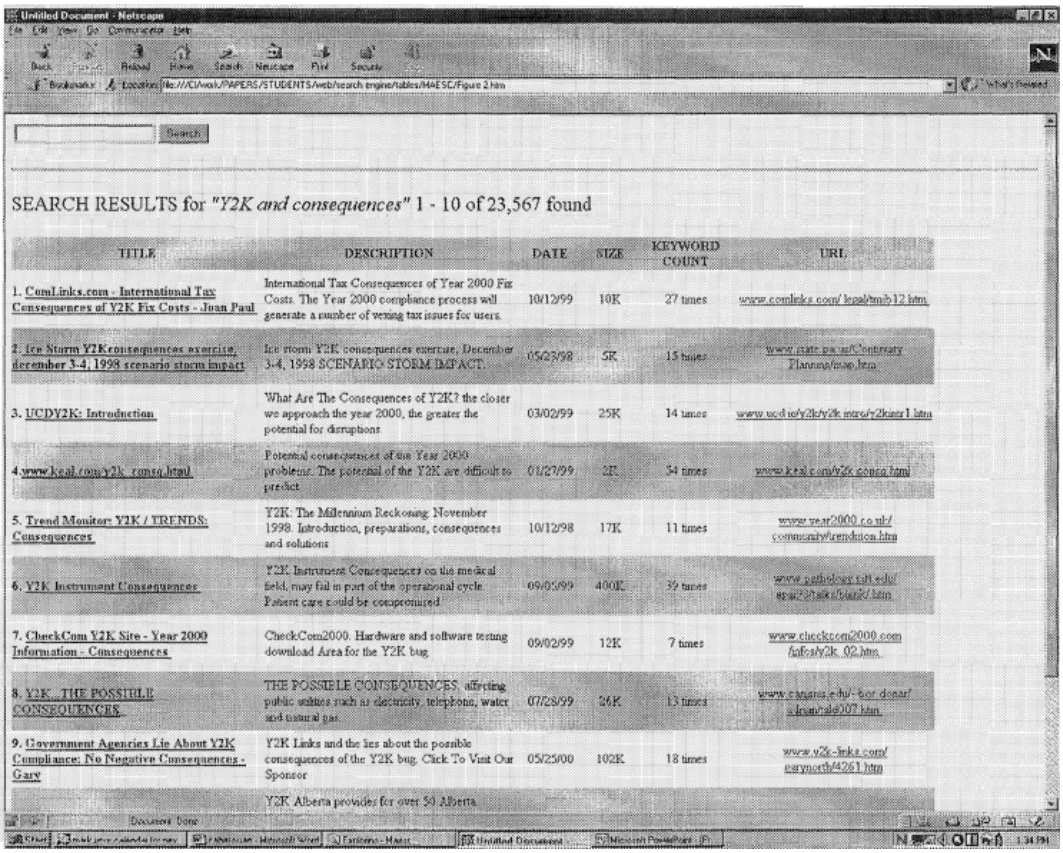



Figure 3.1: Resnick et al. (2001) main interface

Chen and Dumais (2000) designed a user interface where the results were organized into hierarchical categories. As seen in Figure 3.2, under each category, web pages belonging to that category were listed. The category could be expanded or collapsed on demand by the user. To save screen space, only the title of each page was shown and the summary could be viewed by hovering the mouse cursor over it.

PUC-Rio - Certificação Digital Nº 1721764/CA

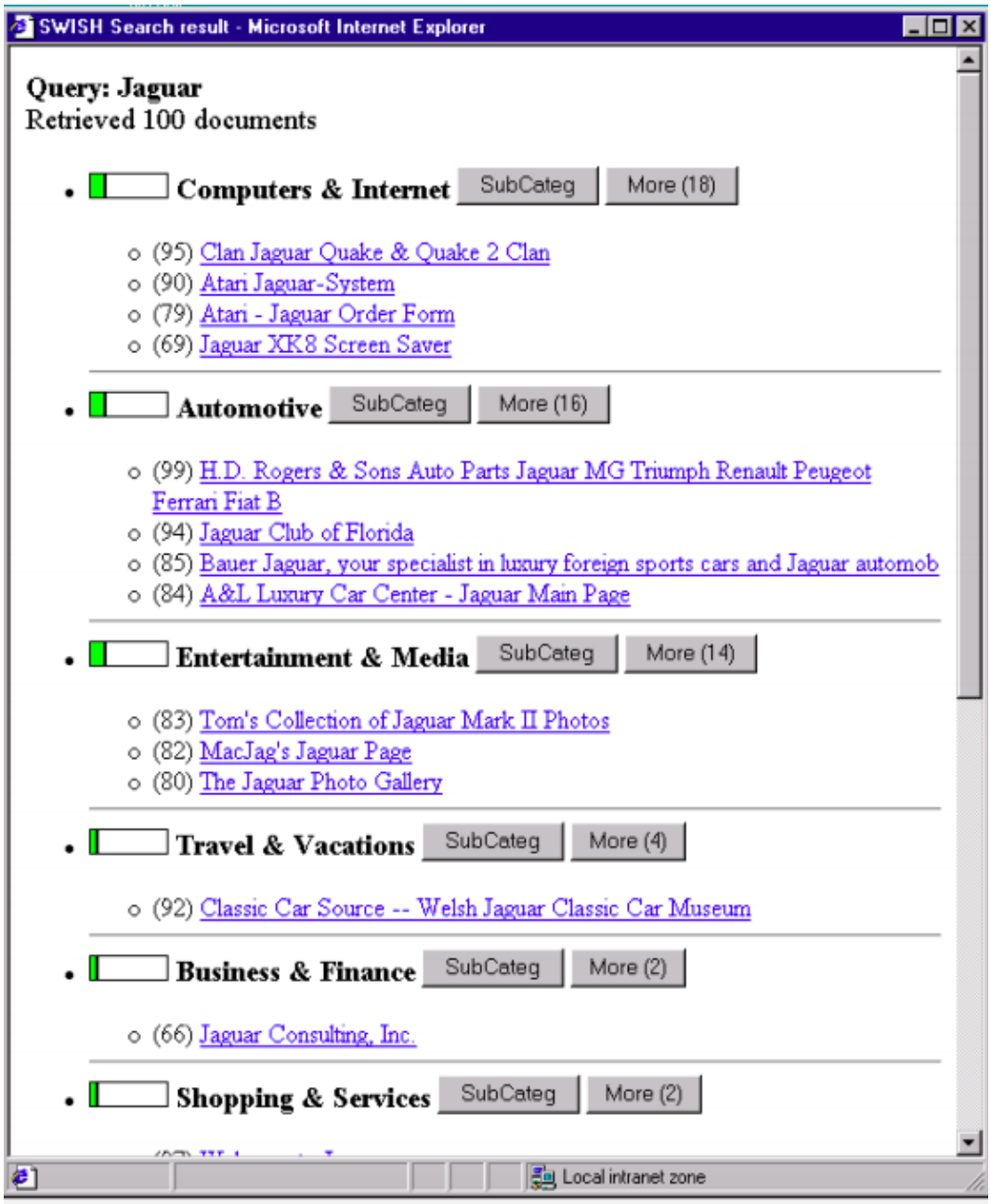



Figure 3.2: Chen and Dumais (2000) main interface

Drori (2003) defined a hierarchical structure containing three levels for displaying search results focused on the use of textual components, which consist of two categories: internal document information, such as significant sentences and words, information from HTML tags and keywords; and external document information, which is "based on the document's subject field and not contained within the actual document", such as the document address, common words, and the publisher. Despite its low resolution, we can see in Figure 3.3 the SRP for the question *"What was the date of birth of the English physicist Newton?"*, where the *titles* are shown in blue, the *keywords* in green, the *lines in context* in black, and *search terms* in red.

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure 3.3: Drori (2003) main interface

Dumais et al. (2001) developed and evaluated seven different interfaces for structuring search results using category information and list. The interfaces presented different ways to display the results on categories or lists of pages, for example, with summaries inline or by hover (Figure 3.4). The combinations are: (a) a list of items (page titles and links) that reveal each page summary on hover; (b) a list of items with inline summaries; (c) a list with category names; (d) a list with category names and items with hover summary; (e) a list with category names and items with summary inline; (f) a list of categories with no page titles.

PUC-Rio - Certificação Digital Nº 1721764/CA

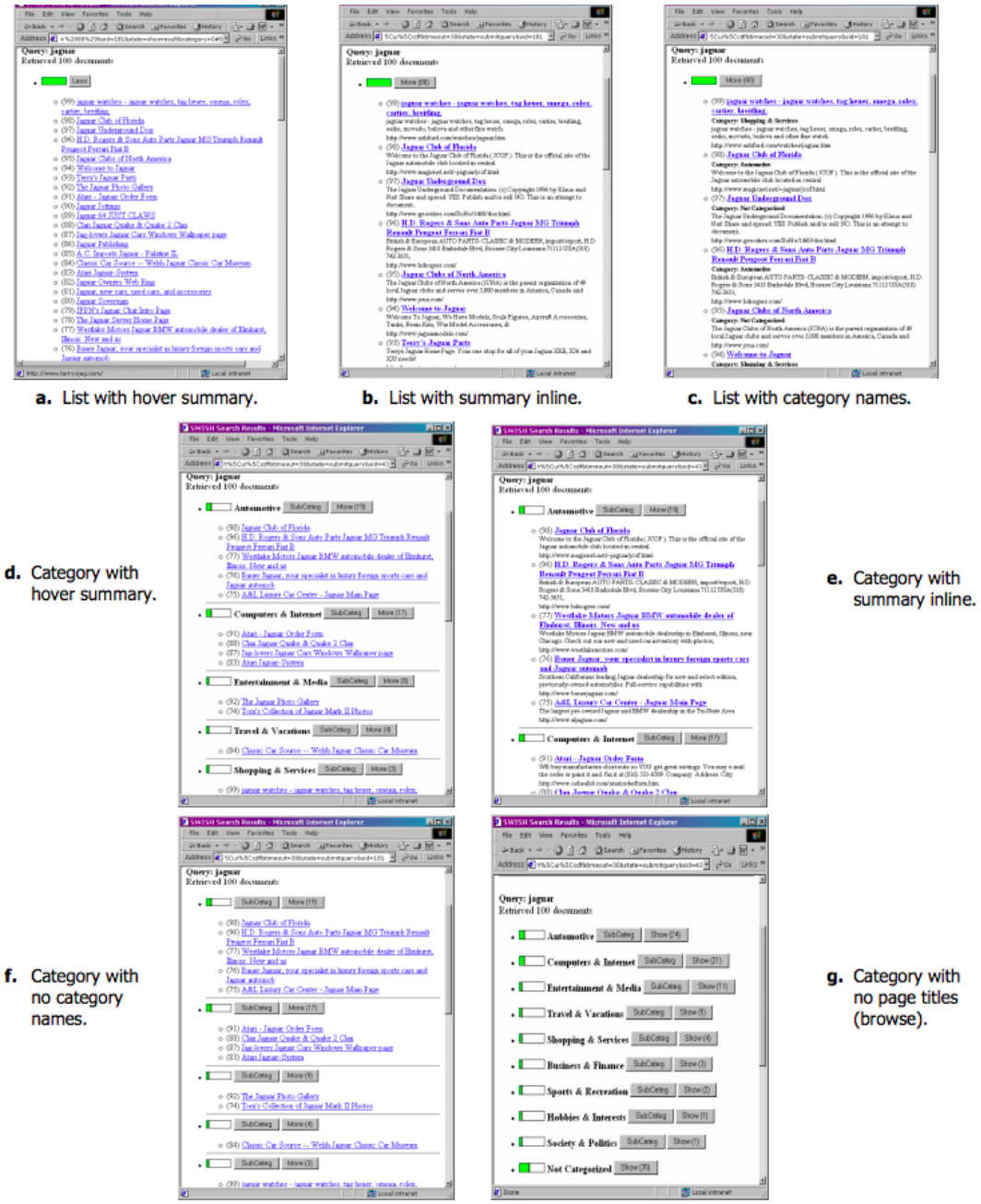



Figure 3.4: Dumais et al. (2001) main user interface

## 3.2 Flexible User Interfaces for Searching Data Visualizations

In this section, we describe some user interfaces that are more closely related to our research.

### 3.2.1 Datatone

Datatone (Gao et al., 2015) is a Web tool to address ambiguity in natural language interfaces for data visualization (Figure 3.5). In Datatone, users type

PUC-Rio - Certificação Digital Nº 1721764/CA

a keyword-based query and the system generates a set of visual representations related to it. The system exhibits only highest ranked visualization and stores a model of the ambiguity. This model is the foundation of the interface widgets, through which the user can, through direct manipulation, or through a combination of user-system interactions, manage ambiguity.

As shown in Figure 3.5, after a user has searched for "hockey medals after 2004", the system shows as the result the sum of silver medals of Ice Hockey after 2004, and allows the user to adjust the search by selecting another option in the "ambiguity widgets". Through these widgets, the user can explore the questions he/she may not have considered before, such as the "What are the silver medals of hockey and skating by country?".

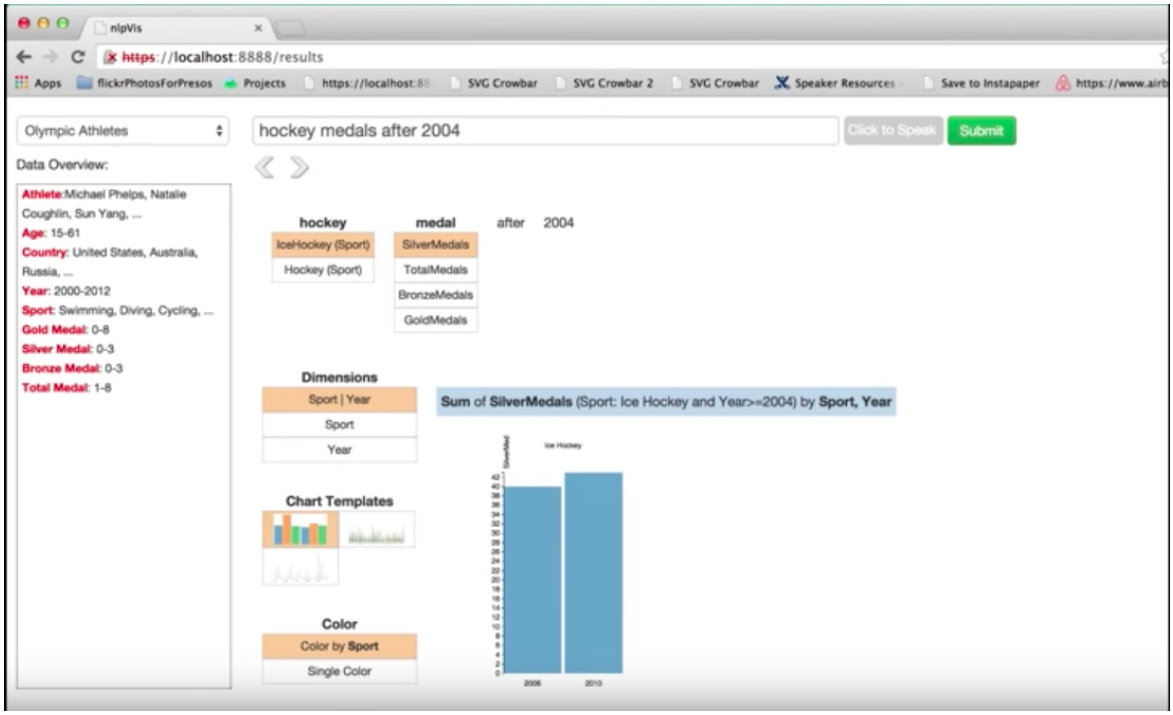



Figure 3.5: Datatone's main user interface

### 3.2.2 Eviza

Eviza (Setlur et al., 2016) is a natural language interface for visualizations that has "rich semantics and expressibility and can support the analytical flow of exploring an *existing* information visualization". Eviza has ambiguity widgets inpired in Datatone (Gao et al., 2015). Figure 3.6 shows the SRP for the question "where are the large earthquakes" The system associates *large* with the attribute *magnitude* and exhibits the ambiguity widget to allow the user to modify this setting, for instance, narrowing the search to the state of California.

PUC-Rio - Certificação Digital Nº 1721764/CA

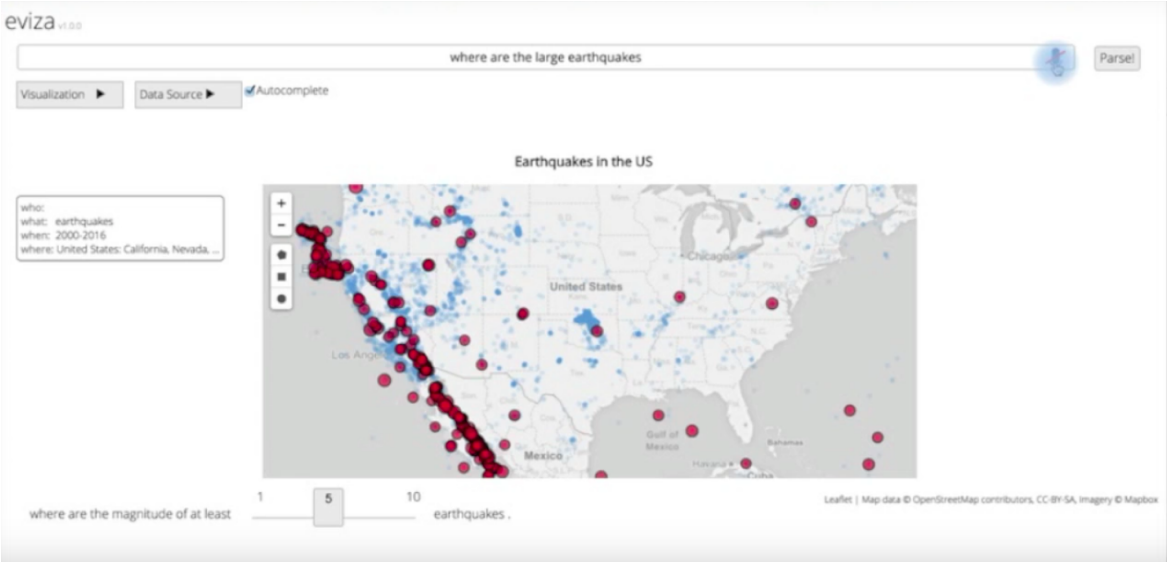



Figure 3.6: Eviza's main interface

### 3.2.3 Articulate

Articulate (Sun et al., 2010) is a "semi-automated visual analytic model that is guided by a conversational user interface". Figure 3.7 shows its main interface. The bottom of the interface is where the user's queries are displayed. In the figure, despite its low resolution, we can see the user query was "compare the gross of apple and bananas". The system shows a set of visualizations that aim to answer the user query, but to explore the data it is necessary to state a new query.

PUC-Rio - Certificação Digital Nº 1721764/CA

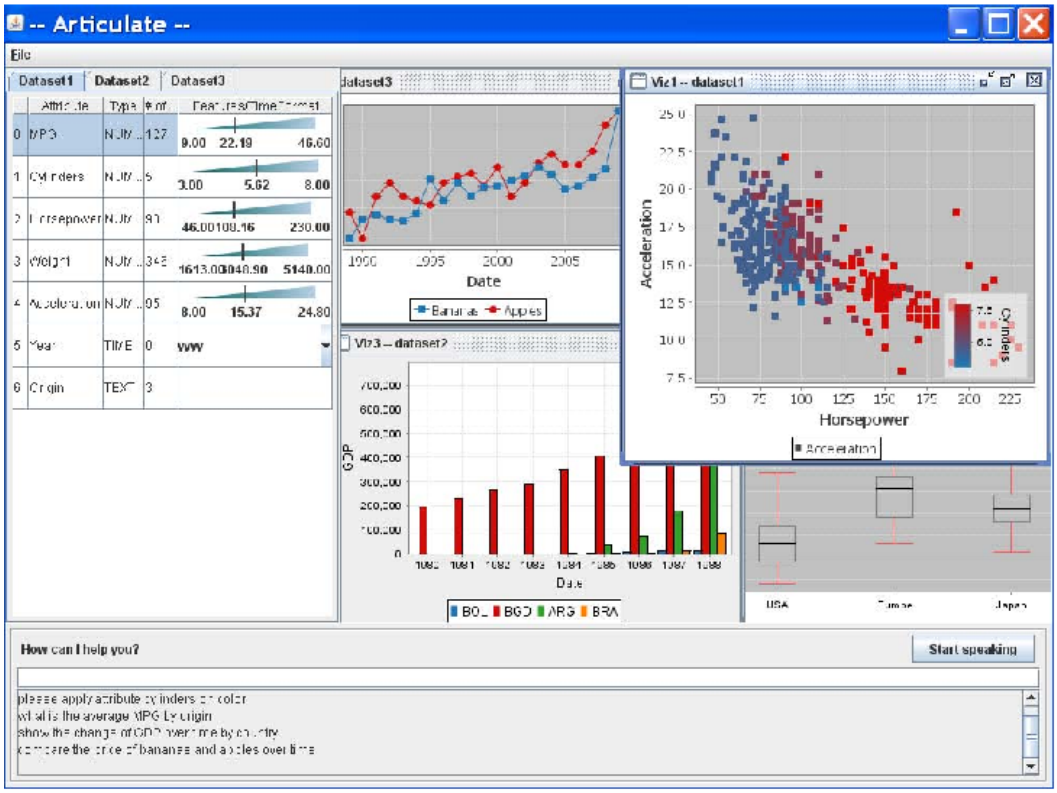



Figure 3.7: Articulate's main user interface

PUC-Rio - Certificação Digital Nº 1721764/CA

# 4
# Proposed Solution: Progressive Disclosure of Related Search Results

As described in the previous chapter, many systems allow the user to navigate through search results by refining the search query. Although effective when the user has a clear vision of their interests, those interfaces may not be very suitable when the user is performing an exploratory search or cannot properly formulate their information need.

As we have seen, Datatone (Gao et al., 2015) allows the user to navigate through related questions by direct manipulation of the query or through manual interactions with its ambiguity widgets. In other words, Datatone requires users to take action in order to obtain related search results.

We hypothesize that, instead of requiring users to manually adjust the queries to amplify their search results, user interfaces for searching data visualizations may continuously offer answers to related queries based on navigation through an underlying ontology. Our proposed search user interface, named *JARVIS- Journey towards Augmenting the Results of Visualization Search* , is based on the progressive disclosure model used by Google Images,[1] where the interface continuously appends search results to the search results page. Rather than requiring users to refine their queries, JARVIS automatically amplifies the set of results with answers to related queries.

To support our hypothesis, our group developed a system to explore the Movies and TV Series domain, using the IMDB Movie Ontology developed by Calvanese et al. (2017), an ontology to describe the movie domain semantically. Their ontology uses the International Movie Database ID[2]) data as its data source. IMDb contains comprehensive movie information such as title, genre, director and actor, *e.g.*, "Taxi Driver", " Crime", "Martin Scorsese" and "Robert De Niro". Figure 4.1 shows a simplified version of the ontology

[1]https://images.google.com/?gws_rd=ssl
[2]https://www.imdb.com/

PUC-Rio - Certificação Digital Nº 1721764/CA

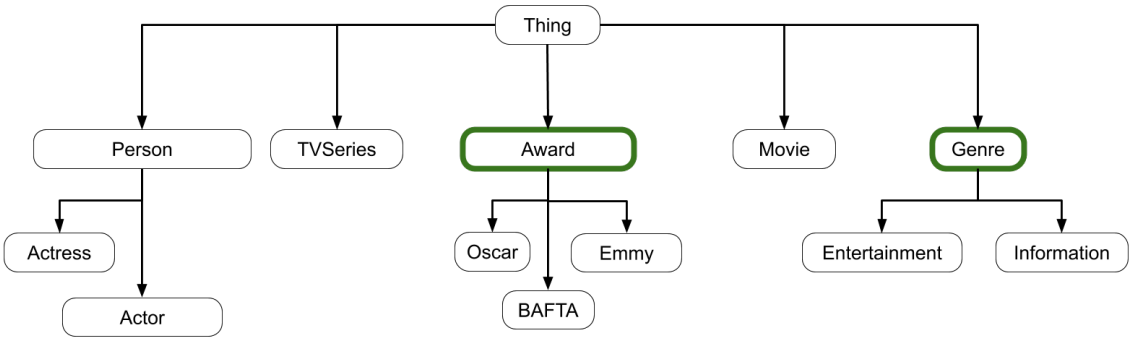



Figure 4.1: Simplified ontology of Calvanese et al. (2017)

## 4.1 Research Context

This section aims at clarifying the larger context within which this work is inserted. Figure 4.2 presents the architecture with the complementary parts of this research. The topmost part represents the interface of the system, where the user can input their questions and visualize the answers.

Our work focuses on understanding how to better support the user by designing a result page that is both effective and efficient. More specifically, we propose and evaluate a progressive disclosure mechanism for related questions. For that to happen, many other parts need to be in place. Figure 4.2 shows the architecture of JARVIS˙The red boxes were the parts of this research and were developed by the authors of this thesis; the orange box is the Related Answer Generator, which was developed in conjunction with other members of our research group but whose scope lies outside of the research presented in this thesis; the grey boxes are the contributions there were fully developed by other members of our research team. There is an essential piece of research that it is concerned about translating the natural language question to a query in RDF, in our case, a SPAQRL query. Take the question "Which actresses won the most Golden Globes last year?" as an illustration. The interpreter, fed with a movie ontology, identifies the known entities, such as *actresses (Actress → is → Person)*, *Golden Globes (Award → has_name → Golden Globe Awards)*, relationships, in this case, *won (Person → awarded → Award)* and the temporal attributes like *last year*. With this information, the system transforms the question into an RDF query that accesses the domain database and gathers the answers. This work is essential for a system like the one we proposed, but this research is not being developed by the author of this dissertation; rather, we use the result of the research of a member of our research

PUC-Rio - Certificação Digital Nº 1721764/CA

group.

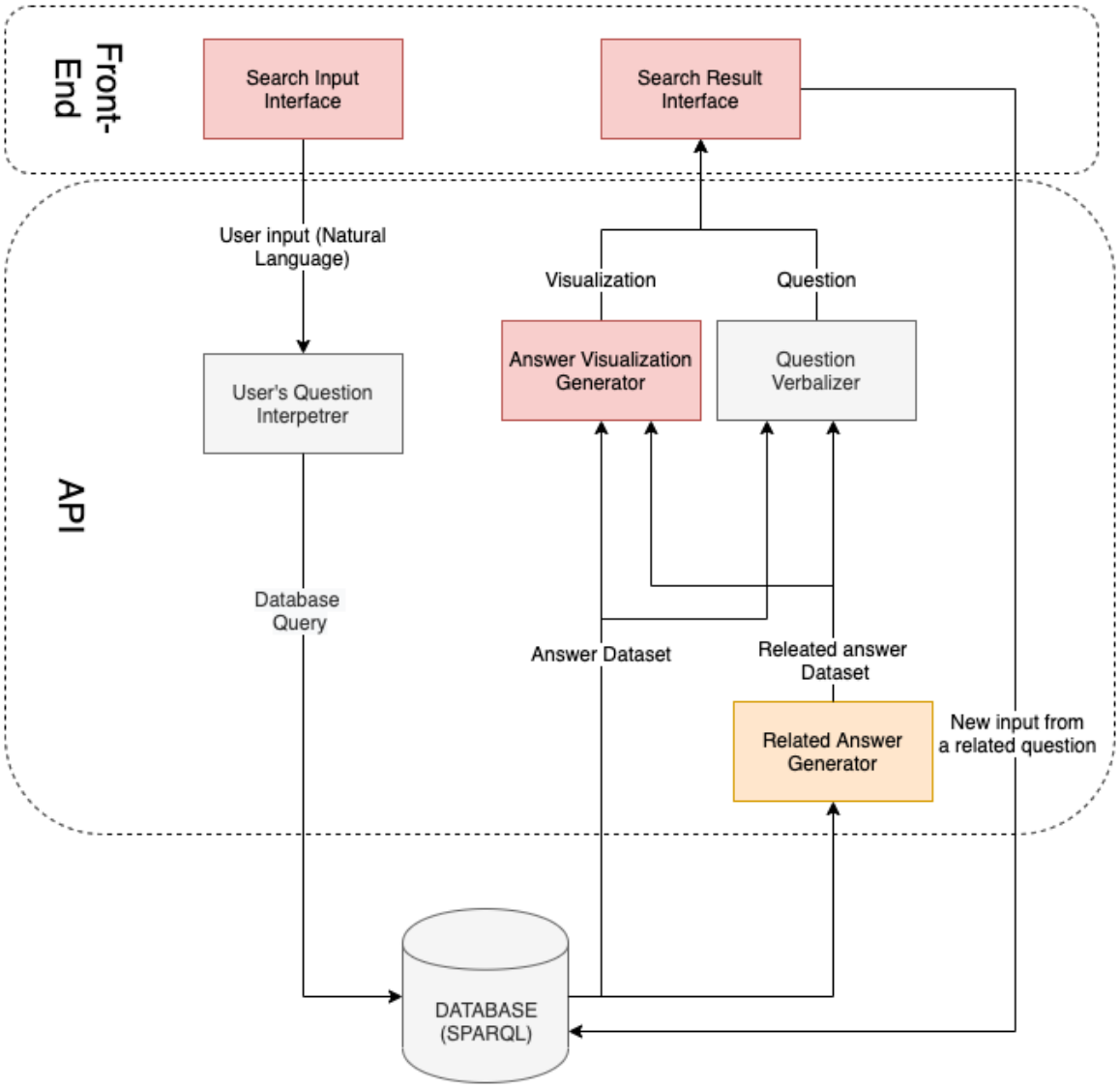



Figure 4.2: JARVIS architecture

Also crucial is the generation of the relevant related questions that will be represented as data visualization to the user. Figure 4.4 shows an example of how the system generates its related questions. In the question "What were the 5 highest rated movies from Viola Davis this decade?", after identifying the known entities from the ontology in the question, it looks for the structure of its relationships. In this case: actress → is_actress_in → movie → has_rating → imdb_rating.

In order to improve the effectiveness of the system, a domain expert is able to enrich the ontology with relationships that he/she finds interesting to the users of the search engine. Figure 4.3 shows the simplified ontology of Calvanese et al. (2017) with annotations. This process is especially important when dealing with large and complex ontology. In our case, the ontology used is

PUC-Rio - Certificação Digital Nº 1721764/CA

fairly small and simple. This aspect enables us to have good results even when we have skipped this stage when building our system. We applied the same methodology for a different domain for a large company of Oil  Gas in Brazil. The ontology, in this case, was significantly larger and presented relationships that were more complex. Because we did not have access to domain expert in this case, the related question could present such variation that would make no sense to the final user of the search system. In a case like this, developers of such system should prioritize direct relationships in the ontology to build the related question mechanism.

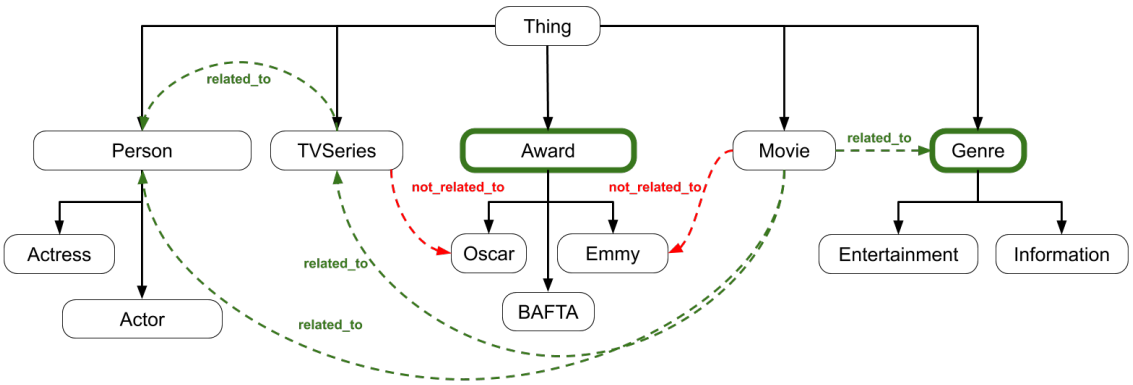



Figure 4.3: Simplified ontology of Calvanese et al. (2017) with annotations

With that structured defined, it starts to scan for traversals it can make in the ontology and which are relevant generate a related question. Those traversals can occur in various ways, for example from an entity to its parent (movie → production) or, like in Figure 4.4, to a sibling entity, that is, an entity that belongs to the same structure to the one identified in the user query. In this case, the system can change Movies to TV Series and select "What were the 5 highest rated *TV series* from Viola Davis this decade?" as the new query.

## 4.2 The Related Question Mechanism

In order to enhance the answers generated by the interpreter and to reduce the user's cognitive effort to formulate other related questions which may interest them, our group developed a mechanism that recommends answers to questions related to the main question which the user searched. This mechanism applies operations to the ontology, taking into consideration the entities that were detected in the initial question (see Figure 4.4).

PUC-Rio - Certificação Digital Nº 1721764/CA

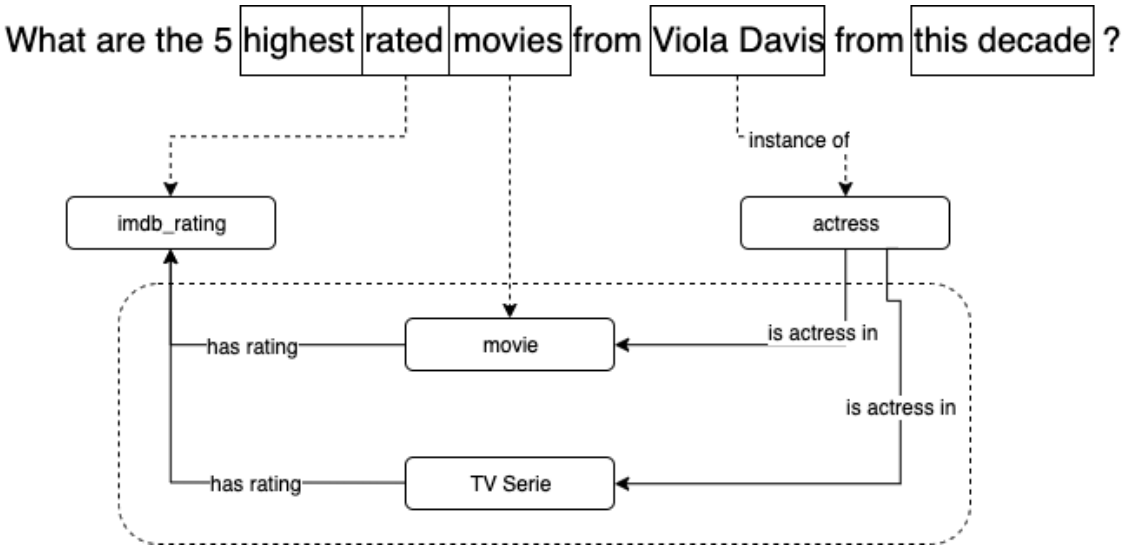



Figure 4.4: Relation between a question in natural language and the corresponding elements in an ontology.

Figure 4.2 depicts the JARVIS architecture. Let us take the example described in Chapter 1: the information needed by the user is the movie *genre* that generated the highest box office in 2018, but when formulating their query they typed: "Which *movies* had the highest gross revenue in 2018?". JARVIS sends, through an API, the natural language query written by the user. The API looks for the literal answer or answers to the question and ranks the results. It then exhibits the $n$ highest ranked direct, literal results for the query on the topmost area of the interface, in a slightly shaded area (Figure 4.5). Below that area, it progressively displays results from related questions, which are gradually received from the API. Those results are the outcomes of a search mechanism that, given a domain ontology (*e.g.*, related to the IMDB), navigates through the ontology looking for useful relationships between the elements presented in the search query to expand the given question into related ones.

JARVIS may offer, for example, results for questions such as "Which *studios* had the highest gross revenue in 2018?" (through a movie–produced by–studio relationship), "Which *movies* had the highest gross revenue in 2018 per country?" (through a movie–produced in–country relationship), and "Which *movie genre* had the highest gross revenue in 2018?" (through a movie–classified as–genre relationship). These related questions may offer the information needed by the user, as well as different perspectives on the data related to the query, without any manual interaction by the user.

In this work, we focus on the delivery mechanism for the results and on how the users interact with it. The challenges of translating a natural language question to a database query, and of navigating in a ontology to find the useful relationships for related questions are relevant research topics, which

PUC-Rio - Certificação Digital Nº 1721764/CA

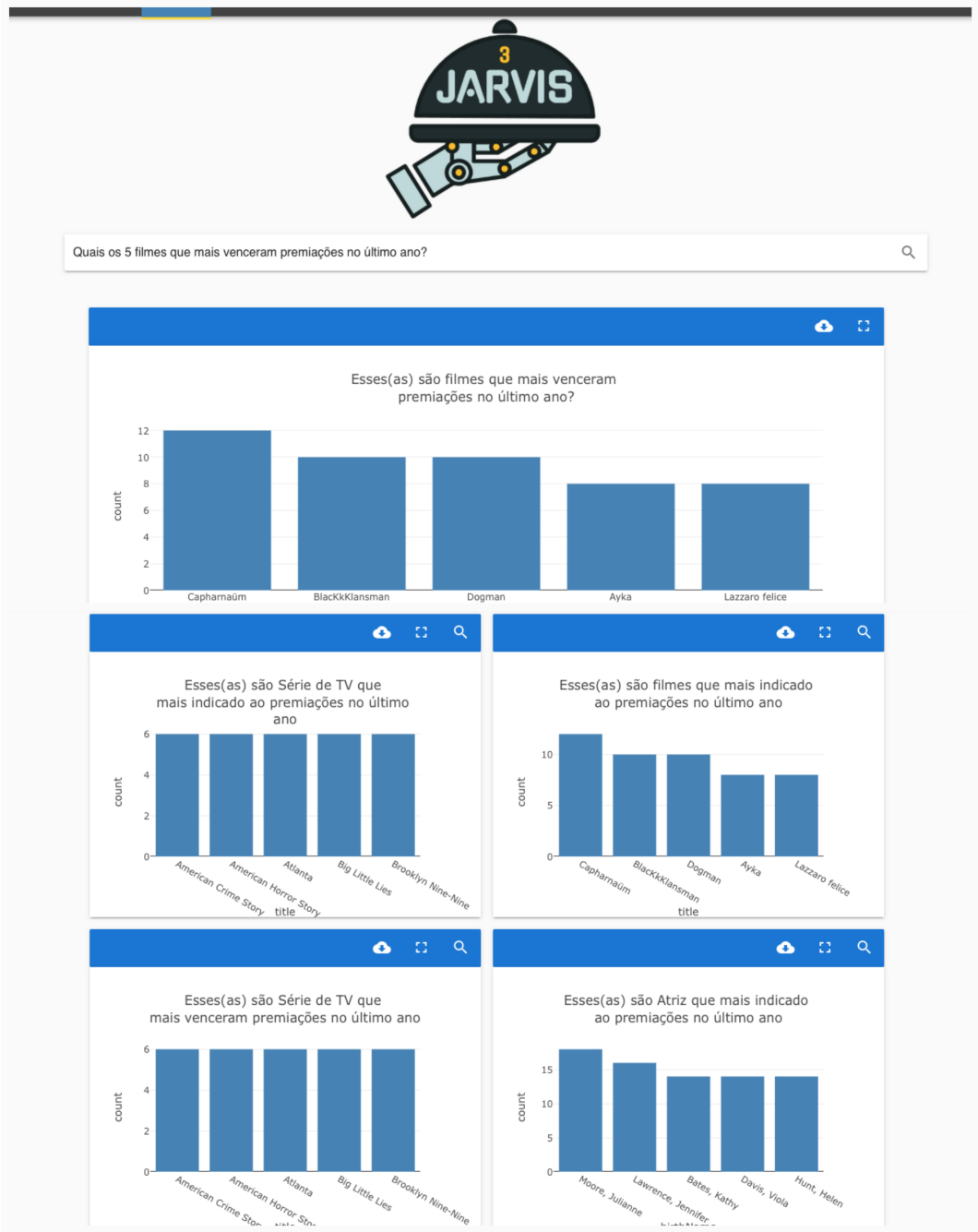



Figure 4.5: JARVIS search user interface

are currently being developed by other members of our research group, and therefore lie outside the scope of this work.

PUC-Rio - Certificação Digital Nº 1721764/CA

# 5
# Evaluation Study

This chapter describes how we evaluated our proposal. Section 5.1 discusses the alternative user interface models we developed to compare with our proposal. Section 5.2 describes the evaluation plan of our study.

## 5.1
## Alternative Search User Interface Models

As described in Chapter 4, our proposed solution progressively discloses results for related queries. To evaluate the effectiveness and efficiency of our solution, we have devised two other search user interface (SUI) models for the same search task. The first uses the traditional search interaction method described by Wilson (1999) (henceforth called *Traditional* SUI (J1)), and the second is built taking Google Images search result interface as inspiration (henceforth called *Related-links* SUI (J2)).

The *Traditional* SUI (J1) (Figure 5.2) is an almost direct representation of the work described by Wilson (1999). It presents a user interface where the user types a search query and receives the highest ranked result for their question. The only way the user can expand the search results is by manually editing or typing a new query for the system, which again will only return the highest ranked result. This model represents a baseline for our work, whereas the interface, although pedestrian, is straightforward and familiar to the participants.

The *Suggested-links* SUI (J2) (Figure 5.2) introduces a suggested list of related questions. The user is now presented not only with the highest ranked result, but also with a set of related questions on a lateral bar. That mechanism allows the user to navigate through related questions more quickly, but still requires manual interaction with the user interface. The model J2 is presented to the participants so we can attempt to understand whether the mere introduction of related questions is enough to reduce the users' cognitive overload and to build a more effective interface for our research context.

PUC-Rio - Certificação Digital Nº 1721764/CA

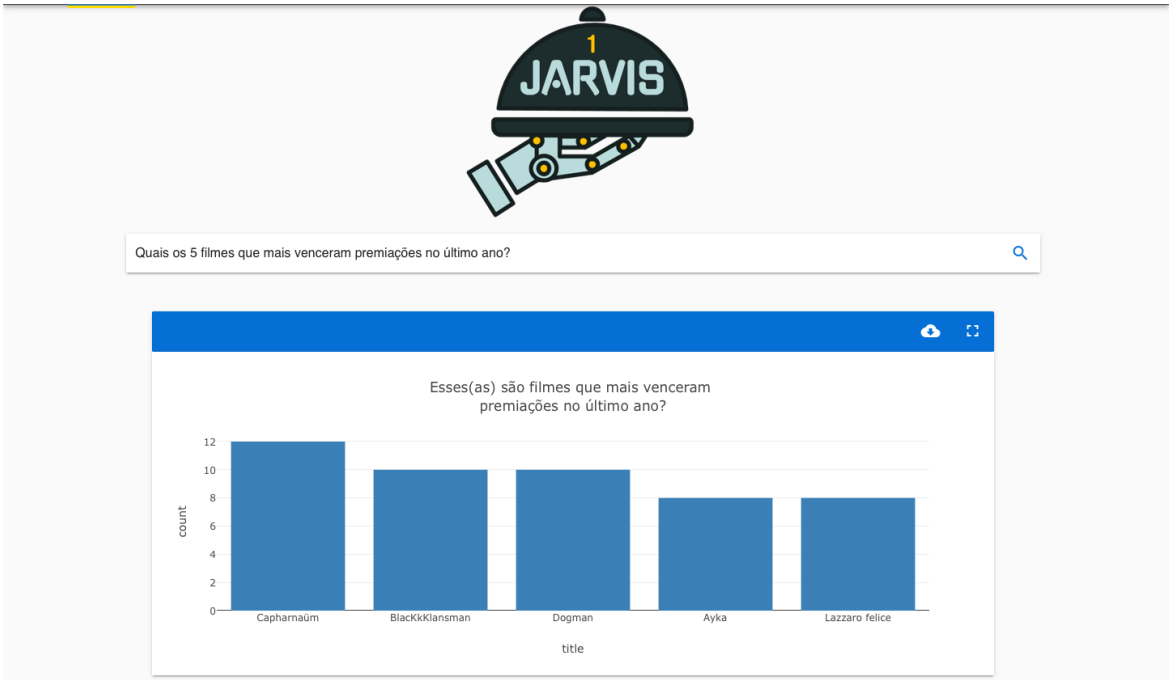



Figure 5.1: User interface of the Traditional SUI (J1)

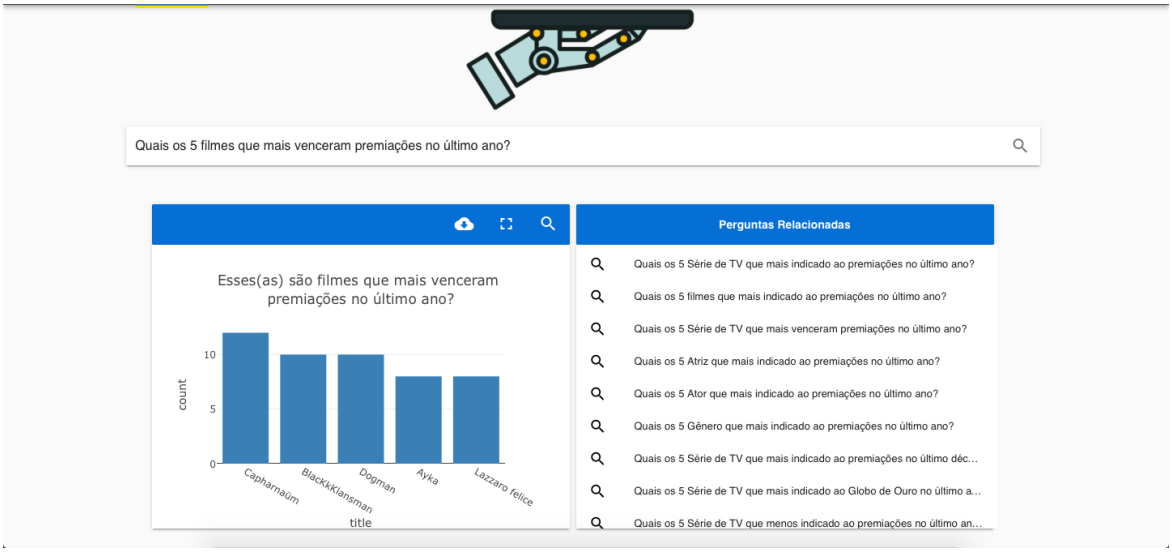



Figure 5.2: User interface of the Related-links SUI (J2)

## 5.2 Experimental Design

To evaluate JARVIS, we conducted a comparison test of the three SUIs. We invited graduate students from different areas to serve as volunteer participants in the study.

To reduce the learning effects, we varied the order in which each SUI is presented to the users, using the configuration shown in Table 5.1. With this, we attempted to see if the order on which the user experienced each model affected their evaluation. Since model J1 required more effort to complete the

PUC-Rio - Certificação Digital Nº 1721764/CA

suggested task, we hypothesized that users that had contact with the model J1 prior to the other models would find the introduction of mechanisms for searching the related queries more useful. That would be especially true with users on experiment A, where the mechanism scaled in complexity gradually. The user started the experiment with only a straightforward search mechanism in J1, to later test an interface that presents he/she with related queries in J2, to finally evaluate . This order would help to gradually raise the awareness of the related questions and its answers. Conversely, we also hypothesized that participants in experiment C could get confused with non-traditional features of model J3 and evaluate it poorly.

| Group | Order of SUIs |
|---|---|
| A | J1, J2, J3 |
| B | J2, J1, J3 |
| C | J3, J2, J1 |

Table 5.1: Experiment groups

PUC-Rio - Certificação Digital Nº 1721764/CA

Before the experiment, we explained the procedure to the participants and asked for their consent. The informed consent form is in Appendix B

For each SUI, the user received six search tasks, each one representing a search query. We devised the queries in two groups, one which had two related queries and other with four related queries, but we did not inform users of such grouping. Such grouping was designed so that in the Related-links SUI (J2) and JARVIS (J3), participants would need to type only two queries, and then they would have quick access to the remaining related queries through the links at the left-hand panel. In the Traditional SUI (J1), however, the user would need to type in each of the six queries manually. For each group, when the participant asked the first question, the results pages also presented either the questions of the following tasks (on J2) or their answers (on J3). The groups were also designed in such a way that the related question ranked high in each related queries mechanism of J2 and J3, except for the last question of the first group (see the question 4 of Appendix A), which was intentionally more distant and thus required the user to scroll the related queries component (on J2) or the screen (on J3).

The content of the tasks varied from asking the participants to discover the five movies that had won the most awards last year (through a *Movie–won award–Awards* relationship) and related information such as the 5 TV Series that won most awards last year (through a *TV Series –won award–Awards* relationship) and in the last decade (through a time variation). Appendix A

shows the material given to the participants, which presents the questions to be answered and a form in which to answer them (in Portuguese).

After interacting with each SUI, we asked the participant to fill out a questionnaire regarding the perceived ease of use and utility of the SUI – based on the Technology Acceptance Model (TAM) (Davis, 1989) –, and their subjective workload assessment – based on the NASA Task Load Index[1] (Hart, 2006). The full questionnaire can be found in Appendix C. At the end of the session, we will briefly interview the users, asking them to choose their preferred SUI and explain the factors that led them to their choice.

Besides the questionnaire and interview data, we also collected performance data in terms of effectiveness (correctness of the result) and efficiency (time on task).In particular, we used the number of searches as a proxy for effectiveness.

We expected that models J3 followed by J2 would present better results in the TAM questionnaire due to the introduction of mechanisms that offer more ways to explore the search results. However, participants may find that the new interfaces require from them a more significant effort, resulting in poor results on the NASA TLX.

The next chapter presents and discusses the results of the experiment.

PUC-Rio - Certificação Digital Nº 1721764/CA

[1]https://humansystems.arc.nasa.gov/groups/TLX/

# 6
# Results

This chapter presents and discusses the experiment results. Section 6.1 describes the study participants. Sections 6.2 and 6.3 discuss the results of the NASA Task Load Index and Technology Acceptance Model Questionnaires, respectively. In Section 6.4, we discuss the insights gathered during the interviews with each participant. Section 6.5 discusses the number of searches participants made with each model, as a proxy for efficiency. Section 6.6 describes how each participant ranked each search model, and how their previous knowledge of experimental group may have affected their perception. Finally, Section 6.7 discusses the study limitations we have identified.

PUC-Rio - Certificação Digital Nº 1721764/CA

## 6.1
## Study participants

Table 6.1 shows the study participants and their characteristics. Fifteen people participated in the experiment: three females (P01, P04, P14) and 12 males (P02, P03, P05, P06, P07, P08, P09, P10, P11, P12, P13, P14). They were all graduate students at PUC-Rio (11 Master's students, and 4 PhD students). Apart from P14, who is a psychology student, all the participants were Computer Science students.

Eleven participants (P01, P02, P03, P04, P05, P06, P07, P08, P10, P11, P15) fell within the 18-24 age group. Only four participants (P09, P12, P13, P14) were 25 to 44 years old.

Regarding their previous knowledge of the models, all participants were familiarized with traditional search tools. Four of the participants had already seen a search user interface similar to J3 in another context, but had not used it (P02, P03, P05, P06). We henceforth call these "participants with little previous knowledge". One participant – henceforth "Developer" – helped develop J3 for an R&D project (P11). The other ten participants had no knowledge of models J2 and J3 – henceforth "participants with no previous knowledge.

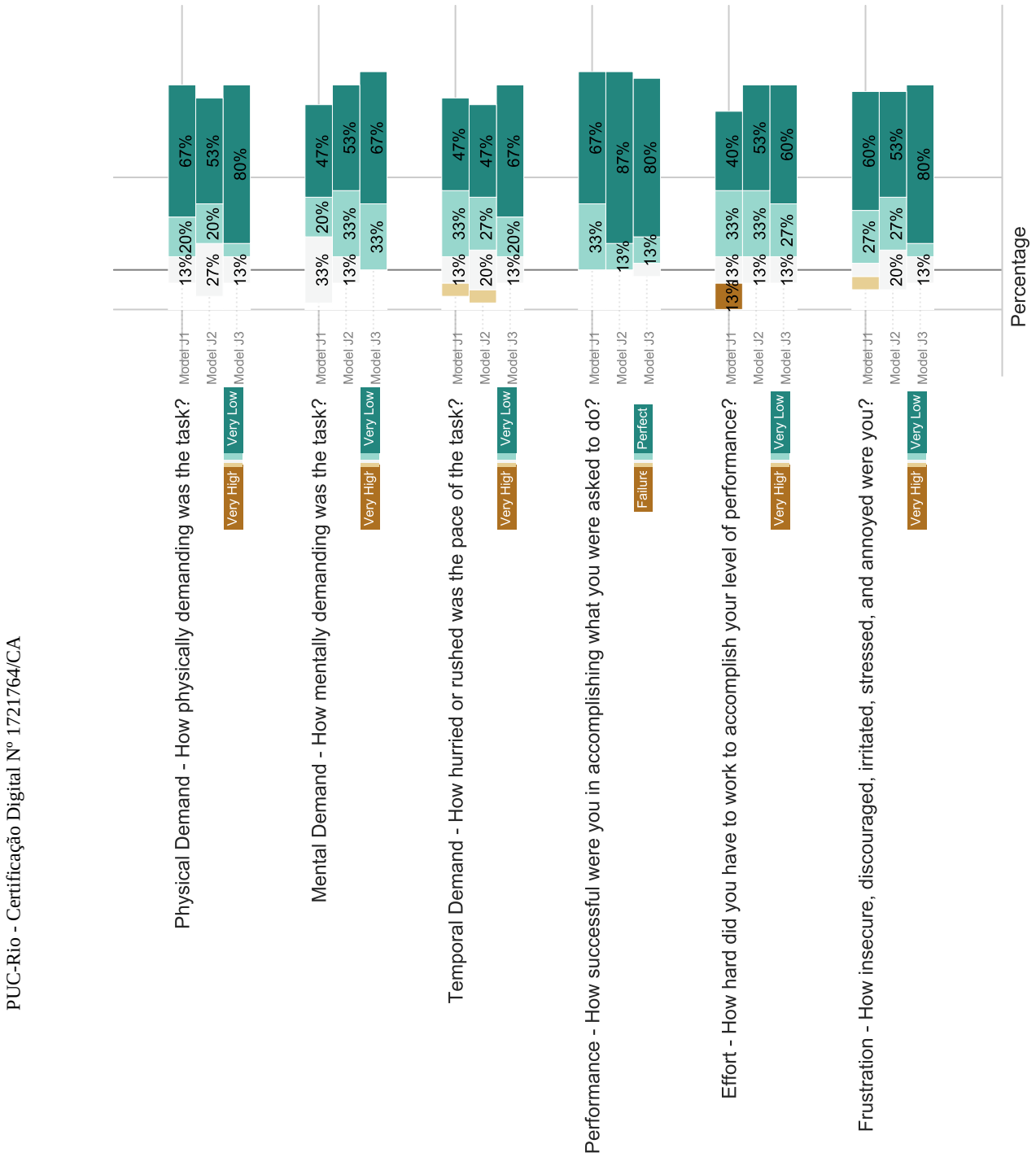
| Participants | Gender | Age group | Degree | Field of Work |
|---|---|---|---|---|
| P01 | Female | 18-24 | Master's student | Computer Science |
| P02 | Male | 18-24 | Master's student | Computer Science |
| P03 | Male | 18-24 | Master's student | Computer Science |
| P04 | Female | 18-24 | Master's student | Computer Science |
| P05 | Male | 18-24 | Master's student | Computer Science |
| P06 | Male | 18-24 | PhD student | Computer Science |
| P07 | Male | 18-24 | Master's student | Computer Science |
| P08 | Male | 18-24 | PhD student | Computer Science |
| P09 | Male | 25-44 | PhD student | Computer Science |
| P10 | Male | 18-24 | Master's student | Computer Science |
| P11 | Male | 18-24 | Master's student | Computer Science |
| P12 | Male | 25-44 | PhD student | Computer Science |
| P13 | Male | 25-44 | Master's student | Computer Science |
| P14 | Female | 25-44 | Master's student | Psychology |
| P15 | Male | 18-24 | Master's student | Computer Science |

Table 6.1: Characterization of study participants

## 6.2 NASA Task Load Index Results

Figure 6.1 shows the results of the entire NASA Task Load Index questionnaire, discussed in detail in the next subsections.

PUC-Rio - Certificação Digital Nº 1721764/CA




PUC-Rio - Certificação Digital Nº 1721764/CA

Figure 6.1: NASA Task Load Index results

The raw data, statistics, and hypotheses tests regarding the NASA TLX questionnaire can be found in Appendix D, Sections D.1.1 (for all users), D.2.1 (for users with little previous knowledge of the models), and D.3.1 (for users with no previous knowledge of the models).

### 6.2.1 Mental Demand (NASA TLX)

Figure 6.2 shows the results of the Mental Demand measurement of the NASA TLX.

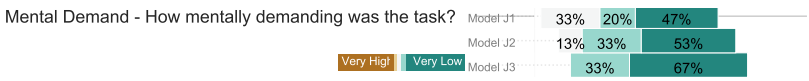



Figure 6.2: NASA Task Load Index - Mental Demand Results

The task in the experiment was relatively simple. However, as the complexity of the models grew (J2 is more complicated than J1, and J3 is more complex than J2), we expected that the mental demand scores for J1 would be better than for J2 and J3. During the interviews, this hypothesis seems to gather even more support mainly because of what participants (P01, P02, P03, P04, P010, P13, P14) called "lack of resources" of model J1 and their assessment of J2 and J3. Regarding the other models, we expected results from the questionnaire to follow the comments of the participants that the models J2 (P11, P14) and J3 (P3, P4, P5, P8, P9, P11, P13, P14) would be better suited to multiple search tasks, thus ranking worse than J1.

Figure 6.2 seems to be in accordance with our hypothesis and shows an advantage of J1 for the *Mental demand* measurement over the other models. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see Appendix D.1.1.1), users with little knowledge (see Appendix D.2.1.1)) or only users with no knowledge (see Appendix D.3.1.1).

### 6.2.2 Physical Demand (NASA TLX)

Figure 6.3 shows the results of the Physical Demand measurement of the NASA TLX.

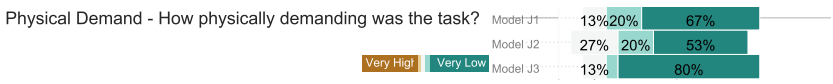



Figure 6.3: NASA Task Load Index - Physical Demand Results

The Physical demand measurement may be a reflection of the number of clicks or the number times the user has to manually inform the query in the main search bar to complete the tasks. In J1, the user had no other choice but to insert the search queries six times, so we might hypothesize that, although the task itself was not physically troublesome, models J2 and J3 would be

PUC-Rio - Certificação Digital Nº 1721764/CA

rated slightly better because they offered options that did not involve typing or copy-and-pasting new queries in to the system.

Figure 6.3 seems to show a slight advantage of J3 for the *Physical demand* measurement over the other models. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.1.2), only users with little knowledge (see D.2.1.2), and only users with no knowledge (see D.3.1.2)).

### 6.2.3 Temporal Demand (NASA TLX)

Figure 6.4 shows the results of the Temporal Demand measurement of the NASA TLX.

Figure 6.4: NASA Task Load Index - Temporal Demand Results

Since model J1 required more interaction and clicks at the user interface than models J2 and J3, we hypothesized that J1 would perform worse on this measurement than the other models. However, during the interviews, we noticed that the model J2 had drawn polarized opinions from participants. This may have been an effect from the limitations of J2. However, we believe that faulty design implementation presented in J2 played a prominent role on those participants commentaries. Considering the component was not very intuitive, and the text font was quite small, the interaction with the component was deeply affected, and its problems may have overshadowed its virtues.

Figure 6.4 seems to be in accordance with our hypothesis and shows an advantage of J3 for the *Temporal demand* measurement over J1. Unsurprisingly, model J2 seems to have a slight worse temporal demand evaluation than the others models. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.1.3), only users with little knowledge (see D.2.1.3), and only users with no knowledge (see D.3.1.3)).

### 6.2.4 Performance (NASA TLX)

Figure 6.5 shows the results of the Performance measurement of the NASA TLX.

PUC-Rio - Certificação Digital Nº 1721764/CA

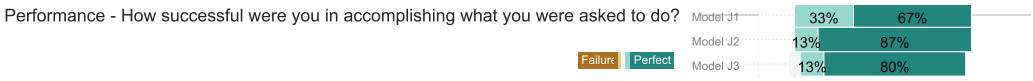



Figure 6.5: NASA Task Load Index - Performance Results

The Performance measurement results may have been profoundly affected by a severe problem with the experiment: the server performance. Because the server was very fragile, the system was slower than usual. This problem affected primarily the user interface of the J3 model, which is, by far, the model that needs to receive a larger volume of data to build the visualizations at the user interface. These issues may have influenced the performance scores in the questionnaire, leaving J2 slightly better ranked than J3.

Figure 6.5 seems to be in accordance with our hypothesis, showing an advantage of J2 for the *Performance demand* measurement over the other models. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.1.4), only users with little knowledge (see D.2.1.4), and only users with no knowledge (see D.3.1.4)).

### 6.2.5 Effort (NASA TLX)

Figure 6.6 shows the results of the Effort measurement of the NASA TLX.

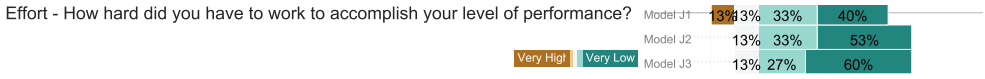



Figure 6.6: NASA Task Load Index - Effort Results

Similar to the Physical demand measurement described in subsection 6.2.2, the effort measurement may be a reflection of the number of times the user had to manually inform a query in the main search bar to complete the tasks. Surprisingly, users ranked even J3 poorly, acknowledging that the user interface could further reduce the user effort on searching. Moreover, J1 is the only model that offers no other option to complete the task, *i.e.*, it requires that all queries be informed, one by one, in the search bar. For that matter, we hypothesized that J3 would perform better than J2 and that J2 would perform better than J3.

Figure 6.6 seems to be in accordance with our hypothesis and shows an advantage of J3 for the *Effort* measurement over the other models. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.1.5), only users with little knowledge (see D.2.1.5), and only users with no knowledge (see D.3.1.5)).

PUC-Rio - Certificação Digital Nº 1721764/CA

### 6.2.6 Frustration (NASA TLX)

Figure 6.7 shows the results of the Frustration measurement of the NASA TLX.

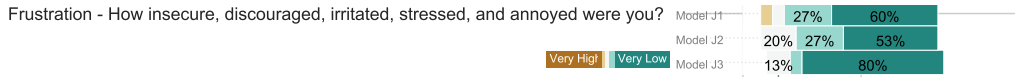



Figure 6.7: NASA Task Load Index - Frustration Results

Because J3 is the model that it was more complicated than the others, it also had the effects of the problems with the server being more prominent in the J3. When the participants were exploring the interface of J3, the server would often crash and require a reboot in other to become functional again. Because of that, we expected to J3 to be lowest-ranked model in this measurement.

Surprisingly, Figure 6.7 seems to debunk our hypothesis and shows a slight advantage of J3 for the *Frustration* measurement over the other models, even with the problems on the interface design and the server malfunctions. J2, for which participants condemned the amount of text required to read in order to interact with the related questions, J1 closely behind J1. However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.1.6), only users with little knowledge (see D.2.1.6), and only users with no knowledge (see D.3.1.6)).

## 6.3 Technology Acceptance Model (TAM) Results

Figures 6.8 and 6.9 show the results of the entire Technology Acceptance Model questionnaire, discussed in detail in the next subsections.

PUC-Rio - Certificação Digital Nº 1721764/CA

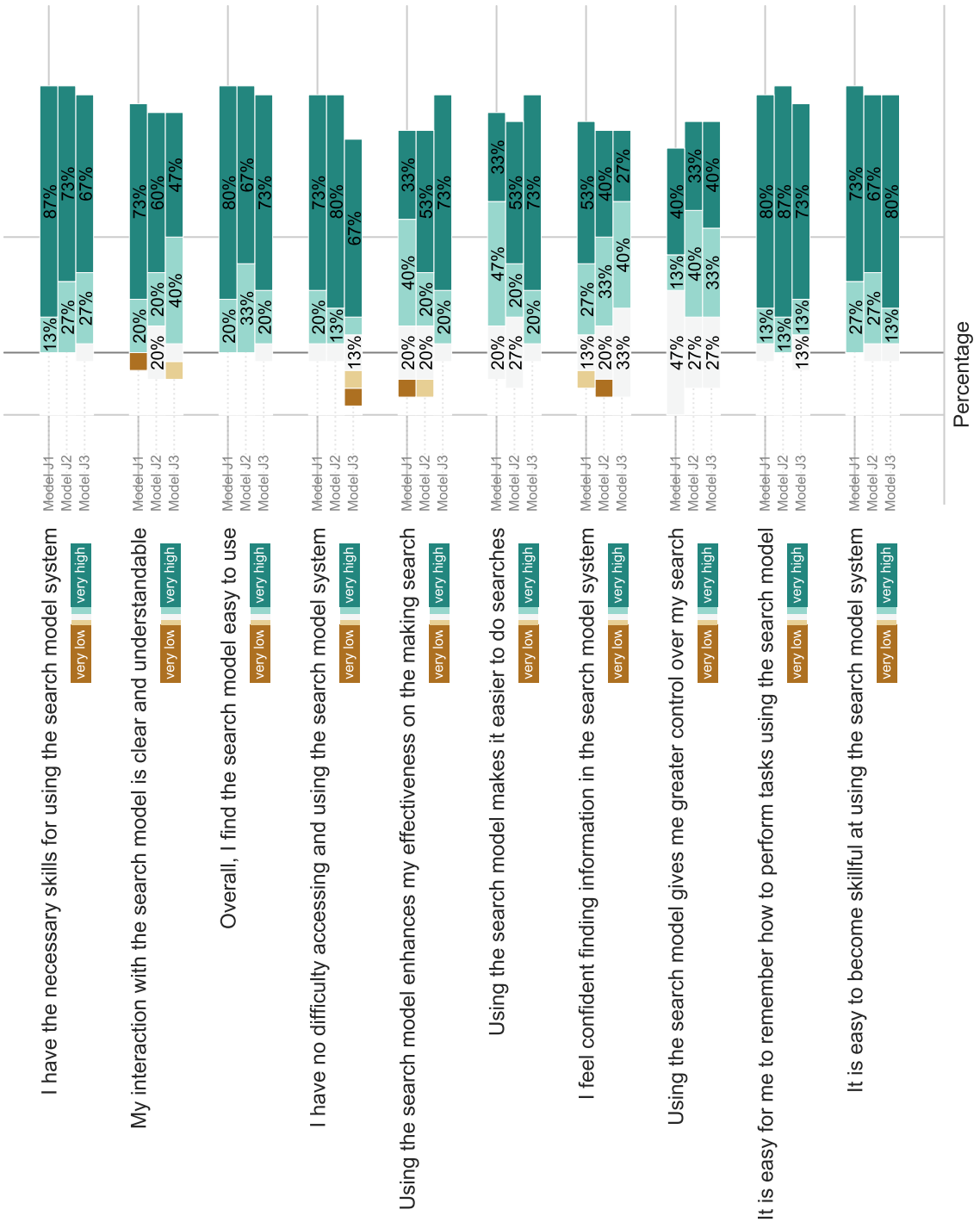



Figure 6.8: TAM results (part 1)

PUC-Rio - Certificação Digital Nº 1721764/CA

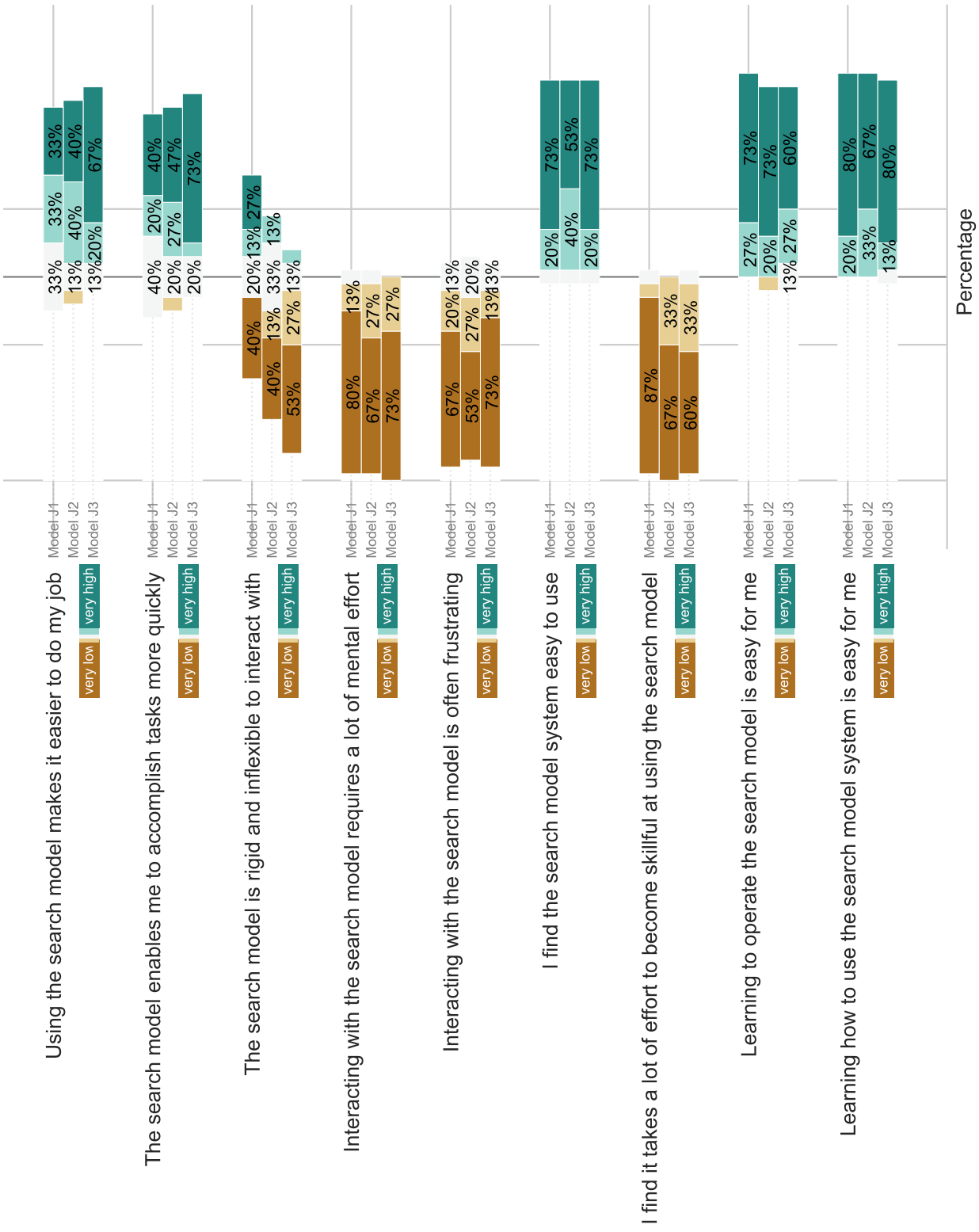



Figure 6.9: TAM results (part 2)

The raw data, statistics, and hypotheses tests regarding the TAM questionnaire can be found in Appendix D, Sections D.1.2 (for all users), D.2.2 (for users with little previous knowledge of the models), and D.3.2 (for

PUC-Rio - Certificação Digital Nº 1721764/CA

users with no previous knowledge of the models).

### 6.3.1 I find the search model X easy to use

Figure 6.10 shows the results of the TAM item “I find the search model system X easy to use”.

Figure 6.10: The search model X is easy to use - Results

Because of the nature of the task, all models performed reasonably well. The figure shows the model J2 as the worst-rated and the models J1 and J3 tied. This seems to contradict our hypothesis that the model J1 is easier to use than models J2 and J3. Instead, it shows identical results to J1 and J3. This means that, at least from the user perspective, JARVIS(J3) is not harder to use than the other models, even having more complex features than the other models. The unfortunate result of the model J2 may be a consequence of their faulty design implementations of the model, that some participants reported as "confusing."

However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.1), only users with little knowledge (see D.2.2.1), and only users with no knowledge (see D.3.2.1)).

### 6.3.2 Using the search model X gives me greater control over my search

Figure 6.11 shows the results of the TAM item “Using the search model X gives me greater control over my search”.

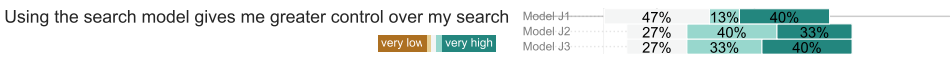



Figure 6.11: The search model X gives me greater control over my search - Results

Some participants (P13, P14) found the question confusing and asked for clarification on what control means in the context of the task they were doing. For them, we defined control as “The degree to which he/she believes that the model supports the use of the system to perform the desired behaviour”. We hypothesize that, because the model J1 offered to the user only one way to interact and control the search, J1 would present worse results than J2 and J3.

PUC-Rio - Certificação Digital Nº 1721764/CA

The results shown in Figure 6.11 seem to confirm our hypothesis, with model J3 having a slight advantage over model J2.

However, A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.2), only users with little knowledge (see D.2.2.2), and only users with no knowledge (see D.3.2.2)).

### 6.3.3 The search model X enables me to accomplish tasks more quickly

Figure 6.12 shows the results of the TAM item "The search model X enables me to accomplish tasks more quickly".

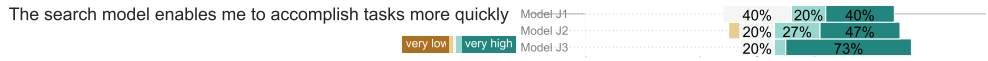



Figure 6.12: The search model X enables me to accomplish tasks more quickly - Results

With this item, our goal was to evaluate whether the perceived evaluation from the participants matched their actual time on task. Unfortunately, due to the problems described in chapter **??**, we are unable to execute this analysis. We expected that model J2 and J3 would perform better than J1 in this item because it offers the answer or question for the other task more efficiently, not requiring the participant to type each search query to finish the experiment tasks. The results shown in Figure 6.12 confirms our prediction for the model J3, which had the best evaluation among the participants. Despite presenting to the user an alternative way to search, the model J2 was the worst-ranked among all three models. This results might be an indication that the mere recommendation of related questions is not enough to support the user while navigating on search result pages more quickly. This result might also be an outcome of an unrefined designed interface.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.3), only users with little knowledge (see D.2.2.3), and only users with no knowledge (see D.3.2.3)).

### 6.3.4 Using the search model X enhances my effectiveness on making a search

Figure 6.13 shows the results of the TAM item "Using the search model X enhances my effectiveness on the job".

PUC-Rio - Certificação Digital Nº 1721764/CA

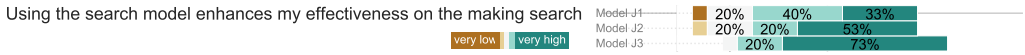



Figure 6.13: Using the search model X enhances enhances my effectiveness on making a search - Results

In this item, we hypothesized that the model J3 would perform better than the other models due to the addition of related answers in the interface. Conversely, the lack of resources from model J1 would result in a worse evaluation from participants regarding the other models. Figure 6.13 shows that the data seems to confirm our expectations, ranking the model J3 better than J2 and J1, respectively.

However, A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.4), only users with little knowledge (see D.2.2.4), and only users with no knowledge (see D.3.2.4)).

### 6.3.5
### Using the search model X makes it easier to do my job

Figure 6.14 shows the results of the TAM item "Using the search model X makes it easier to do my job".

Figure 6.14: Technology Assessment Model - Makes it easier to do my job Results

In this item, some participants were asked whether each model makes it easier for them to do their job. We hypothesized that our suggested model J3 would perform better than models J1 and J2. Moreover, that J2 would perform better than J1. The figure appears to confirm our expectation regarding J3, but it debunks our expectation showing J2 with a worse evaluation than J1. This effect may occur due to the known problems with the design of the interface, that participants (P3, P4, P7) said *"had to too much text"*.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.5), only users with little knowledge (see D.2.2.5), and only users with no knowledge (see D.3.2.5)).

### 6.3.6
### Using the search model X makes it easier to do searches

Figure 6.15 shows the results of the TAM item "Using the search model X makes it easier to do searches".

PUC-Rio - Certificação Digital Nº 1721764/CA

Figure 6.15: Using the search model X makes it easier to do searches - Results

The figure shows that the model J1 had the worst evaluation by the participants. This result is in line with our expectations, since the model offers the user only one way to search, which makes it harder for the users to perform multiple searches. The excellent results from the model J3 (all participants rated the model with high or very high scores) are promising and in line with the comments from the participants. Those participants believe that the related questions or answer are features that support multiple search task better than the model J1. Eight (P3, P4, P5, P7, P8, P10, P13, P14) of the participants cited the model J3 as a better alternative to multiple search tasks while five (P3, P7, P12. P13, P14) also cited J2 as a more suitable option than J1 for the same kind of tasks.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.6), only users with little knowledge (see D.2.2.6), and only users with no knowledge (see D.3.2.6)).

### 6.3.7
### Learning how to use the search model X system is easy for me

Figure 6.16 shows the results of the TAM item "Learning how to use the search model X system is easy for me".

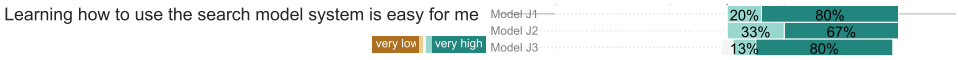



Figure 6.16: Learning how to use the search model X system is easy for me - Results

In this item, we expected the model J1 to present the best result because it represents a simple, but familiar, interaction model to the user, more than the models J2 and J3. In particular, model J3 introduces to the user features that he/she may not grasp at first sight. So we expected that participants would rate the model lower. Surprisingly the figure seems to debunk our hypothesis showing slightly better results to J3 than to J1.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.7), only users with little knowledge (see D.2.2.7), and only users with no knowledge (see D.3.2.7)).

PUC-Rio - Certificação Digital Nº 1721764/CA

### 6.3.8
### Learning to operate the search model X is easy for me

Figure 6.17 shows the results of the TAM item "Learning to operate the search model X is easy for me".

Figure 6.17: Learning to operate the search model X is easy for me - Results

In this item, the results shown in Figure 6.17 seem to align much better with our hypothesis. The model J1 is very straightforward and familiar and requires less effort to learn how to operate than models J2 and J3, which introduce novice elements to the search result page. The better result from J2 may be an indication that because effectively it shows fewer elements than model J3, it fits better the screen, as said by two of the participants (P3 e P6).

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.9), only users with little knowledge (see D.2.2.9), and only users with no knowledge (see D.3.2.9)).

### 6.3.9
### Interacting with the search model X is often frustrating

Figure 6.18 shows the results of the TAM item "Interacting with the search model X is often frustrating".

Figure 6.18: Interacting with the search model X is often frustrating - Results

Due to the problems with the experiment, we hypothesized that model J3 would perform worse than the other models, since most of the problems with the server deeply affected its resources. Model J3 also was the model where the server failed multiple times. Nevertheless, Figure 6.18 shows that the results of the TAM item "Interacting with the search model X is often frustrating" contradicts our hypothesis. It shows the model J3 with better evaluation among the models, and model J2 with the worst evaluation.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.10), only users with little knowledge (see D.2.2.10), and only users with no knowledge (see D.3.2.10)).

PUC-Rio - Certificação Digital Nº 1721764/CA

### 6.3.10
### The search model X is rigid and inflexible to interact with

Figure 6.19 shows the results of the TAM item "The search model X is rigid and inflexible to interact with".

Figure 6.19: The search model X is rigid and inflexible to interact with - Results

For this item, we hypothesized that the model J1 would present the worse ratings between all models, because it is rigid and inflexible by design. Participants of the experiment (P1, P11, P13, P14) added that the model J1 is the best for when the user does not know what he or she wants. In contrast, models J2 and J3 presented alternative ways for the user to explore the data, and we expected that they would not be perceived as rigid and inflexible. The results in Figure 6.19 seem to confirm our hypothesis, with model J3 being favored by participants over model J2.

However, a Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.11), only users with little knowledge (see D.2.2.11), and only users with no knowledge (see D.3.2.11)).

### 6.3.11
### It is easy for me to remember how to perform tasks using the search model X

Figure 6.20 shows the results of the TAM item "It is easy for me to remember how to perform tasks using the search model X".

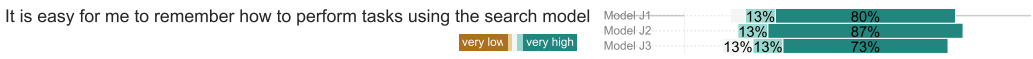



Figure 6.20: It is easy for me to remember how to perform tasks using the search model X - Results

This item had surprising results when taking into consideration the other items, because it is the only one in which J2 performed better than the J1 and J3. We expected model J1 to be better than J2 and J3 because as almost all participants (with the exception of P10 and P11) noted that it has a clear and straightforward interface. This is also the only item whereas the model J2 outperformed both model J1 and J3. Debunking our expectations that model J1, due to its lack of complexity, would rank better than both more complex models. The worst result from the model J3 could be an effect of the confusing and distracting interface, as noted by some participants (P04, P05, P11).

PUC-Rio - Certificação Digital Nº 1721764/CA

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.12), only users with little knowledge (see D.2.2.12), and only users with no knowledge (see D.3.2.12)).

### 6.3.12
### Interacting with the search model X requires a lot of mental effort

Figure 6.21 shows the results of the TAM item "Interacting with the search model X requires a lot of mental effort".

Figure 6.21: Interacting with the search model X requires a lot of mental effort - Results

Again, the model J1 performed better than the other models as we expected and in line with the results presented in Section 6.2 by the NASA TLX questionnaire. Since J1 presented a clean and simple interface, the effort to interact with the search model was almost minimal. The results in Figure 6.21 support this idea and are in line with the comments from the participants, who believe that the model J3 has an information overload problem (P12) and an interface that can be confusing to the user (P04, P05, P11).

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.13), only users with little knowledge (see D.2.2.13), and only users with no knowledge (see D.3.2.13)).

### 6.3.13
### My interaction with the search model X is clear and understandable

Figure 6.22 shows the results of the TAM item "My interaction with the search model X is clear and understandable".

Figure 6.22: My interaction with the search model X is clear and understandable - Results

Figure 6.22 seems to confirm our hypothesis that the model J1 is, due to its lack of features, more clear and understandable than models J2 and J3. The figure also shows a slight advantage of model J2 over model J3. That may be related to the "information overload" problem that a participant (P12) reported in the interviews.

PUC-Rio - Certificação Digital Nº 1721764/CA

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.14), only users with little knowledge (see D.2.2.14), and only users with no knowledge (see D.3.2.14)).

### 6.3.14
### I find it takes a lot of effort to become skillful at using the search model X

Figure 6.23 shows the results of the TAM item "I find it takes a lot of effort to become skillful at using the search model X".

Figure 6.23: I find it takes a lot of effort to become skillful at using the search model X - Results

This item also follows a pattern that we expected, with the most familiar user interface, model J1, with better results. This effect was deeply influenced by the simple and familiar structure of the interface. Regarding the other two models, J3 was favored by participants over J2. This may be a result of the exhausting interface, as noted by some participants (P08, P11, P13).

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.15), only users with little knowledge (see D.2.2.15), and only users with no knowledge (see D.3.2.15)).

### 6.3.15
### Overall, I find the search model X easy to use

Figure 6.24 shows the results of the TAM item "Overall, I find the search model X easy to use".

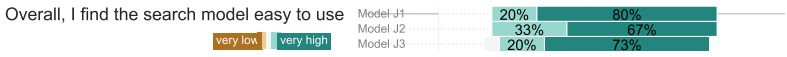



Figure 6.24: Overall, I find the search model X easy to use - Results

Later in the questionnaire, we asked a question very similar to the one in Section 6.3.1. We intended to check whether, after evaluating various aspects in the questionnaire, the user's initial opinion remained the same. The results Figure 6.24 shows that the distribution stayed the same except for one participant on he/she evaluation of model J3. Because of that, the model J3 rating has decreased, but still performing better than model J2.

PUC-Rio - Certificação Digital Nº 1721764/CA

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.16), only users with little knowledge (see D.2.2.16), and only users with no knowledge (see D.3.2.16)).

### 6.3.16
### It is easy to become skillful at using the search model X

Figure 6.25 shows the results of the TAM item "It is easy to become skillful at using the search model X".

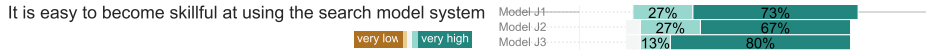



Figure 6.25: It is easy to become skillful at using the search model X - Results

Figure 6.25 debunks our expectations for this item. Although it shows the participants evaluated model J1 better than the other models, it also shows model J2 behind J3 in the evaluation. This may be a reflection of the participants' perception of the model J2 presented during the interviews; some participants reported that the related questions component of J2 was exhausting (P08, P11, P13). Most of this perception was based on the amount of text in the component. This notion is further expanded when participants (P01, P04, P11, P12) called J3 "too confusing".

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.17), only users with little knowledge (see D.2.2.17), and only users with no knowledge (see D.3.2.17)).

### 6.3.17
### I feel confident finding information in the search model X system

Figure 6.26 shows the results of the TAM item "I feel confident finding information in the search model X system".

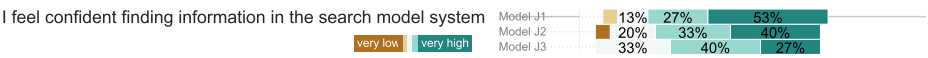



Figure 6.26: I feel confident finding information in the search model X system - Results

The surprise in this item comes from the evaluation of model J2, which performed even worse than model J1. That may be because the related questions are presented in a very similar structure. For example, in a scenario where the user asks "What are the 2018 best-rated movies?" and is also interested in the "What are the 2017 best-rated movies?" this related question, even

PUC-Rio - Certificação Digital Nº 1721764/CA

when highly ranked in the component, may get confused in between many year variations of “What are the [year] best-rated TV series?” and “What are the [year] worst-rated TV series?”.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.18), only users with little knowledge (see D.2.2.18), and only users with no knowledge (see D.3.2.18)).

### 6.3.18
### I have the necessary skills for using the search model X system

Figure 6.27 shows the results of the TAM item “I have the necessary skills for using the search model X system”.

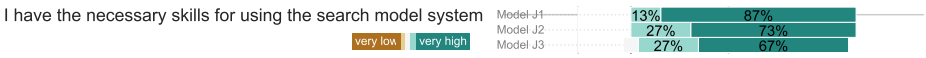



Figure 6.27: I have the necessary skills for using the search model X system - Results

As the complexity of the models grow (J2 is more complex than J1, and J3 is more complex than J2), we expected that the participants would already feel skillful to use model J1, but would need more time to grasp the mechanisms of the models J2 and J3 truly. With model J3 being more complex, it was no surprise that Figure 6.27 confirmed our expectation showing the model J1 as the highest-rated, and model J3 as lowest-rated.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.19), only users with little knowledge (see D.2.2.19), and only users with no knowledge (see D.3.2.19)).

### 6.3.19
### I have no difficulty accessing and using the search model X

Figure 6.28 shows the results of the TAM item “I have no difficulty accessing and using the search model X”.

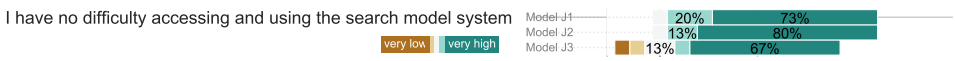



Figure 6.28: I have no difficulty accessing and using the search model X - Results

Figure 6.27 confirms our expectation that model J1 is more familiar to users than model J2 and J3. Model J1 requires users to type in all queries, whereas models J2 and J3 require users to type in fewer queries and select

PUC-Rio - Certificação Digital Nº 1721764/CA

related results. In contrast, J2 and J3 require users to read the related queries, searching for those of interest.

A Kruskal-Wallis hypothesis test showed no significant difference among the models, at $\alpha = 0.05$ (either considering all users (see D.1.2.20), only users with little knowledge (see D.2.2.20), and only users with no knowledge (see D.3.2.20)).

## 6.4 Interviews

Table 6.2 shows a compilation of the common critiques reported by the participants of the study. We categorized each comment into four categories: Model, Design, Configuration, and Implementation. In *Model*, we selected the comments more closely related to intrinsic aspects of each model and which would likely remain true even if significant changes to the design or implementation of the system were made. In *Design*, we summarized comments related to how the model was implemented at the user interface level; critiques made in this category could be solved or reduced by a better user interface design. The comments reported under the *Configuration* category are those that may be related to the parametrization and flexibility of how the related questions are calculated and prioritized. In *Implementation*, we compile the observations strongly associated with the system performance.

### 6.4.1 Model

Regarding the Model, most of the comments from the participants are related to in which occasions they seem to be adequate and what the model's intrinsic, positive or negative characteristics are which had the most significant influence on the participants' experience. Participants mostly commented on the differences between model J1 and model J2 and J3. The majority (12 participants) of them found that the model J1 is simple or straightforward, a characteristic that they consider a positive aspect of the model. In contrast, they found that executing the tasked asked in the model J1 was very time consuming for multiple searches (5 participants) and better suited for when the user knows what he or she wants (4) or only needs to do a quick search (4 participants). Although in the participants' view the model J1 lacks resources (7 participants), some of the participants perceive this as a positive aspect, because the interface does not distract the user from the task he/she is executing (2 participants).

These perspectives of model J1 contrast with what participants said about the models J2 and J3. Furthermore, whereas they found it difficult to use model

PUC-Rio - Certificação Digital Nº 1721764/CA

J1 in long or exploratory tasks, they highlighted the benefits of using the other models for those scenarios. For example, participants reported (2 participants) that, because both models presented them with related questions, they were able to have insights of new questions that they may not have had if they were only exploring the data following the traditional interface of model J1. They also found that both models support scenarios where the user needs to make multiple searches or search tasks that are broader or exploratory (2 participants for J2 and 4 for J3). Those comments are in line with the conceptual design of those models, as described in Chapter 5.

Besides that, a crucial point for this research is how the recommendations affected the user interaction. We did not directly ask of participants questions about the perceived effectiveness of the recommendation to avoid inducing certain answers. However, participants spontaneously evaluated aspects of the recommendation. The majority (7 for the model J2 and 10 for model J3) agreed that the recommendation makes the search task easier to be achieved. Those comments are an indication that the models J2 and J3 are potentially efficient and useful for the search task, especially when the user is on an exploratory task.

PUC-Rio - Certificação Digital Nº 1721764/CA

### 6.4.2 Design

In the Design category, we outlined the most common complaint from the participants that we believe could be solved by redesign the model interface. These are features such as the related questions component from model J2. Although it does what is supposed to do, the component could have been better designed to highlight the differences between questions to reduce the amount of text the user needs to read. Eight participants evaluated that the model J2 was harder to interact because it had too much text to read, and they evaluated it as worse to read than charts. Three of those participants even added that the model J2 was exhausting for them to use.

However, regarding the differences within model J2 and J3, there is a trade-off between using the chart or the text component. Although participants noted that the use of text could be exhausting, they also elucidated on problems with model J3, such as the need to scroll the interface (3 participants) and some (3 participants) even stated that they found the interface distracted them from the task they need to perform.

### 6.4.3 Configuration

The Configuration category represents issues that may be related to the parametrization and flexibility of how the related questions are calculated and prioritized.

In both models, J2 and J3, users complained about the selection of the related question/answers that were exhibited. Although this mechanism is out of the scope of this dissertation, we must develop a better algorithm for related questions/answer in order to better evaluate the models we are proposing. Participants even commented about the lack of coherence between the related questions and the main query the typed. In other words, in order for models such as J2 and J3 to thrive, as a better alternative to the design of search result pages, is vital that the ranking engine of those questions/answer be effective.

There are also other notable problems with the configurations of the models, such as views that should not be exhibited because it has no data or the data it is inconsistent. Nevertheless, the participants did not mention those problems during the interviews, either because they did not notice them or because it was not important enough for them.

### 6.4.4 Implementation

Concerning the implementation category, we are interested in how the system performance affected the participants' experience with each model. Since our implementation suffered from many issues, as mentioned before, we looked in the interviews on what are how those problems their influenced their evaluation.

In that regard, two participants reported that the model J3 is too slow. This is a direct effect of the fact that the model J3 was the one most affected by those problems.

PUC-Rio - Certificação Digital Nº 1721764/CA

PUC-Rio - Certificação Digital Nº 1721764/CA



| | | | Total | P01 | P02 | P03 | P04 | P05 | P06 | P07 | P08 | P09 | P10 | P11 | P12 | P13 | P14 | P15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J1 | The model fits the screen better | Implementantion | 5 | | | | X | X | X | | | | | X | | | X | |
| | Simple / Straightfoward | Model | 12 | X | X | X | | X | X | X | X | X | | | X | X | X | X |
| | Lacks resources | | 7 | X | X | X | X | | | | | | X | | | X | X | |
| | Time-consuming for multiple searches | | 5 | X | | | X | | | | | | X | X | | X | | |
| | Better for a quick search | | 4 | X | | | | | | | | | | X | | X | X | |
| | Better for when the user knows what it wants | | 4 | X | | | | | | | | | | X | | X | X | |
| | Don't distract the user | | 2 | | | | | X | | | | | | | | | X | |
| J2 | Can be confusing for the user | Configuration | 4 | X | | | X | | | | | | | X | X | | | |
| | The recommendation can be better ranked | | 2 | | X | | | | | | X | | | | | | | |
| | Reading questions is worse than reading charts | Design | 8 | X | | X | | X | X | | X | | | X | | X | X | |
| | Exhaustive | | 3 | | | | | | | | X | | | X | | X | | |
| | The model fits the screen better | Implementation | 2 | | | X | | | X | | | | | | | | | |
| | Recommendations make the task easier | Model | 7 | | | X | X | X | X | | | X | | | | X | X | |
| | Makes the search task broader/explorative | | 5 | | | X | | | | X | | | | | X | X | X | |
| | Better for multiple search tasks | | 2 | | | | | | | | | | | X | | | X | |
| | The model gives me insights of new questions | | 2 | | X | | | | | | | | | | | | X | |
| J3 | The recommendation can be better ranked | Configuration | 5 | | X | | X | | | | | X | | X | | | | X |
| | The need to scroll is bad | Design | 3 | | | | | | X | | | X | | | | X | | |
| | Can be confusing/distracting for the user | | 3 | | | | | X | X | | | | | | X | | | |
| | Too slow | Implementation | 2 | | | | X | | | | | X | | | | | | |
| | Has an information overload problem | | 1 | | | | | | | | | | | | X | | | |
| | Recommendations make the task easier | Model | 10 | X | | X | X | X | X | | X | X | | X | | X | X | |
| | Better for multiple search tasks | | 8 | | | X | X | X | | | X | X | | X | | X | X | |
| | Makes the search task broader/explorative | | 4 | | X | | | X | | X | | | | | | | X | |
| | The model gives me insights of new questions | | 2 | | X | | | | | | | | | | | | X | |
| | Better for when the user doesn't know what it wants | | 2 | X | | | | | | | | | | | | | X | |
| | | | | | | | | | | | | | | | | | | |

Table 6.2: Compilation of Interview Results

## 6.5 Efficiency

As described in the limitations of the study in Section 6.7, we were unable to directly measure the time on task. However, we counted the number of explicit searches the participants did during the experiment. This can be considered as an indirect indication of efficiency of each model. It is flawed, however, as it does not take into account the time it took for participants to locate the related questions when using models J2 and J3.
Figure 6.29 shows the number of searches made in each model, by each group, according to the participants' previous knowledge.

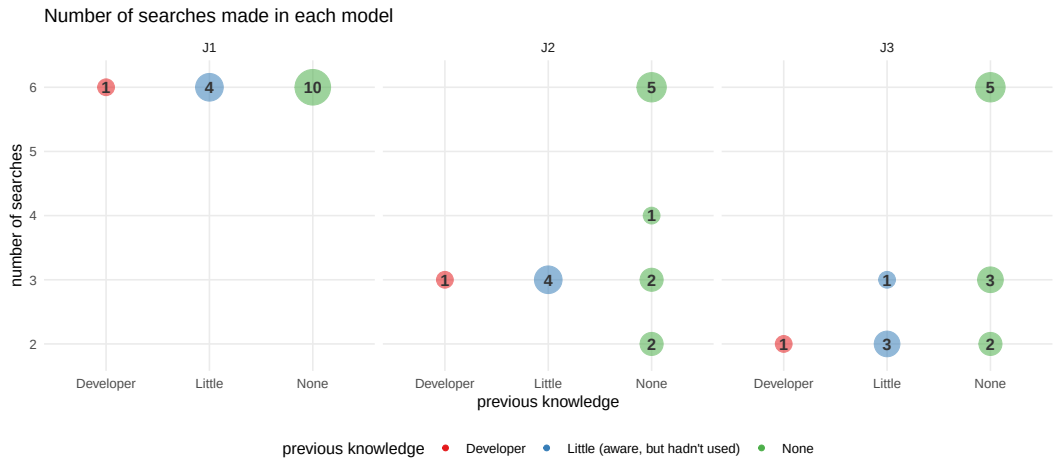



Figure 6.29: Bubble plot of number of searches using each model

We conducted statistical analyses of the differences in the number of searches across models in three different groups: all users, users with little previous knowledge of the models, and users with no previous knowledge of the models, as described in the next subsections.

### 6.5.1 Was there a significant difference between the number of searches across models, considering all users?

Table 6.3 shows the median number of searches and corresponding interquartile range using each model, in the group of all users. We note that model J1 always requires six searches to complete the task.

Table 6.3: Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering all users?

| model | median | IQR |
|---|---|---|
| J1 | 6 | 0 |

PUC-Rio - Certificação Digital Nº 1721764/CA

| model | median | IQR |
|---|---|---|
| J2 | 3 | 3 |
| J3 | 3 | 4 |

We ran a Kruskal-Wallis rank sum test, which showed a **significant difference** at $\alpha = 0.05$ level ($\chi^2 = 14.600$, df=3, p-value=0.0022). We therefore ran a Conover-Iman post-hoc test, with Bonferroni correction. As Table 6.4 shows, there was a significant difference in the number of searches between J1-J2 and J1-J3, in the group of all users.

Table 6.4: Conover-Iman post-test results with Bonferroni correction - Was there a significant difference between the number of searches across models, considering all users?

| model pair | p-value | significance |
|---|---|---|
| J1 - J2 | 0.0005360 | * |
| J1 - J3 | 0.0000386 | * |
| J2 - J3 | 0.6068743 | |

PUC-Rio - Certificação Digital Nº 1721764/CA

#### 6.5.1.1 Was there a significant difference between the number of searches across models, considering users with little knowledge of the models?

Table 6.5 shows the median number of searches and corresponding interquartile range using each model, in the group of users with little previous knowledge of the models.

Table 6.5: Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering users with little knowledge of the models?

| model | median | IQR |
|---|---|---|
| J1 | 6 | 0.00 |
| J2 | 3 | 0.00 |
| J3 | 2 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed a **significant difference** at $\alpha = 0.05$ level ($\chi^2 = 9.900$, df=2, p-value=0.0071). We therefore ran a Conover-Iman post-hoc test, with Bonferroni correction. As Table 6.6 shows, there was a significant difference in the number of searches between

J1-J2, J1-J3, and J2-J3, in the group of users with little previous knowledge of the models.

Table 6.6: Conover-Iman post-test results with Bonferroni correction - Is there a significant difference between the number of searches across models, considering users with little knowledge of the models?

| model pair | p-value | significance |
|---|---|---|
| J1 - J2 | 0.0005620 | * |
| J1 - J3 | 0.0000108 | * |
| J2 - J3 | 0.0076816 | * |

#### 6.5.1.2
#### Was there a significant difference between the number of searches across models, considering users with NO knowledge of the models?

Table 6.5 shows the median number of searches and corresponding interquartile range using each model, in the group of users with no previous knowledge of the models.

Table 6.7: Median and interquartile range of scores per model - Was there a significant difference between the number of searches across models, considering users with NO knowledge of the models?

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0 |
| J2 | 5.0 | 3 |
| J3 | 4.5 | 3 |

We ran a Kruskal-Wallis rank sum test, which showed no significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.872$, df=3, p-value=0.1180).

## 6.6
## Users' Overall Preferences

At the end of the experiment, we asked participants to answer two questions: "Which model better supported you in the task?" and "Which model hampered you more during the task?". Based on their answers, we ranked their preferences regarding the model, summarized in Table 6.8.

The table associates the participants and their preferences with the corresponding experimental group (which defines the order in which participants were exposed to each model) and their previous knowledge about the J3 model.

PUC-Rio - Certificação Digital Nº 1721764/CA

In Table 6.8, 'Favorite' indicates the model that a participant said that better supported them in the task; and 'Least favorite' indicates the model that the participant said that hampered them more.

| | | | Model preference | | |
|---|---|---|---|---|---|
| Participant | Group | Previous knowledge | J1 | J2 | J3 |
| P01 | A (J1, J2, J3) | None | Neutral | Least Favorite | **Favorite** |
| P02 | A (J1, J2, J3) | Little | Least Favorite | Neutral | **Favorite** |
| P03 | A (J1, J2, J3) | Little | Least Favorite | Neutral | **Favorite** |
| P04 | A (J1, J2, J3) | None | Neutral | Least Favorite | **Favorite** |
| P05 | A (J1, J2, J3) | Little | Least Favorite | Neutral | **Favorite** |
| P06 | B (J2, J1, J3) | Little | Least Favorite | Neutral | **Favorite** |
| P07 | B (J2, J1, J3) | None | Least Favorite | Neutral | **Favorite** |
| P08 | B (J2, J1, J3) | None | Neutral | Least Favorite | **Favorite** |
| P10 | B (J2, J1, J3) | None | Neutral | Least Favorite | **Favorite** |
| P11 | B (J2, J1, J3) | Developer | Least Favorite | Neutral | **Favorite** |
| P13 | C (J3, J2, J1) | None | Neutral | Least Favorite | **Favorite** |
| P09 | C (J3, J2, J1) | None | Neutral | **Favorite** | Least Favorite |
| P12 | C (J3, J2, J1) | None | **Favorite** | Neutral | Least Favorite |
| P14 | C (J3, J2, J1) | None | Least Favorite | **Favorite** | Neutral |
| P15 | C (J3, J2, J1) | None | Least Favorite | Neutral | Neutral |

Table 6.8: Study participants' preferences

All participants who used J3 last (after using J1 and J2, in any order), ranked J3 as their favorite search user interface.

In the case of participants in group A, who used the models models in the order J1, J2, J3, three participants had little previous knowledge of the J3 model. In group B, three participants had no previous knowledge, one participant had little knowledge, and one participant had participated in the development of J3 in an industry project.

Although we used a heterogenous group of participants between experiments across groups, we failed to ensure that the groups had similar composition in terms of the participants' previous knowledge of the models.

All participants in group C, who used the models in the order J3, J2, and J1, had no knowledge of the proposed model (J3). Three of them did not notice or did not pay much attention to the related questions (in J2) or related visualizations (in J3), and therefore used J2 and J3 as if they were identical to J1. In the interviews, they stated that they were "too focused on the task" to notice the additional features. This may indicate that models J2 and J3 are not very suitable for non-exploratory or direct search tasks. Moreover, the more complex user interfaces in J2 and J3, without having a perceived benefit, may have also influenced those users not to rate J3 highly.

These results indicate that some familiarity with the model may be necessary to use it efficiently, and that it may be useful to, through the user interface itself, call users' attentions to the added features. In this sense, the familiarity with J1 provides an advantage over the other models.

PUC-Rio - Certificação Digital Nº 1721764/CA

## 6.7 Study Limitations

Unfortunately, a few days before the experiments, the API we used as a service to our work became faulty and behaved erratically. During the pilot test, we identified that too often the user interactions with the system would cause the server to restart or not to load data correctly. To build a workaround solution significant changes in the performance of the interface were needed, slowing down the chart load in several seconds in order to ensure that the data presented to all the participants would be consistent. This prevented us from analyzing the time on task.

The faulty service included the verbalization component of the API, responsible for creating sentences in natural language that a user could understand. This meant that questions that should have been presented as "What were the highest rated movies from last year?" had to be shown to the user as phrases that lacked grammar, cohesion and coherence, such as "This is high rated movies in 2018?". Such changes deeply affected J3, as there was not time to create a workaround for the experiment.

We also had to limit the user interaction with the models when using the search bar, because the use of certain characters would cause the service to malfunction. Therefore, we presented the task queries in a file, and asked participants to copy and paste them to the search field as needed.

The lack of statistically significant differences in the questionnaire results may be due to the small sample of participants (15), so further studies should be conducted with a larger sample of participants to check whether the results would be reproduced.

PUC-Rio - Certificação Digital Nº 1721764/CA

## 6.8 Discussion

Although most of the results presented in this work are not statistically significant, it is essential to note that the evaluation of /myTool (model J3) did not perform much worse than the other models and, in most cases, it received a better evaluation from the user than models J1 and J3. These results are an indication that even though /myTool it is significantly more complex and less familiar than the other models, from the user perspective it is a potential solution for the design of search result pages that enhance the user experience when doing an exploratory search.

The results on Section **??** imply that between the traditional search behaviour of model J1 and the exploratory based of models J2 and J3 there is a significant difference. These differences imply that since the results from

the TAM and NASA TLX were mostly in pair with the other models, the alternative models are possible useful solutions and should be explored more in-depth in future works.

Based on our results, we believe a possible alternative solution for the design of search results pages such as the one we proposed in this study may be a hybrid of from the models evaluated in this research. This hybrid model can be displayed in at least two forms of interface design:

a) A new model that shows the related question as a preview and only unfolds the visualization when the user explicitly ask.

b) An adaptative interface model that increases the personalization of search results pages by showing or hiding the related answers to the user.

An intermediate interface such as a) could be designed, developed and evaluated in the additional studies proposed on Section 7.2.1. However, building an interface to improve user interaction by building a user model based on the partial knowledge of that user is a much more complex issue. To build an interface that is genuinely adaptive is not enough to memorize the interaction of the user with the interface. Instead, improvement should emerge from a generalization over previous interactions and carry over to new ones.

PUC-Rio - Certificação Digital Nº 1721764/CA

# 7
# Conclusions

This chapter describes the lessons learned in this dissertation. Section 7.1 presents an overview of the main contributions of this research work and Section 7.2 describes issues that are closely related to this work but fell outside the scope of this dissertation.

## 7.1
## Contributions

The main contribution of this dissertation is a model to amplify cognition for search tasks. The model involves generating and presenting related queries to expand the search space and progressively disclosing the corresponding results.

We conducted an evaluation of the proposed model (J3) in comparison with two distinct search user interface models for data visualization: a *Traditional* SUI Model (J1), inspired by the work of Wilson (1999), and a *Suggested-links* SUI (J2), which combines J1 with a suggested list of related questions.

Despite the limitations of the experiment, the outcomes of the analysis suggest that the model proposed with JARVIS may be a promising path for the design of new SUIs. The results from the Task Load Index and Technology Acceptance Model Questionnaires showed that, although J3 presents a more complex user interface and more features, it did not perform worse than the other models evaluated by the questionnaires. Moreover, even in the simple experiment we conducted, the number of searches made with J3 was significantly lower than the number of searches made with the Traditional SUI (J1), as shown in Section 6.6.

## 7.2
## Future work

This section describes some future work to advance this research.

PUC-Rio - Certificação Digital Nº 1721764/CA

### 7.2.1 Additional Studies

After the service in which our work depends is fixed, we plan to conduct additional studies.

As the statistical power of the study was low because of the small number of participants, especially when segmented into groups according to their previous knowledge, we plan to conduct additional studies with a larger number of participants, to attempt to reproduce the results or obtain more robust evidence about the various quality aspects of our proposed solution. In those studies, we plan to adequately measure the efficiency of each model, based not on the number of searches as a proxy, on time on task.

As we have obtained some preliminary indication that the degree of previous knowledge may affect the results, we also plan a more controlled experiment, in which we will provide different levels of introduction to each model, and allow for different periods of time for exploring the user interfaces before diving into the tasks.

PUC-Rio - Certificação Digital Nº 1721764/CA

### 7.2.2 Progressive Query Expansion

To reduce the user cognitive overload when making a query, JARVIS could be enhanced by a mechanism whereby the user could, as they enter the search query, obtain terms related to the what they are typing and be able to progressively expand the query. We believe such a mechanism may be interesting, especially because JARVIS works on a database described by an ontology. The structure of the ontology could help users optimize the query expansion and turn the proposal into a more effective search system.

### 7.2.3 User Feedback - Implicit or Explicit

A challenging aspect when evaluating a search result page like the one we propose with JARVIS involves taking into consideration the ranking algorithm that recommends visualizations for related questions. One way to gather those pieces of information would be to collect explicit feedback from the user. This feedback can be materialized in the interface in numerous ways. For example, the system may ask for the user to evaluate the visualizations that better answer their query through a like/dislike button. Another way to collect those data is by selecting keywords in the search bar or even presenting quizzes to determine the user's preferences and clusters. Those methods, unfortunately, distract the user from their normal search behaviour (Kelly and Teevan, 2003). To mitigate

those issues, we could adopt implicit feedback techniques in order to expand the user's query and to build a user profiling mechanism that suggests tailored related questions.

PUC-Rio - Certificação Digital Nº 1721764/CA

# Bibliography

N.J. Belkin, R.N. Oddy, and H.M. Brooks. Ask for information retrieval: Part i. background and theory. *Journal of Documentation*, 38(2):61–71, 1982. doi: 10.1108/eb026722.

Diego Calvanese, Benjamin Cogrel, Sarah Komla-Ebri, Roman Kontchakov, Davide Lanti, Martin Rezk, Mariano Rodriguez-Muro, and Guohui Xiao. Ontop: Answering sparql queries over relational databases. *Semantic Web*, 8(3):471–487, 2017.

Hao Chen and Susan Dumais. Bringing order to the web: Automatically categorizing search results. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 145–152. ACM, 2000.

G. Chowdhury and S. Chowdhury. *Introduction to digital libraries*. Facet Publishing, UK, 2002. URL https://strathprints.strath.ac.uk/2603/.

Fred D. Davis. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319–340, 1989. ISSN 0276-7783. doi: 10.2307/249008.

Offer Drori. How to display search results in digital libraries - user study. In *New Technologies for Information Systems, Proceedings of the 3rd International Workshop on New Developments in Digital Libraries, NDDL 2003, and the 1st International Workshop on Validation and Verification of Software for Enterprise Information Systems, VVEIS 2003, In conjunction with ICEIS 2003, Angers, France, April 2003*, pages 13–28. ICEIS Press, 2003.

Susan Dumais, Edward Cutrell, and Hao Chen. Optimizing search by showing results in context. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 277–284. ACM, 2001.

Deborah Fallows. Search Engine Use | Pew Research Center, August 2008. URL https://www.pewinternet.org/2008/08/06/search-engine-use/.

Tong Gao, Mira Dontcheva, Eytan Adar, Zhicheng Liu, and Karrie G. Karahalios. Datatone: Managing ambiguity in natural language interfaces for data visualization. In *Proceedings of the 28th Annual ACM Symposium on User Interface*

PUC-Rio - Certificação Digital Nº 1721764/CA

*Software & Technology*, UIST '15, pages 489–500, New York, NY, USA, 2015. ACM Press. ISBN 978-1-4503-3779-3. doi: 10.1145/2807442.2807478.

Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, June 1993. ISSN 1042-8143. doi: 10.1006/knac.1993.1008.

Sandra G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, October 2006. ISSN 1541-9312. doi: 10.1177/154193120605000909.

Marti A. Hearst. *User Interfaces and Visualization*, volume 2. Addison-Wesley, 1999. URL https://ci.nii.ac.jp/naid/10022005353/en/.

Marti A. Hearst. *Search User Interfaces*. Cambridge University Press, New York, NY, USA, 1st edition, 2009. ISBN 0521113792, 9780521113793.

Tomi Heimonen. *Design and evaluation of user interfaces for mobile web search*. Doctoral Dissertation, Tampere University Press, 2012.

Hideo Joho and Joemon M. Jose. A comparative study of the effectiveness of search result presentation on the web. In Mounia Lalmas, Andy MacFarlane, Stefan Rüger, Anastasios Tombros, Theodora Tsikrika, and Alexei Yavlinsky, editors, *Advances in Information Retrieval*, pages 302–313, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-33348-7.

Diane Kelly and Jaime Teevan. Implicit feedback for inferring user preference: a bibliography. In *SIGIR forum*, volume 37(2), pages 18–28, 2003.

Gary Marchionini. Information-seeking strategies of novices using a full-text electronic encyclopedia. *J. Am. Soc. Inf. Sci.*, 40(1):54–66, January 1989. ISSN 0002-8231. doi: 10.1002/(SICI)1097-4571(198901)40:1<54::AID-ASI6>3.0.CO;2-R.

Gary Marchionini. *Information Seeking in Electronic Environments*. Cambridge Series on Human-Computer Interaction. Cambridge University Press, 1997. ISBN 9780521586740.

Gary Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41–46, April 2006. ISSN 0001-0782. doi: 10.1145/1121949.1121979.

Vicki L. O'Day and Robin Jeffries. Orienteering in an information landscape: How information seekers get from here to there. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93,

PUC-Rio - Certificação Digital Nº 1721764/CA

pages 438–445, New York, NY, USA, 1993. ACM. ISBN 0-89791-575-5. doi: 10.1145/169059.169365.

Marc L. Resnick, Carlos Maldonado, Juan Martin Santos, and Rebeca Lergier. Modeling on-line search behavior using alternative output structures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45(15):1166–1170, 2001. doi: 10.1177/154193120104501503.

Vidya Setlur, Sarah E Battersby, Melanie Tory, Rich Gossweiler, and Angel X Chang. Eviza: A natural language interface for visual analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 365–377. ACM, 2016.

Yiwen Sun, Jason Leigh, Andrew Johnson, and Sangyoon Lee. Articulate: A semi-automated model for translating natural language queries into meaningful visualizations. In Robyn Taylor, Pierre Boulanger, Antonio Krüger, and Patrick Olivier, editors, *Smart Graphics*, pages 184–195, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-13544-6.

Paul Thompson. Finding out about: A cognitive perspective on search engine technology and the www. *Information Retrieval*, 5(2):269–271, Apr 2002. ISSN 1573-7659. doi: 10.1023/A:1015706230423.

Ryen W. White, Bill Kules, and Ben Bederson. Exploratory search interfaces: Categorization, clustering and beyond: Report on the xsi 2005 workshop at the human-computer interaction laboratory, university of maryland. *SIGIR Forum*, 39(2):52–56, December 2005. ISSN 0163-5840. doi: 10.1145/1113343.1113356.

Max L Wilson, Bill Kules, Ben Shneiderman, et al. From keyword search to exploration: Designing future search interfaces for the web. *Foundations and Trends® in Web Science*, 2(1):1–97, 2010.

T.D. Wilson. Models in information behaviour research. *Journal of Documentation*, 55(3):249–270, 1999. doi: 10.1108/EUM0000000007145.

PUC-Rio - Certificação Digital Nº 1721764/CA

# A
# Experiment Tasks

Experimento A - Avaliação de Interface de Busca

Esse experimento visa a avaliação de diferentes modelos de interface de busca.

Na seguinte de modelos ordem você deve executar responder 6 tarefas de visualização.

Ordem das Ferramentas:

Jarvis 1 - Jarvis 2 - Jarvis 3

Tarefas de visualização:

Quais os 5 filmes que mais receberam premiações no último ano?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |

Quais os 5 séries de TV que mais receberam premiações no último ano?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |

Quais as atrizes que mais receberam Oscars no último ano?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |

PUC-Rio - Certificação Digital Nº 1721764/CA

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |

Quais os 5 filmes mais bem avaliados no IMDB em 2001?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |

Quais os 5 Série de TV mais bem avaliados no IMDB em 2001?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |

Quais as 5 Atriz mais bem avaliados no IMDB em 2001?

| JARVIS 1 | JARVIS 2 | JARVIS 3 |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |

PUC-Rio - Certificação Digital Nº 1721764/CA

# B
# Informed Consent Form

Pontifícia Universidade Católica
do Rio de Janeiro

## Termo de Consentimento

O objetivo deste trabalho é **comparar o uso de 3 modelos de interface para busca visuais**. Você foi indicado para nos auxiliar nesse experimento de avaliação de desempenho dessas ferramentas.

Solicitamos seu consentimento para a sua participação e gravação em áudio da entrevista, fornecendo algumas informações adicionais importantes para embasar sua decisão:

1. Os dados coletados durante a entrevista destinam-se **estritamente** a atividades de pesquisa e avaliação desta proposta.
2. A utilização dos resultados desta entrevista pauta-se no **respeito à privacidade**, portanto o **anonimato** dos participantes é preservado em quaisquer documentos que elaborarmos.
3. O consentimento para participação é uma escolha livre, feita mediante a prestação de todos os esclarecimentos necessários sobre a entrevista. **A participação pode ser interrompida a qualquer momento**.
4. O pesquisador Dalai Ribeiro encontra-se disponível para contato pelo e-mail **dribeiro@inf.puc-rio.br**.

De posse das informações apresentadas, gostaríamos que você se pronunciasse acerca da sua participação.

( ) Dou meu consentimento para a realização desta entrevista e sua gravação.

( ) Não autorizo a realização desta entrevista.

Rio de Janeiro, ___ de _____ de 2019

| **Participante** | **Pesquisador** |
|---|---|
| Nome: | Nome: |
| _____ | *Dalai Ribeiro* |
| Assinatura: | Assinatura: |
| _____ | _____ |

PUC-Rio - Certificação Digital Nº 1721764/CA

# C
# Questionnaire



## SAMPLE - JARVIS Questionnaire

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

* Required

1. **Name** *

2. **Gender** *
*Mark only one oval.*
- ◯ Female
- ◯ Male
- ◯ Non-Binary
- ◯ Prefer not say

3. **Educational degree** *
*Mark only one oval.*
- ◯ Undergraduate
- ◯ Graduate
- ◯ Postgraduate

4. **Field of Work**

5. **Age group** *
*Mark only one oval.*
- ◯ 18 - 24 years old
- ◯ 25 - 44 years old
- ◯ 44 - 60 years old
- ◯ more than 60 years old

## NASA Task Load Index

6. **Mental Demand - How mentally demanding was the task?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Very Low | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very High |



PUC-Rio - Certificação Digital Nº 1721764/CA

7. **Physical Demand - How physically demanding was the task?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Very Low | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very High |

8. **Temporal Demand - How hurried or rushed was the pace of the task?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Very Low | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very High |

9. **Performance - How successful were you in accomplishing what you were asked to do?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Perfect | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Failure |

10. **Effort - How hard did you have to work to accomplish your level of performance?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Very Low | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very High |

11. **Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you?** *
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Very Low | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very High |

## Technology Assessment Model

12. **I find the search model A system easy to use**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

13. **Using the search model A gives me greater control over my search**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |



PUC-Rio - Certificação Digital Nº 1721764/CA

14. **The application enables me to accomplish tasks more quickly**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

15. **Using the search model A enhances my effectiveness on the job**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

16. **Using the search model A makes it easier to do my job**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

17. **Using the search model A makes it easier to do searches**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

18. **Learning how to use an the search model A system is easy for me**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

19. **I find the search model A complicated to use**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

20. **Learning to operate the search model A is easy for me**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |



PUC-Rio - Certificação Digital Nº 1721764/CA

21. **Interacting with the search model A is often frustrating**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

22. **The search model A is rigid and inflexible to interact with**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

23. **It is easy for me to remember how to perform tasks using the search model A**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

24. **Interacting with the search model A requires a lot of mental effort**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

25. **My interaction with the search model A is clear and understandable**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

26. **I find it takes a lot of effort to become skillful at using the search model A**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

27. **Overall, I find the search model A easy to use**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |



PUC-Rio - Certificação Digital Nº 1721764/CA

28. **It is easy to become skillful at using the search model A system**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

29. **I feel confident finding information in the search model A system**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

30. **I have the necessary skills for using the search model A system**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

31. **I have no difficulty accessing and using the search model A system**
*Mark only one oval.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Agree |

Powered by
Google Forms



PUC-Rio - Certificação Digital Nº 1721764/CA

# D
# Questionnaire Results

## D.1
## Hypothesis tests - All users

### D.1.1
### TLX

#### D.1.1.1
#### TLX: 6 Mental Demand - How mentally demanding was the task? - All users

Table D.1: Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - All users

| model | median | IQR |
|---|---|---|
| J1 | 2 | 2 |
| J2 | 1 | 1 |
| J3 | 1 | 1 |

We ran a Kruskal-Wallis rank sum test, which showed a **significant difference** at $\alpha = 0.05$ level ($\chi^2 = 6.217$, df=2, p-value=0.0447). We therefore ran a Conover-Iman post-hoc test, with Bonferroni correction, but it did not find any significant difference.

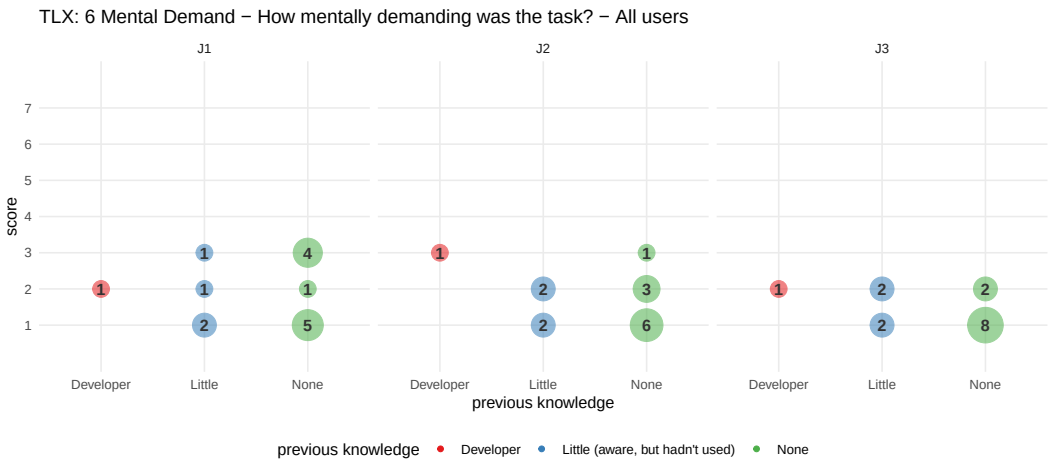



Figure D.1: Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - All users

PUC-Rio - Certificação Digital Nº 1721764/CA

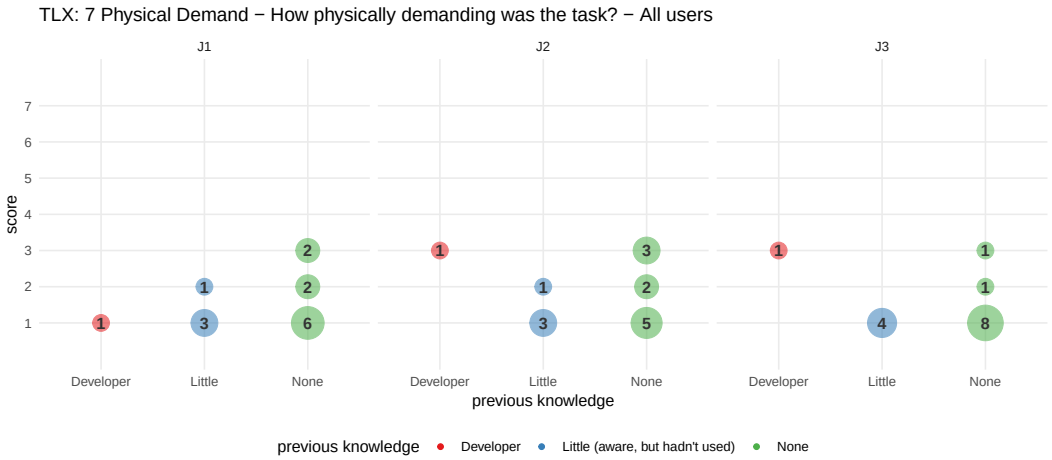



Figure D.2: Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - All users

### D.1.1.2
### TLX: 7 Physical Demand - How physically demanding was the task? - All users

Table D.2: Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 1.0 |
| J2 | 1 | 1.5 |
| J3 | 1 | 0.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.034$, df=2, p-value=0.5964).

PUC-Rio - Certificação Digital Nº 1721764/CA

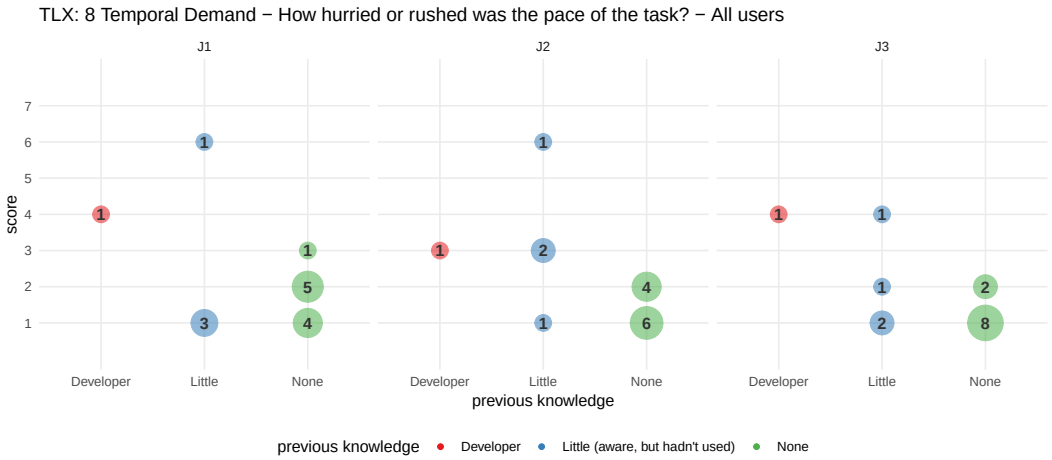



Figure D.3: Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users

### D.1.1.3
### TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users

Table D.3: Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - All users

| model | median | IQR |
|---|---|---|
| J1 | 2 | 1.0 |
| J2 | 2 | 1.5 |
| J3 | 1 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.628$, df=4, p-value=0.6219).

PUC-Rio - Certificação Digital Nº 1721764/CA

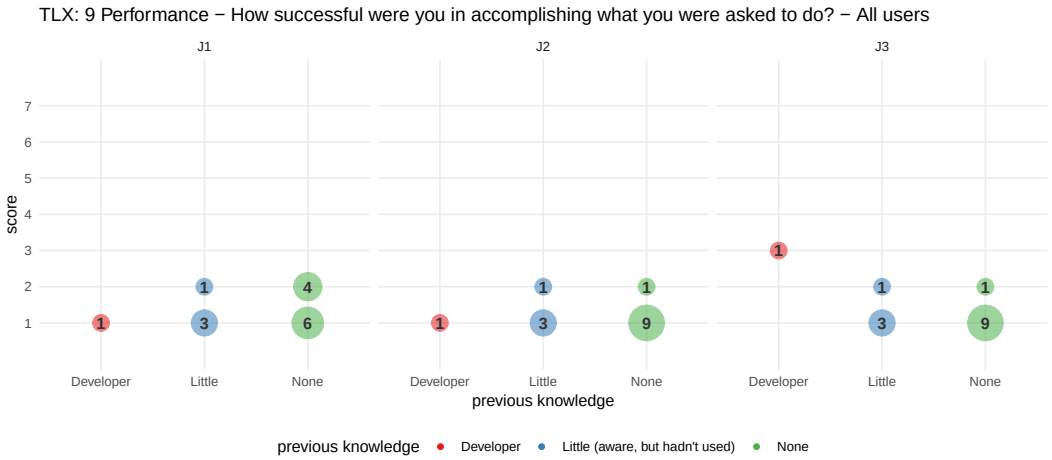



Figure D.4: Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users

### D.1.1.4
### TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users

Table D.4: Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 1 |
| J2 | 1 | 0 |
| J3 | 1 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.101$, df=2, p-value=0.2121).

PUC-Rio - Certificação Digital Nº 1721764/CA

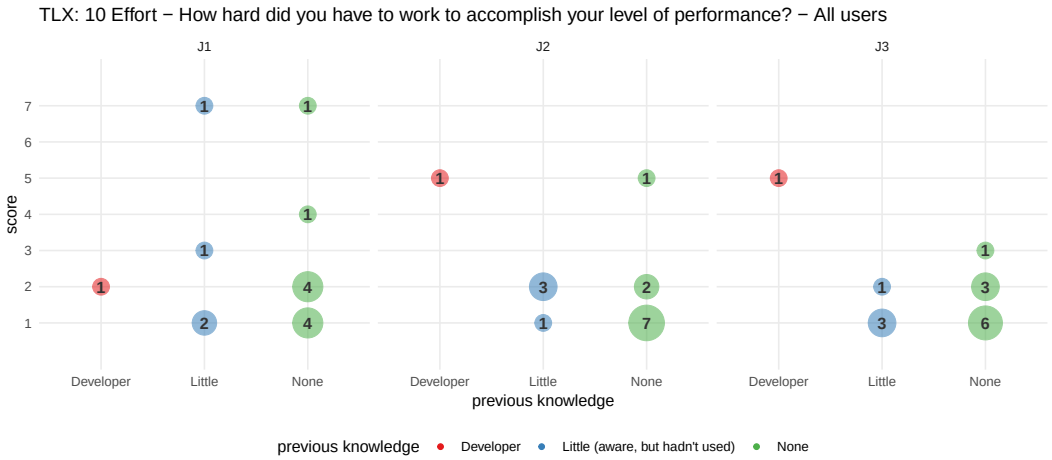



Figure D.5: Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users

### D.1.1.5
### TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users

Table D.5: Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - All users

| model | median | IQR |
|---|---|---|
| J1 | 2 | 1.5 |
| J2 | 1 | 1.0 |
| J3 | 1 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.568$, df=5, p-value=0.3506).

PUC-Rio - Certificação Digital Nº 1721764/CA

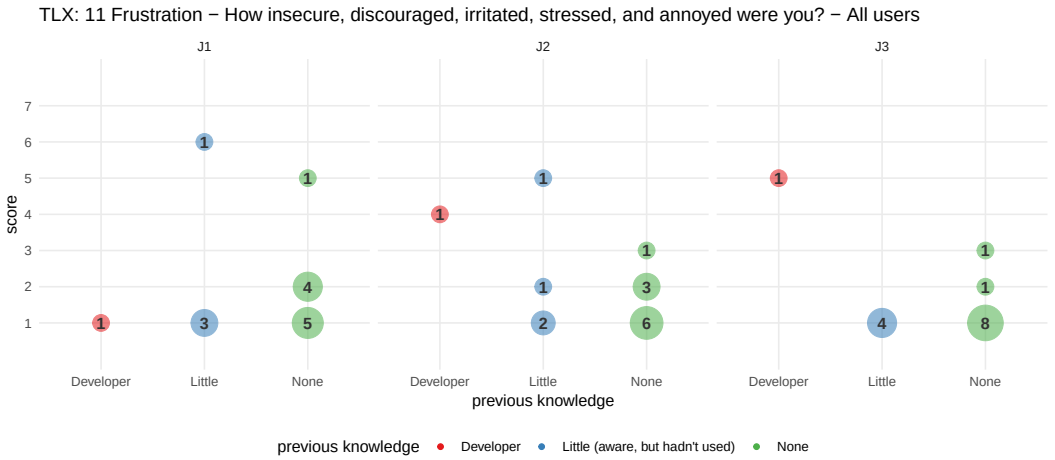



Figure D.6: Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users

### D.1.1.6
### TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users

Table D.6: Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 1 |
| J2 | 1 | 1 |
| J3 | 1 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.122$, df=5, p-value=0.5320).

PUC-Rio - Certificação Digital Nº 1721764/CA

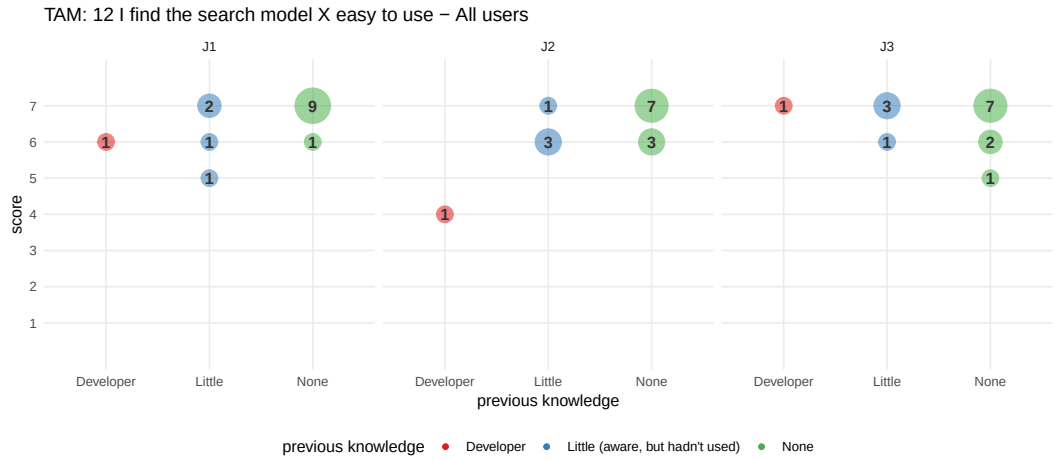



Figure D.7: Response to survey - TAM: 12 I find the search model X easy to use - All users

### D.1.2 TAM

#### D.1.2.1 TAM: 12 I find the search model X easy to use - All users

Table D.7: Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.5 |
| J2 | 7 | 1.0 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.000$, df=3, p-value=1.0000).

PUC-Rio - Certificação Digital Nº 1721764/CA

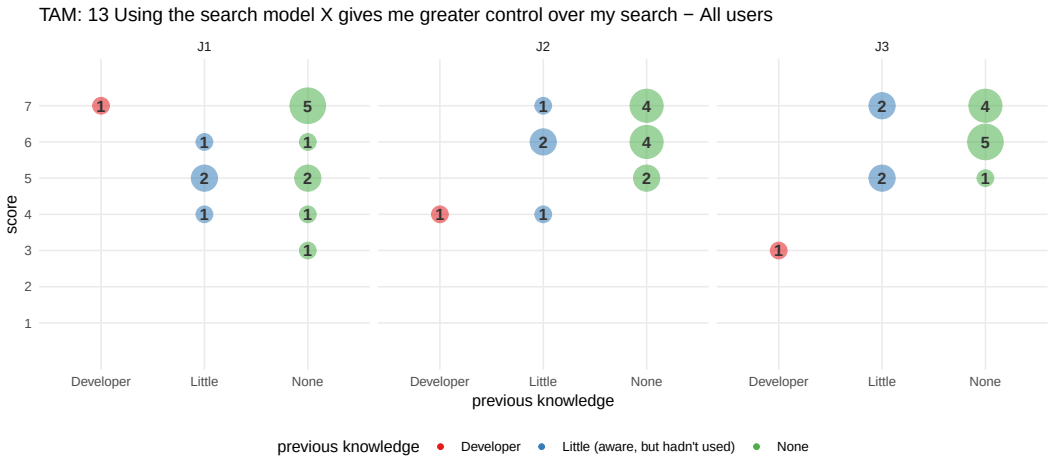



Figure D.8: Response to survey - TAM: 13 Using the search model X gives me greater control over my search - All users

### D.1.2.2
### TAM: 13 Using the search model X gives me greater control over my search - All users

Table D.8: Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - All users

| model | median | IQR |
|---|---|---|
| J1 | 6 | 2.0 |
| J2 | 6 | 1.5 |
| J3 | 6 | 1.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.645$, df=4, p-value=0.6189).

PUC-Rio - Certificação Digital Nº 1721764/CA

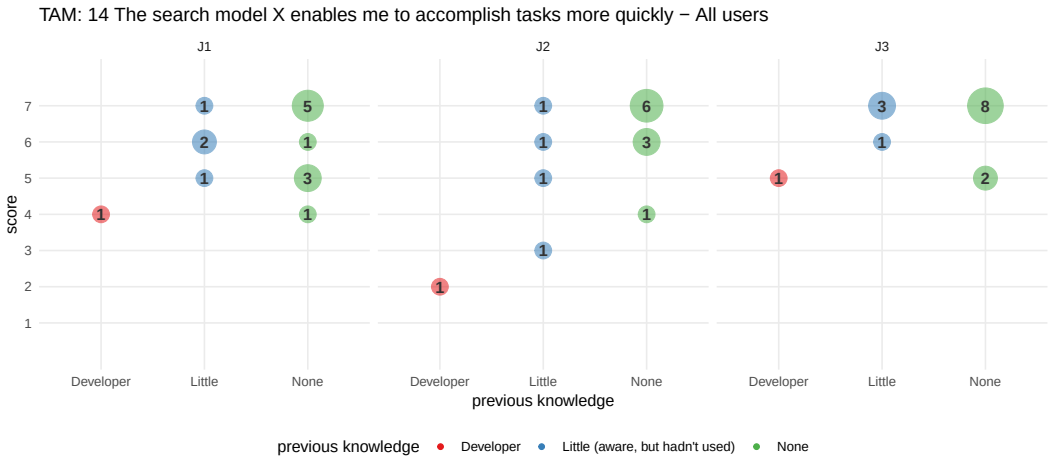



Figure D.9: Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - All users

### D.1.2.3
### TAM: 14 The search model X enables me to accomplish tasks more quickly - All users

Table D.9: Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - All users

| model | median | IQR |
|---|---|---|
| J1 | 6 | 2.0 |
| J2 | 6 | 1.5 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.400$, df=5, p-value=0.4934).

PUC-Rio - Certificação Digital Nº 1721764/CA

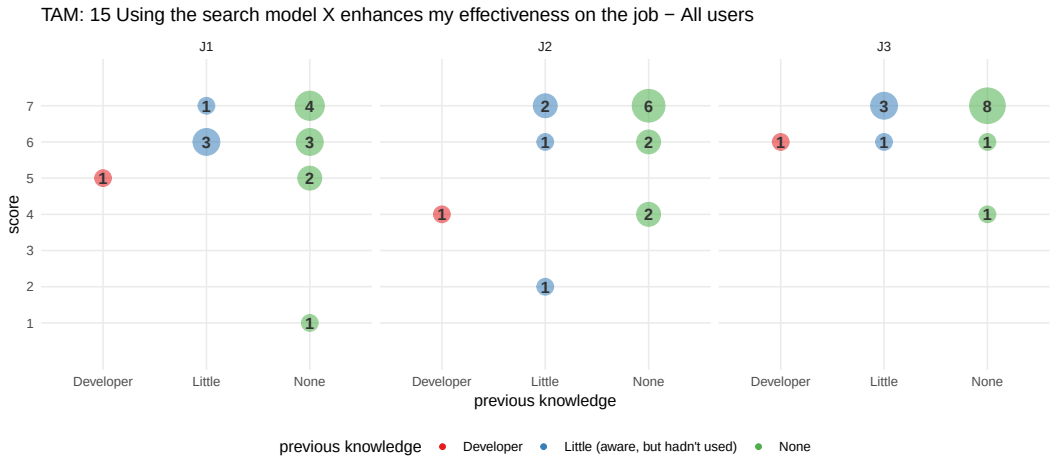



Figure D.10: Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - All users

### D.1.2.4
### TAM: 15 Using the search model X enhances my effectiveness on the job - All users

Table D.10: Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - All users

| model | median | IQR |
|---|---|---|
| J1 | 6 | 1.5 |
| J2 | 7 | 2.0 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 9.533$, df=5, p-value=0.0896).

PUC-Rio - Certificação Digital Nº 1721764/CA

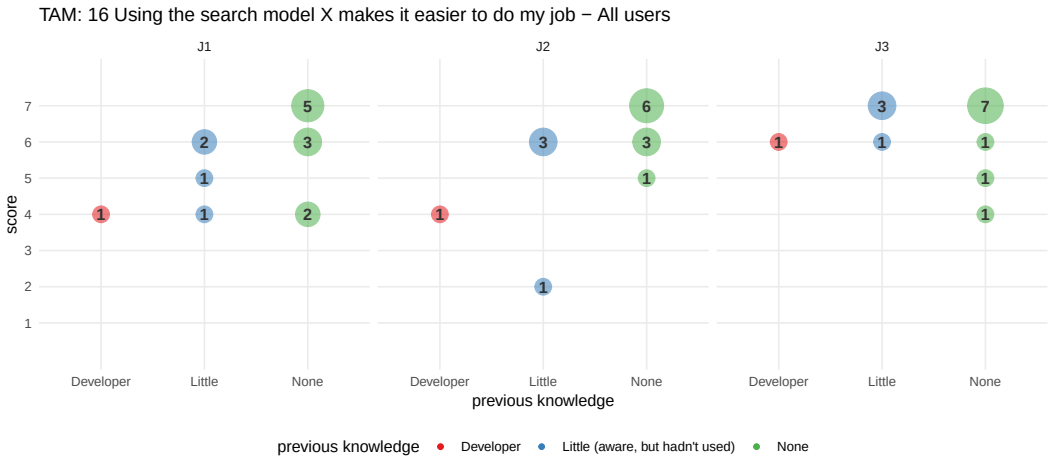



Figure D.11: Response to survey - TAM: 16 Using the search model X makes it easier to do my job - All users

### D.1.2.5
### TAM: 16 Using the search model X makes it easier to do my job - All users

Table D.11: Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - All users

| model | median | IQR |
|---|---|---|
| J1 | 6 | 2.5 |
| J2 | 6 | 1.0 |
| J3 | 7 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.365$, df=4, p-value=0.3588).

PUC-Rio - Certificação Digital Nº 1721764/CA

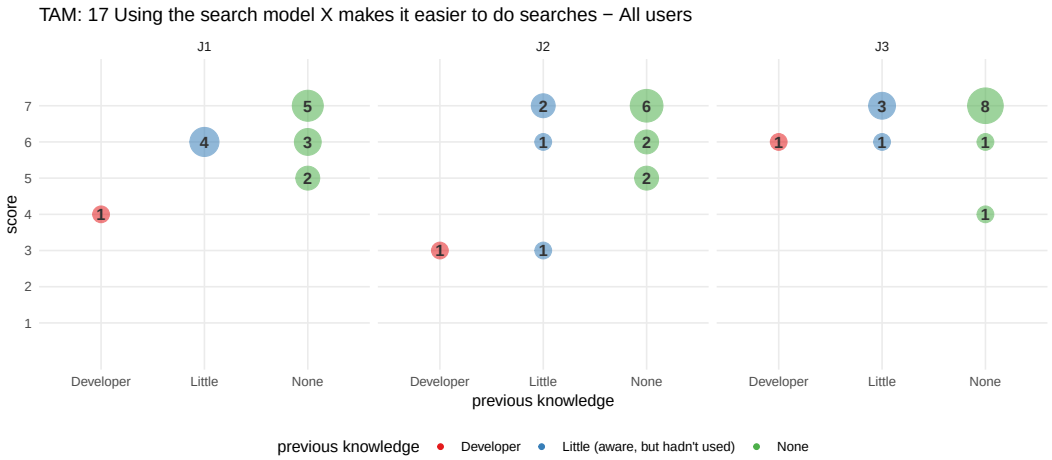



Figure D.12: Response to survey - TAM: 17 Using the search model X makes it easier to do searches - All users

### D.1.2.6
### TAM: 17 Using the search model X makes it easier to do searches - All users

Table D.12: Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - All users

| model | median | IQR |
|---|---|---|
| J1 | 6 | 1.0 |
| J2 | 7 | 1.5 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.472$, df=4, p-value=0.2422).

PUC-Rio - Certificação Digital Nº 1721764/CA

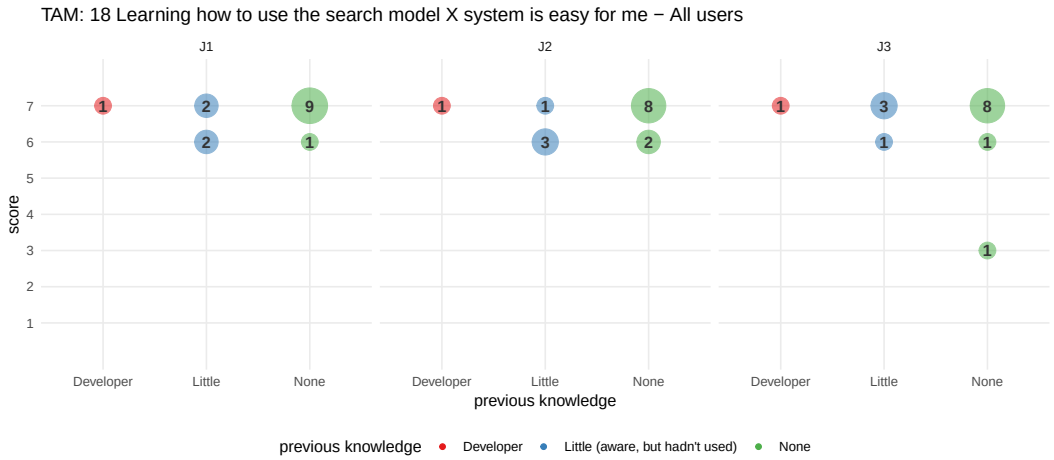



Figure D.13: Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - All users

### D.1.2.7
### TAM: 18 Learning how to use the search model X system is easy for me - All users

Table D.13: Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0 |
| J2 | 7 | 1 |
| J3 | 7 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.613$, df=2, p-value=0.4463).

PUC-Rio - Certificação Digital Nº 1721764/CA

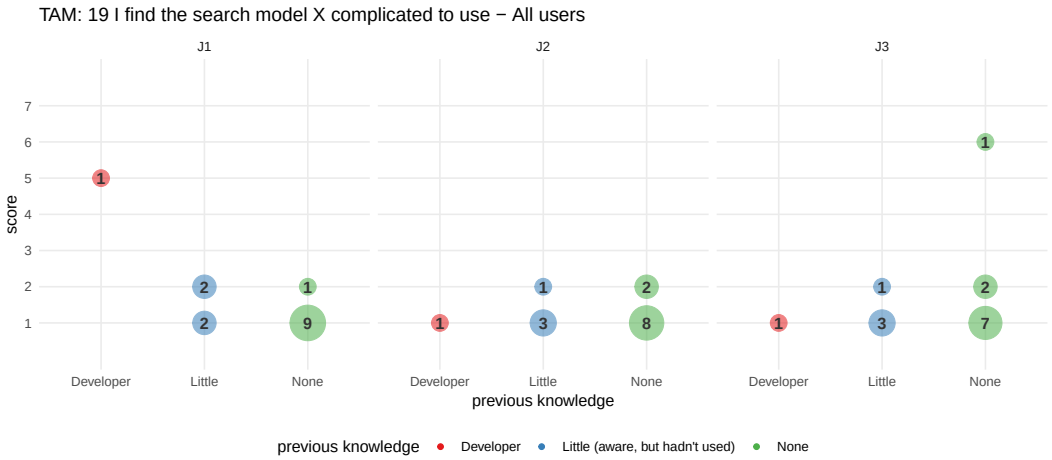



Figure D.14: Response to survey - TAM: 19 I find the search model X complicated to use - All users

### D.1.2.8
### TAM: 19 I find the search model X complicated to use - All users

Table D.14: Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.5 |
| J2 | 1 | 0.0 |
| J3 | 1 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.933$, df=3, p-value=0.4020).

PUC-Rio - Certificação Digital Nº 1721764/CA

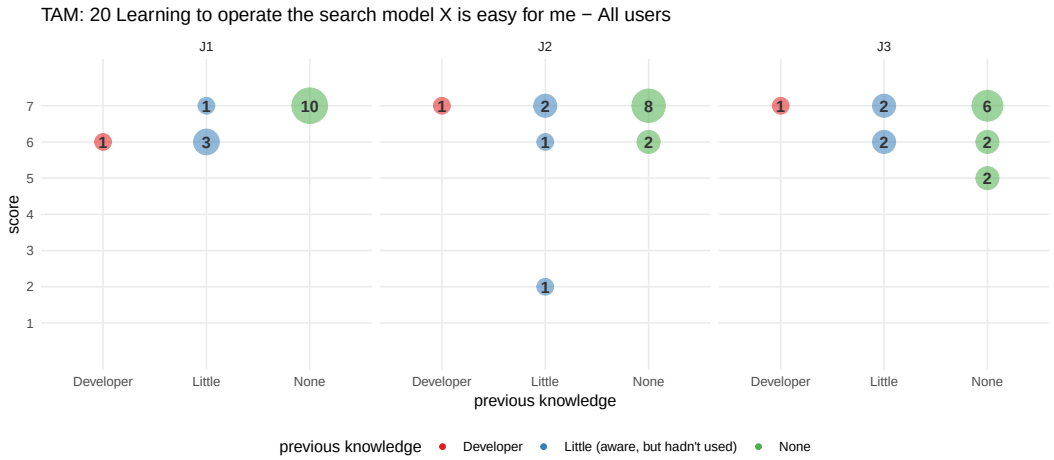



Figure D.15: Response to survey - TAM: 20 Learning to operate the search model X is easy for me - All users

### D.1.2.9
### TAM: 20 Learning to operate the search model X is easy for me - All users

Table D.15: Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.5 |
| J2 | 7 | 0.5 |
| J3 | 7 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.123$, df=3, p-value=0.3731).

PUC-Rio - Certificação Digital Nº 1721764/CA

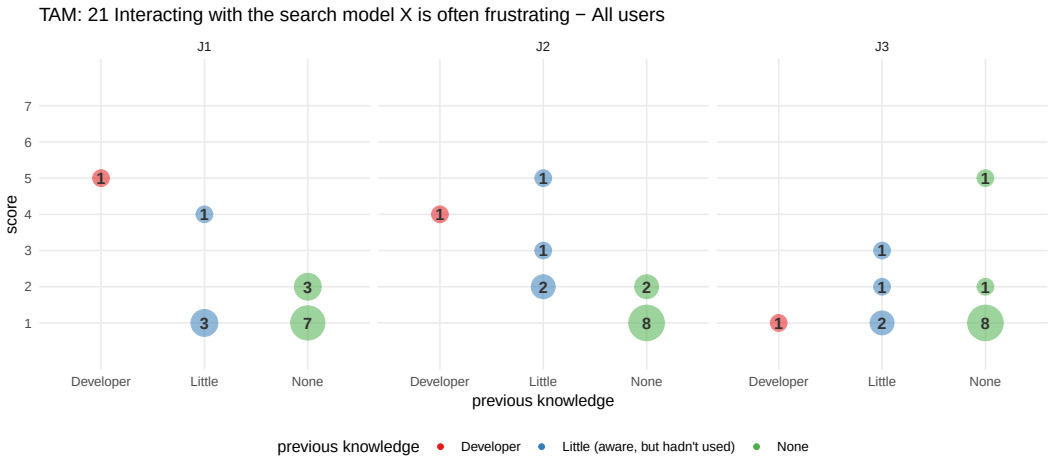



Figure D.16: Response to survey - TAM: 21 Interacting with the search model X is often frustrating - All users

### D.1.2.10
### TAM: 21 Interacting with the search model X is often frustrating - All users

Table D.16: Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 1.0 |
| J2 | 1 | 1.0 |
| J3 | 1 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.680$, df=4, p-value=0.7943).

PUC-Rio - Certificação Digital Nº 1721764/CA

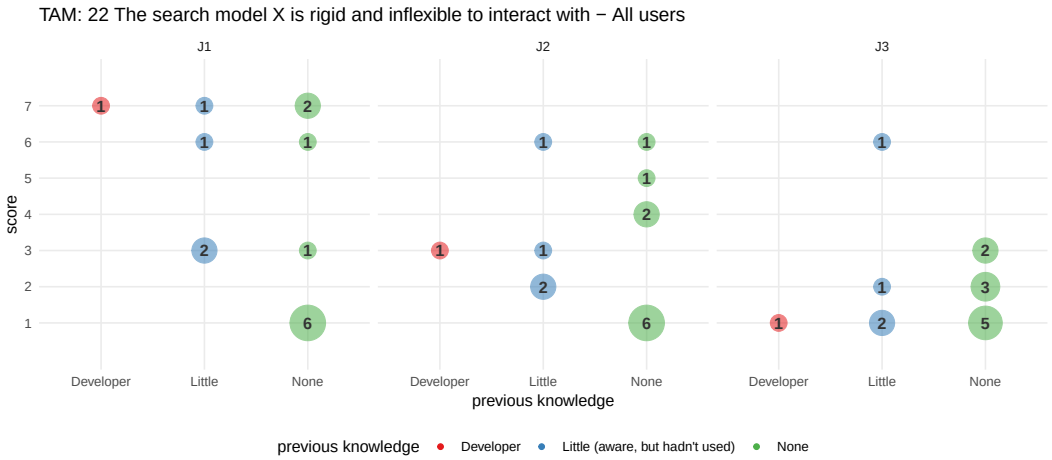



Figure D.17: Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - All users

### D.1.2.11
### TAM: 22 The search model X is rigid and inflexible to interact with - All users

Table D.17: Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - All users

| model | median | IQR |
|---|---|---|
| J1 | 3 | 5.5 |
| J2 | 2 | 3.0 |
| J3 | 1 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 10.574$, df=6, p-value=0.1025).

PUC-Rio - Certificação Digital Nº 1721764/CA

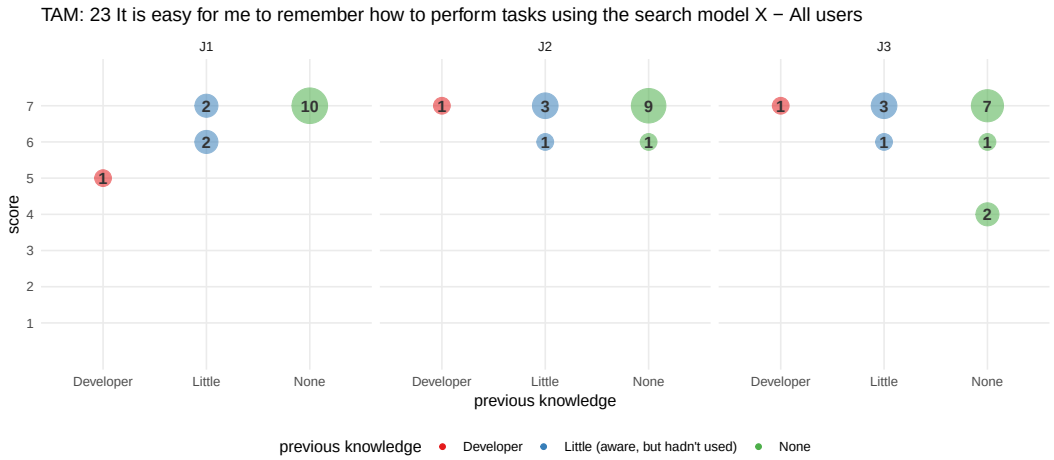



Figure D.18: Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users

### D.1.2.12
### TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users

Table D.18: Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.0 |
| J2 | 7 | 0.0 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.441$, df=3, p-value=0.2176).

PUC-Rio - Certificação Digital Nº 1721764/CA

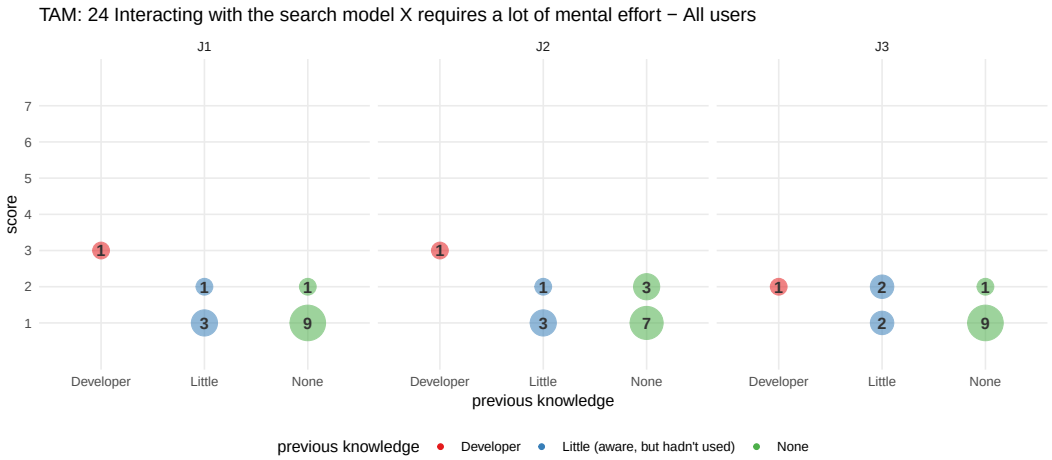



Figure D.19: Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - All users

### D.1.2.13
### TAM: 24 Interacting with the search model X requires a lot of mental effort - All users

Table D.19: Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.0 |
| J2 | 1 | 1.0 |
| J3 | 1 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.364$, df=2, p-value=0.5055).

PUC-Rio - Certificação Digital Nº 1721764/CA

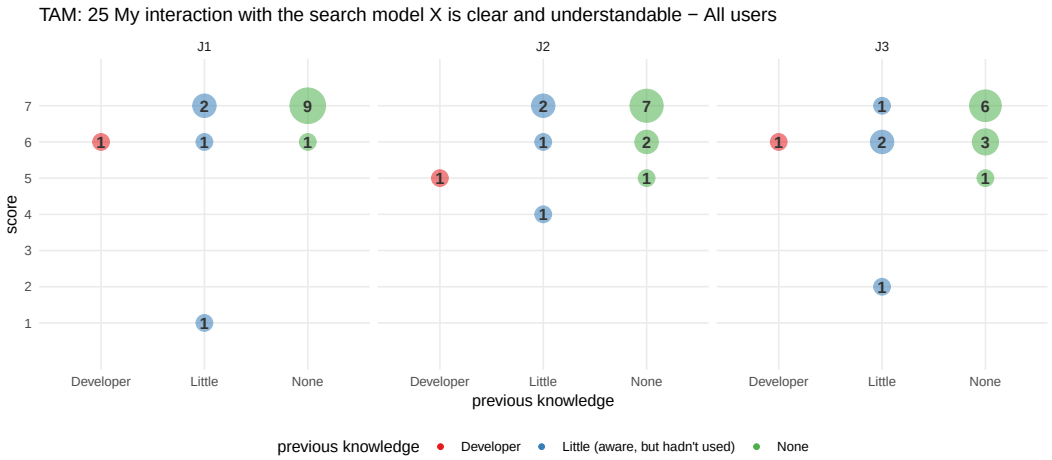



Figure D.20: Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - All users

### D.1.2.14
### TAM: 25 My interaction with the search model X is clear and understandable - All users

Table D.20: Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.5 |
| J2 | 7 | 1.0 |
| J3 | 6 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.391$, df=5, p-value=0.3700).

PUC-Rio - Certificação Digital Nº 1721764/CA

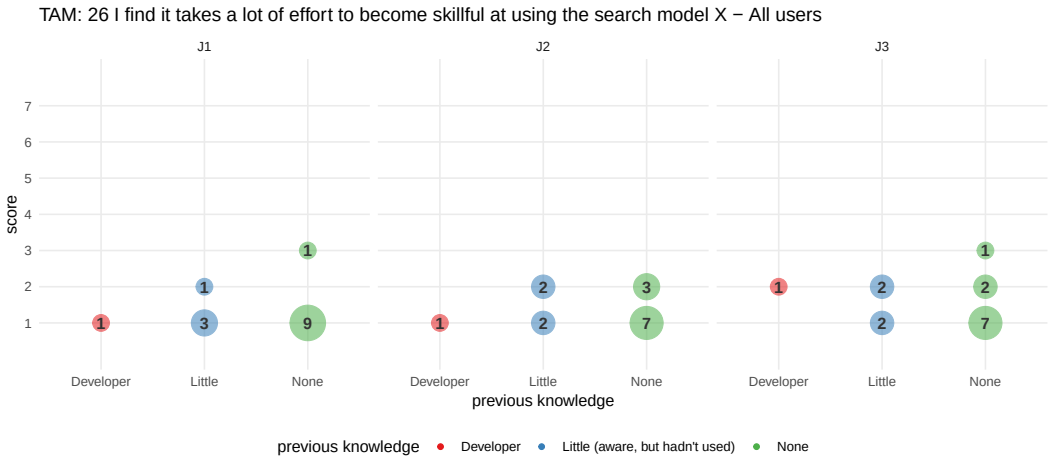



Figure D.21: Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users

### D.1.2.15
### TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users

Table D.21: Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - All users

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0 |
| J2 | 1 | 1 |
| J3 | 1 | 1 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.867$, df=2, p-value=0.2385).

PUC-Rio - Certificação Digital Nº 1721764/CA

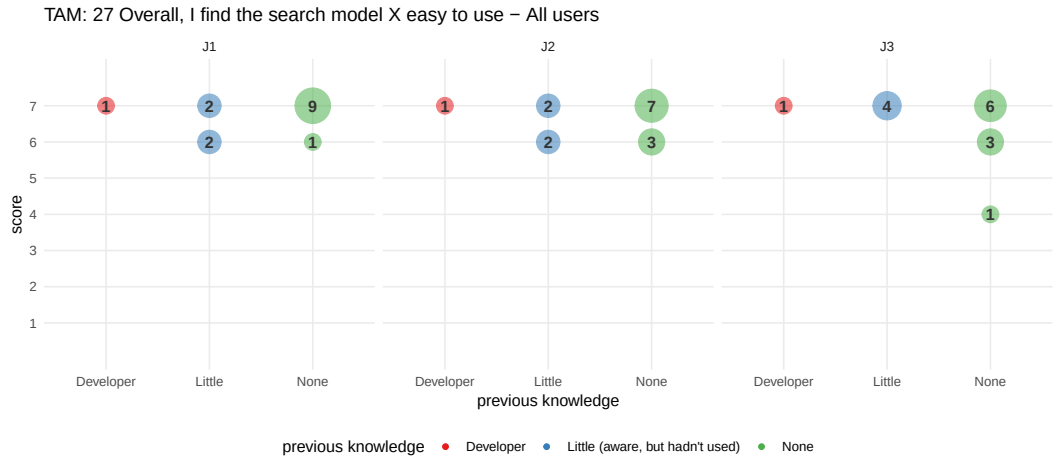



Figure D.22: Response to survey - TAM: 27 Overall, I find the search model X easy to use - All users

### D.1.2.16
### TAM: 27 Overall, I find the search model X easy to use - All users

Table D.22: Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.0 |
| J2 | 7 | 1.0 |
| J3 | 7 | 0.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.511$, df=2, p-value=0.4697).

PUC-Rio - Certificação Digital Nº 1721764/CA

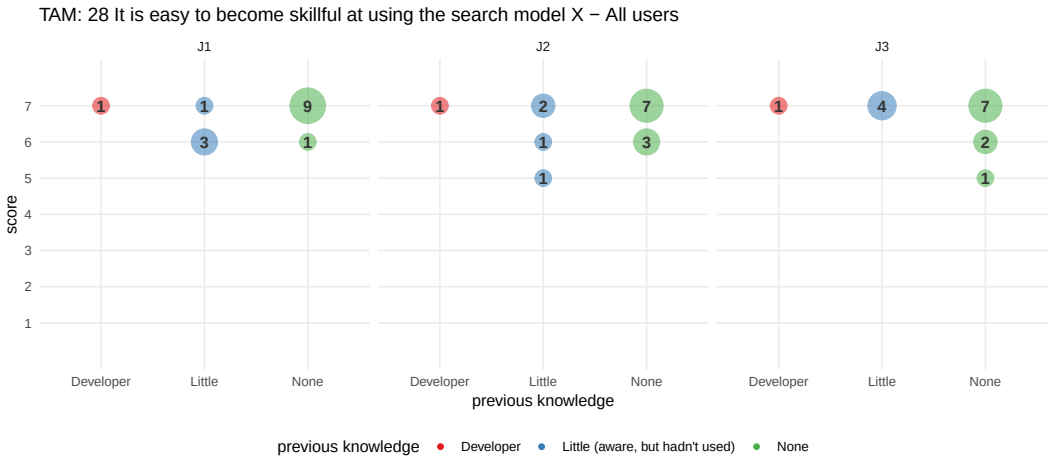



Figure D.23: Response to survey - TAM: 28 It is easy to become skillful at using the search model X - All users

### D.1.2.17
### TAM: 28 It is easy to become skillful at using the search model X - All users

Table D.23: Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.5 |
| J2 | 7 | 1.0 |
| J3 | 7 | 0.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.364$, df=2, p-value=0.5055).

PUC-Rio - Certificação Digital Nº 1721764/CA

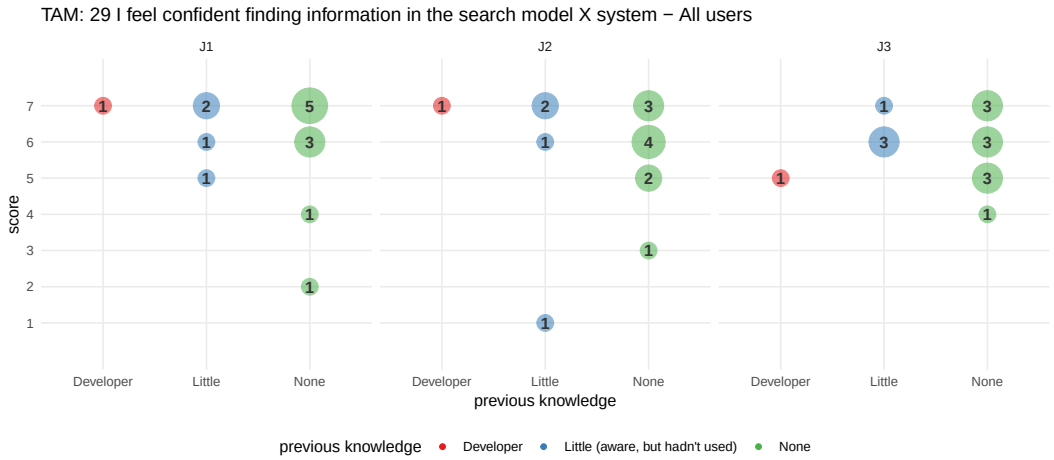



Figure D.24: Response to survey - TAM: 29 I feel confident finding information in the search model X system - All users

### D.1.2.18
### TAM: 29 I feel confident finding information in the search model X system - All users

Table D.24: Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 1.0 |
| J2 | 6 | 1.5 |
| J3 | 6 | 1.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.047$, df=6, p-value=0.5378).

PUC-Rio - Certificação Digital Nº 1721764/CA

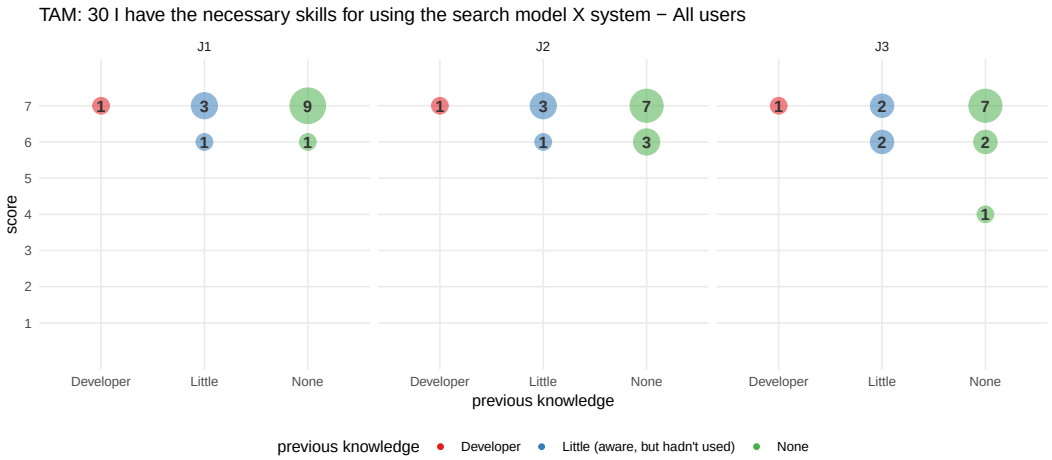



Figure D.25: Response to survey - TAM: 30 I have the necessary skills for using the search model X system - All users

### D.1.2.19
### TAM: 30 I have the necessary skills for using the search model X system - All users

Table D.25: Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.0 |
| J2 | 7 | 0.5 |
| J3 | 7 | 1.0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.442$, df=2, p-value=0.2950).

PUC-Rio - Certificação Digital Nº 1721764/CA

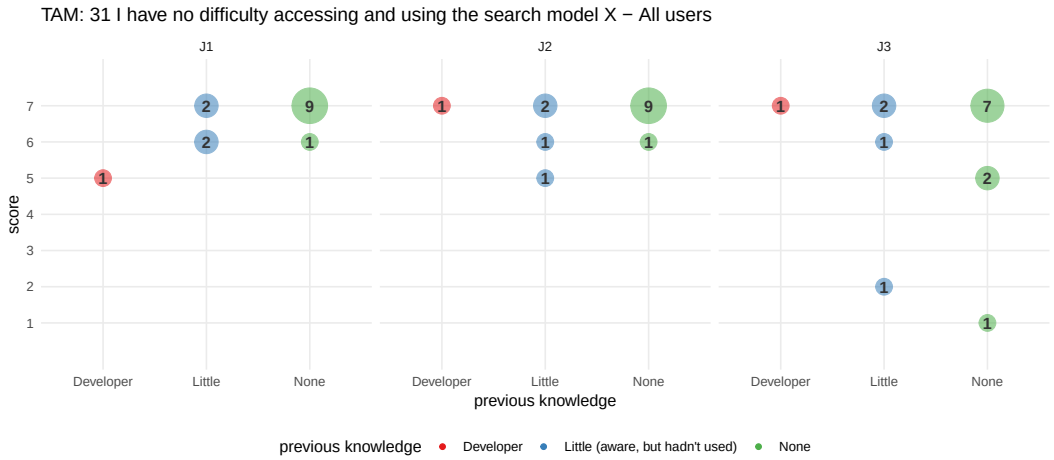



Figure D.26: Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - All users

### D.1.2.20
### TAM: 31 I have no difficulty accessing and using the search model X - All users

Table D.26: Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - All users

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.5 |
| J2 | 7 | 0.0 |
| J3 | 7 | 1.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.322$, df=4, p-value=0.3642).

PUC-Rio - Certificação Digital Nº 1721764/CA

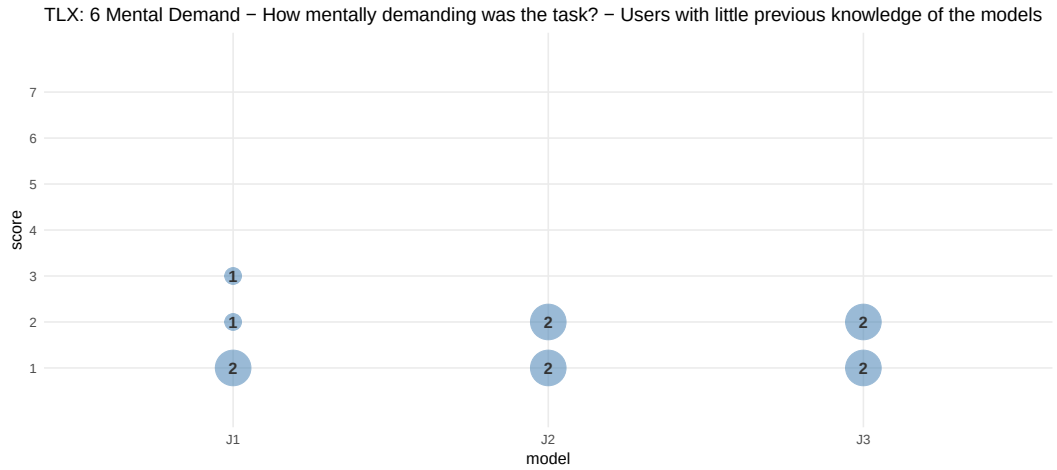



Figure D.27: Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models

## D.2
## Hypothesis tests - Users with little knowledge of model J3

### D.2.1
### TLX

#### D.2.1.1
#### TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models

Table D.27: Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.5 | 1.25 |
| J2 | 1.5 | 1.00 |
| J3 | 1.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.650$, df=2, p-value=0.4382).

PUC-Rio - Certificação Digital Nº 1721764/CA

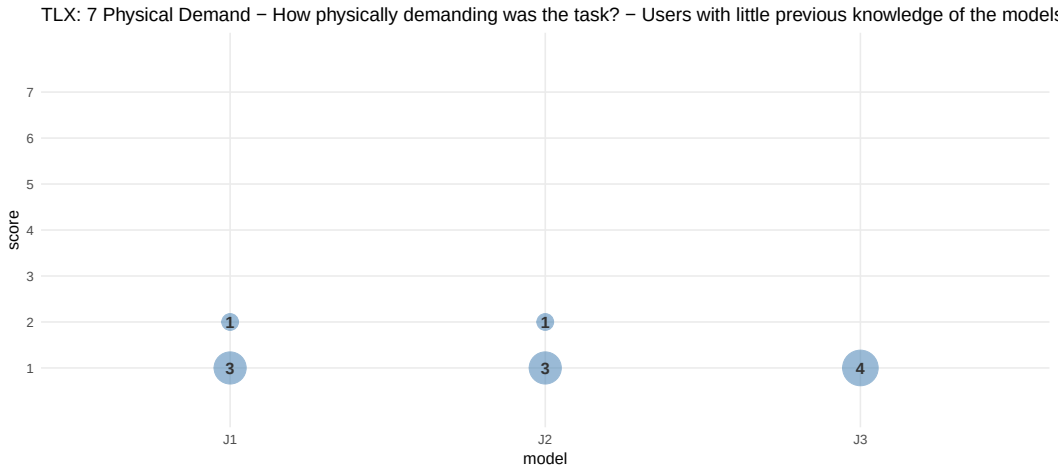



Figure D.28: Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models

### D.2.1.2
### TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models

Table D.28: Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.25 |
| J2 | 1 | 0.25 |
| J3 | 1 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.825$, df=1, p-value=0.3637).

PUC-Rio - Certificação Digital Nº 1721764/CA

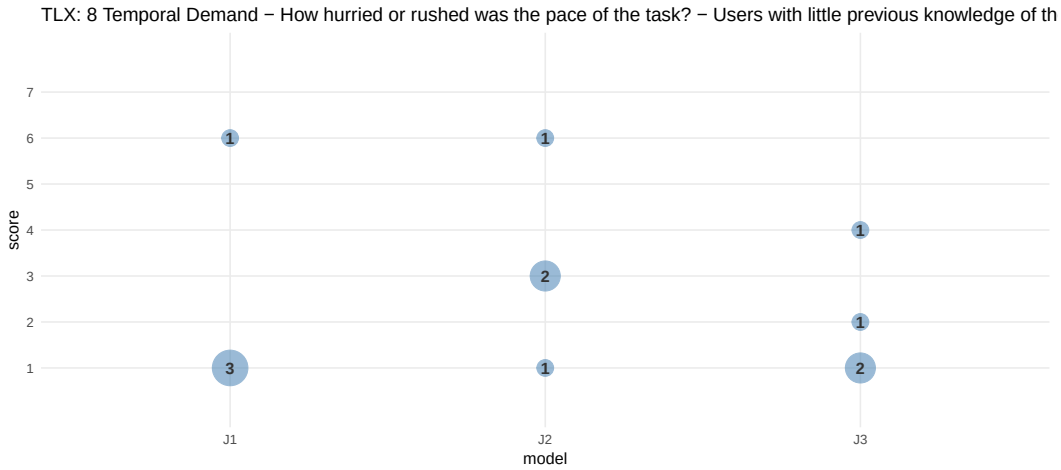



Figure D.29: Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models

### D.2.1.3
### TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models

Table D.29: Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 1.25 |
| J2 | 3.0 | 1.25 |
| J3 | 1.5 | 1.50 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.667$, df=4, p-value=0.4530).

PUC-Rio - Certificação Digital Nº 1721764/CA

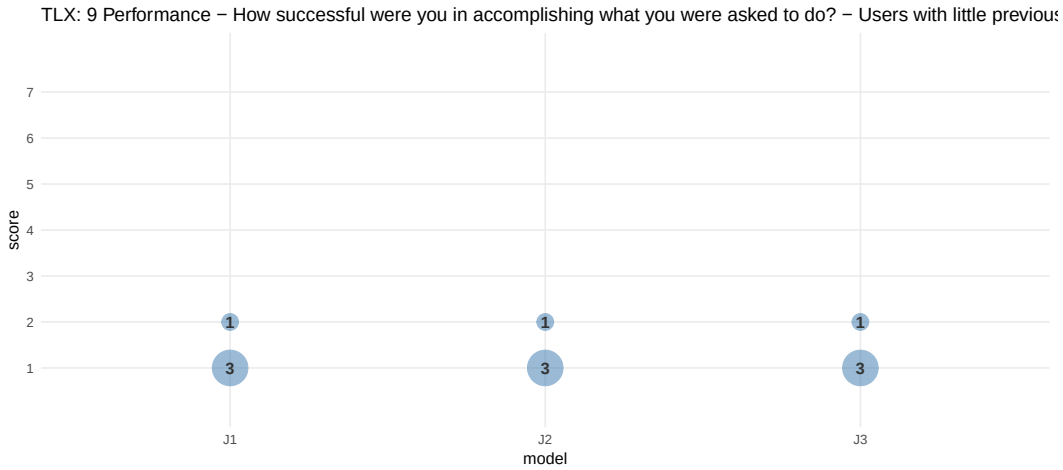



Figure D.30: Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models

### D.2.1.4
### TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models

Table D.30: Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.25 |
| J2 | 1 | 0.25 |
| J3 | 1 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.000$, df=1, p-value=1.0000).

PUC-Rio - Certificação Digital Nº 1721764/CA

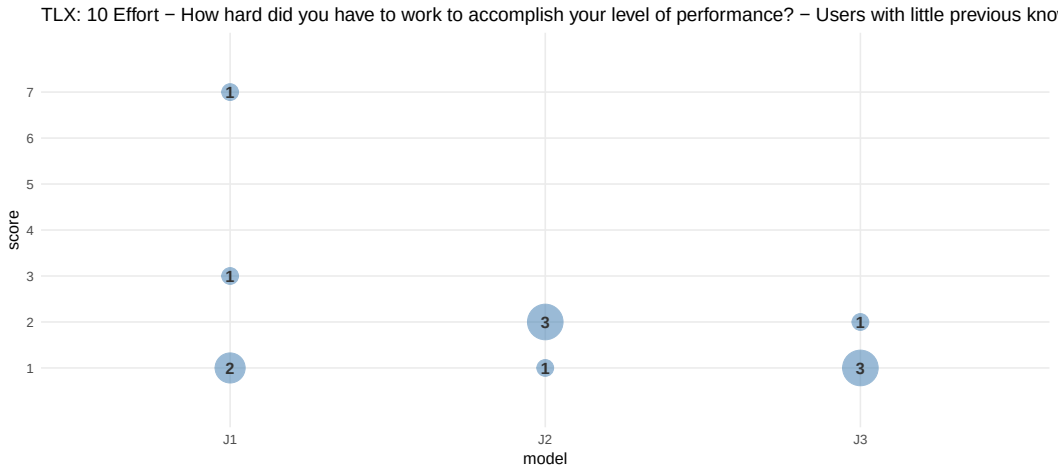



Figure D.31: Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models

### D.2.1.5
### TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models

Table D.31: Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 2 | 3.00 |
| J2 | 2 | 0.25 |
| J3 | 1 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.323$, df=3, p-value=0.3445).

PUC-Rio - Certificação Digital Nº 1721764/CA

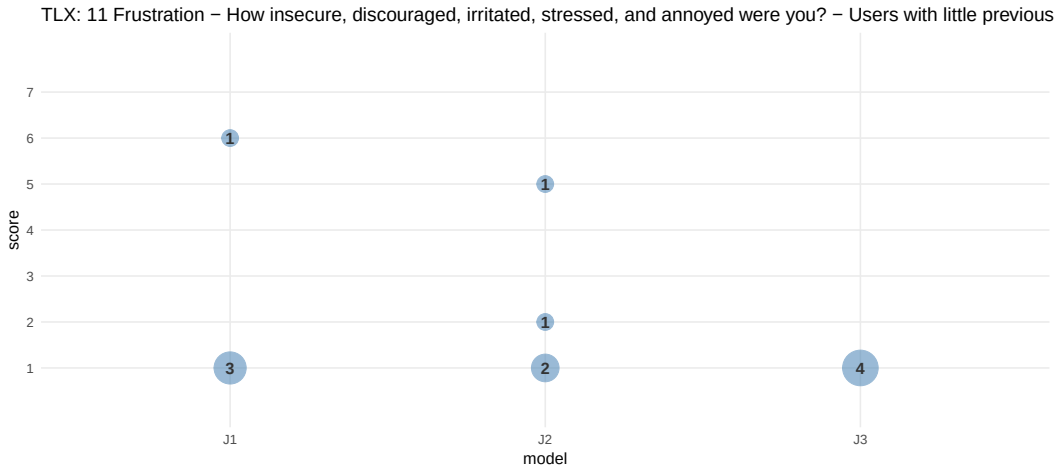



Figure D.32: Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models

### D.2.1.6
### TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models

Table D.32: Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 1.25 |
| J2 | 1.5 | 1.75 |
| J3 | 1.0 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.528$, df=3, p-value=0.6759).

PUC-Rio - Certificação Digital Nº 1721764/CA

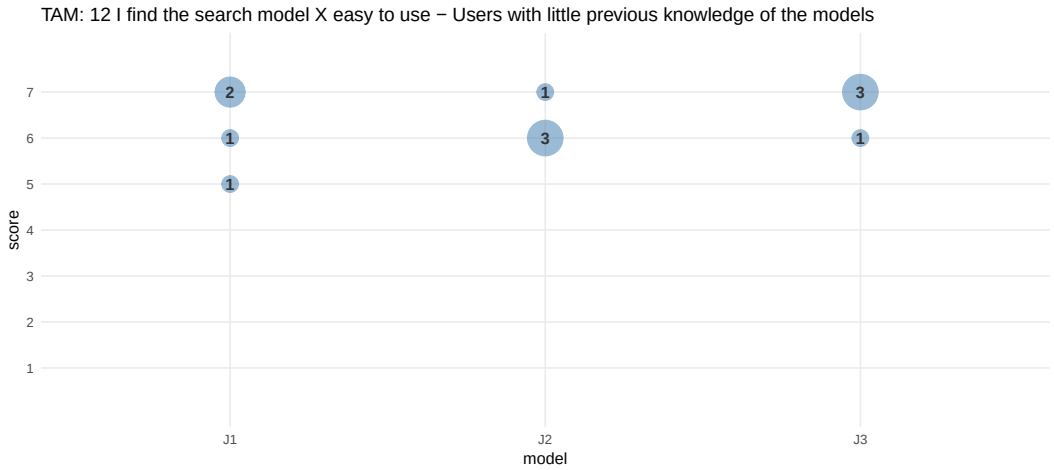



Figure D.33: Response to survey - TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models

## D.2.2 TAM

### D.2.2.1 TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models

Table D.33: Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.25 |
| J2 | 6.0 | 0.25 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.604$, df=2, p-value=0.4484).

PUC-Rio - Certificação Digital Nº 1721764/CA

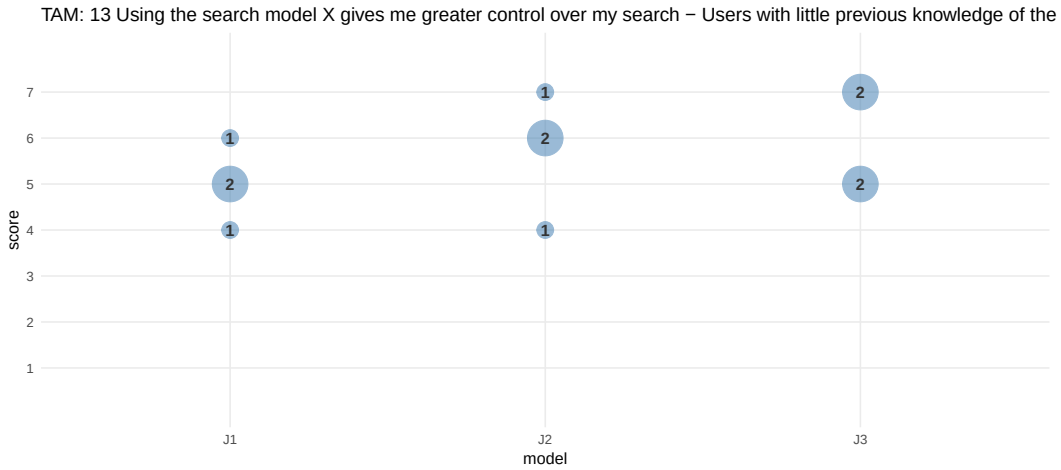



Figure D.34: Response to survey - TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models

### D.2.2.2
### TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models

Table D.34: Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 5 | 0.50 |
| J2 | 6 | 0.75 |
| J3 | 6 | 2.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.979$, df=3, p-value=0.3948).

PUC-Rio - Certificação Digital Nº 1721764/CA

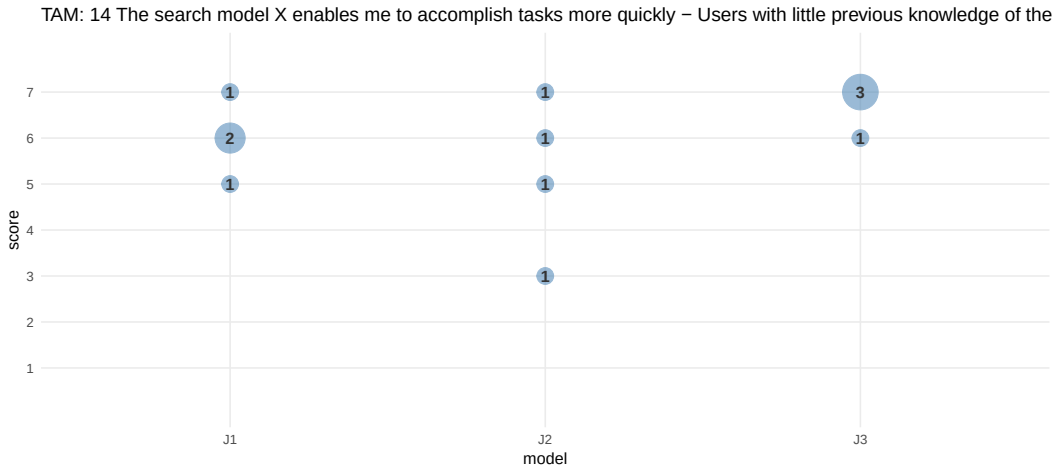



Figure D.35: Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models

### D.2.2.3
### TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models

Table D.35: Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0.50 |
| J2 | 5.5 | 1.75 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.131$, df=3, p-value=0.5456).

PUC-Rio - Certificação Digital Nº 1721764/CA

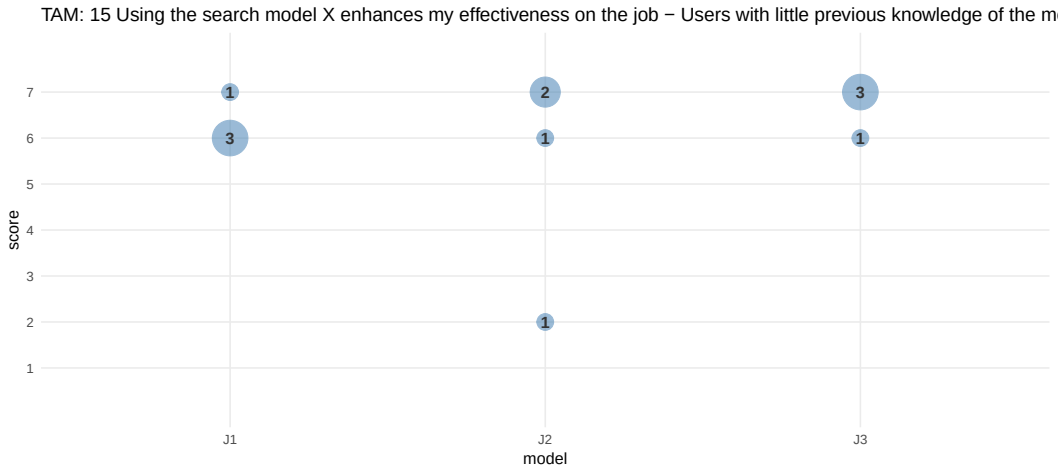



Figure D.36: Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models

### D.2.2.4
### TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models

Table D.36: Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0.25 |
| J2 | 6.5 | 2.00 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.017$, df=2, p-value=0.3648).

PUC-Rio - Certificação Digital Nº 1721764/CA

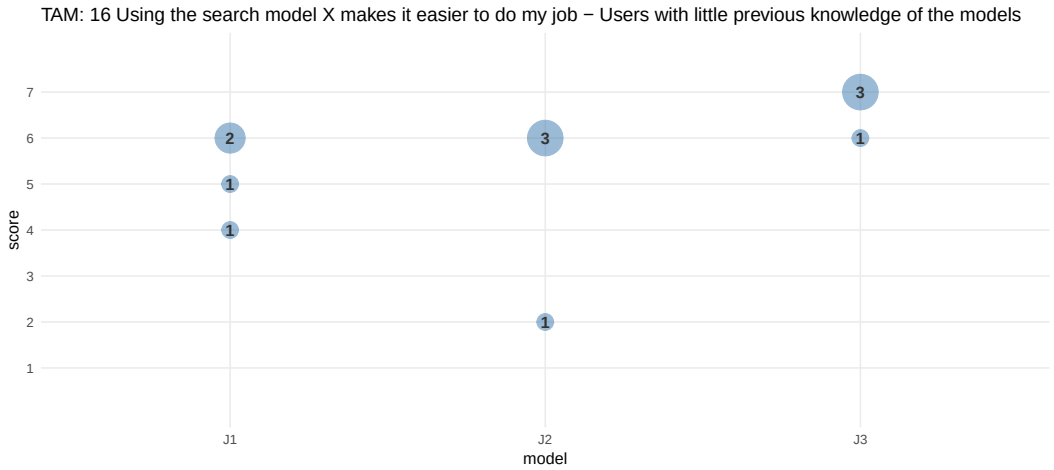



Figure D.37: Response to survey - TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models

### D.2.2.5
### TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models

Table D.37: Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 5.5 | 1.25 |
| J2 | 6.0 | 1.00 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 7.104$, df=4, p-value=0.1305).

PUC-Rio - Certificação Digital Nº 1721764/CA

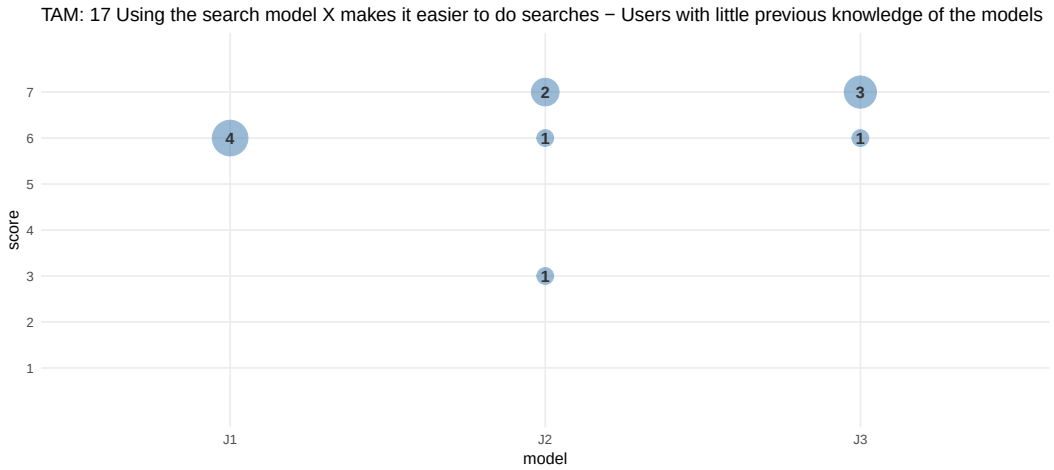



Figure D.38: Response to survey - TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models

### D.2.2.6
### TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models

Table D.38: Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0.00 |
| J2 | 6.5 | 1.75 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.537$, df=2, p-value=0.1034).

PUC-Rio - Certificação Digital Nº 1721764/CA

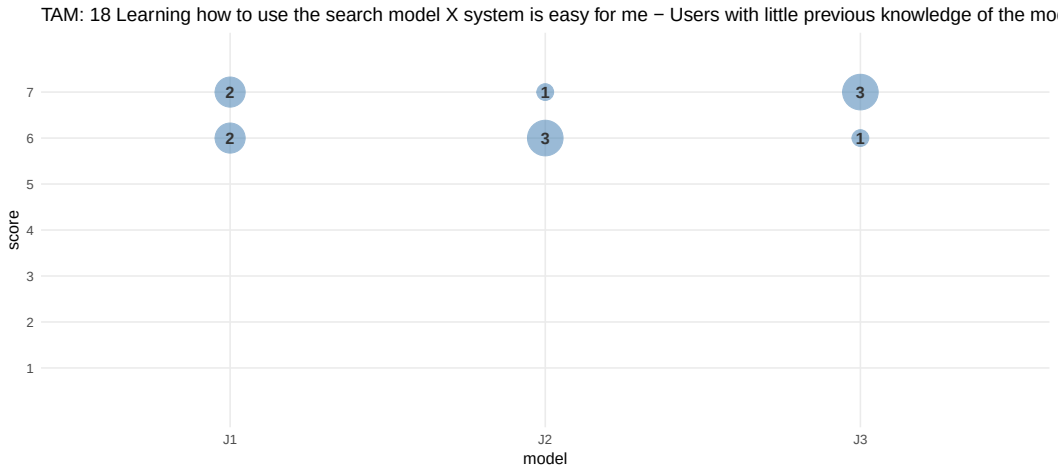



Figure D.39: Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models

### D.2.2.7
### TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models

Table D.39: Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.00 |
| J2 | 6.0 | 0.25 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.458$, df=1, p-value=0.4984).

PUC-Rio - Certificação Digital Nº 1721764/CA

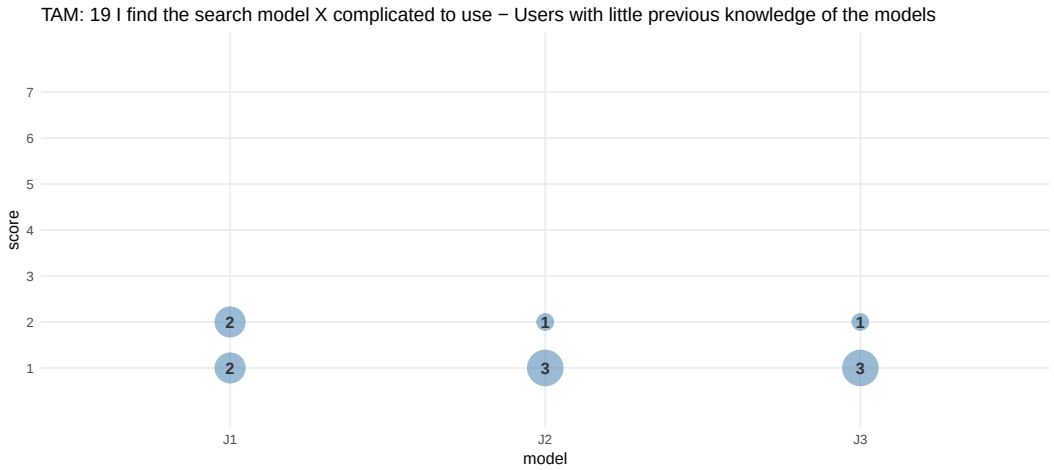



Figure D.40: Response to survey - TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models

### D.2.2.8
### TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models

Table D.40: Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.5 | 1.00 |
| J2 | 1.0 | 0.25 |
| J3 | 1.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.516$, df=1, p-value=0.4727).

PUC-Rio - Certificação Digital Nº 1721764/CA

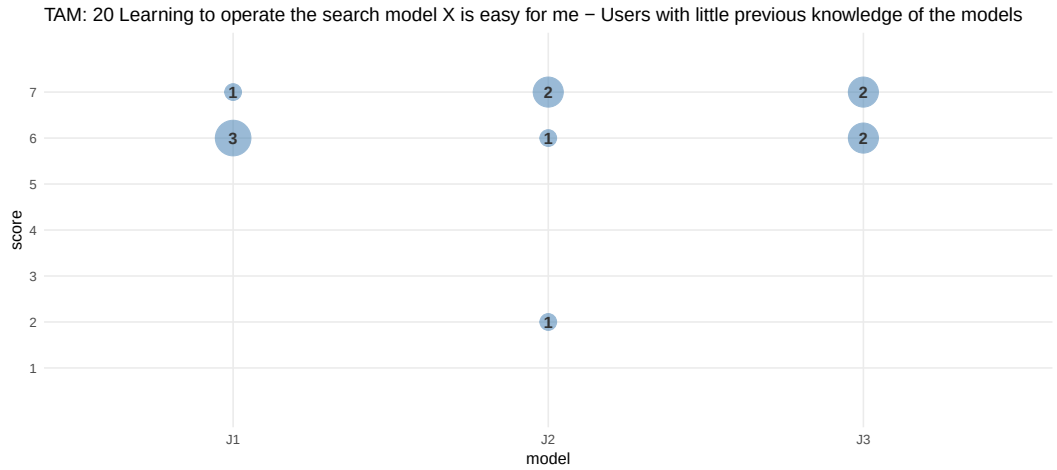



Figure D.41: Response to survey - TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models

### D.2.2.9
### TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models

Table D.41: Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0.25 |
| J2 | 6.5 | 2.00 |
| J3 | 6.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.504$, df=2, p-value=0.7772).

PUC-Rio - Certificação Digital Nº 1721764/CA

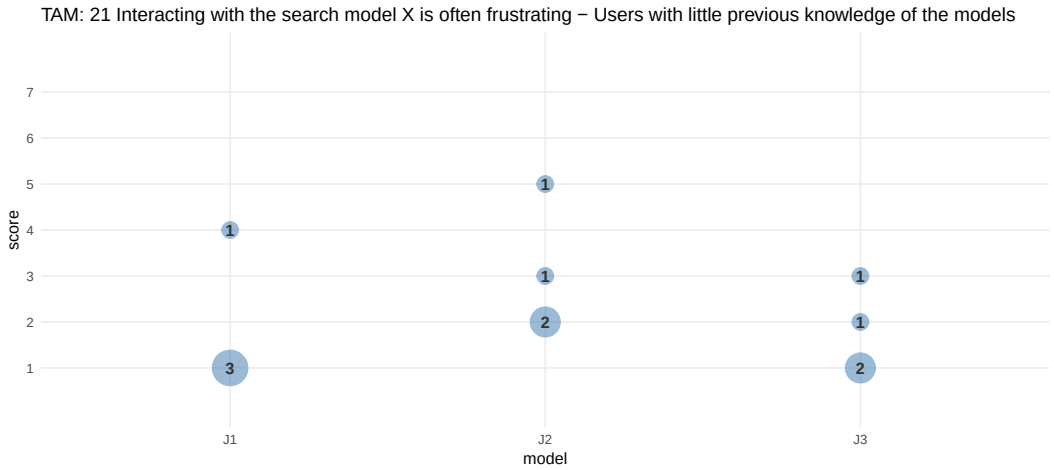



Figure D.42: Response to survey - TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models

### D.2.2.10
### TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models

Table D.42: Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 0.75 |
| J2 | 2.5 | 1.50 |
| J3 | 1.5 | 1.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.796$, df=4, p-value=0.5926).

PUC-Rio - Certificação Digital Nº 1721764/CA

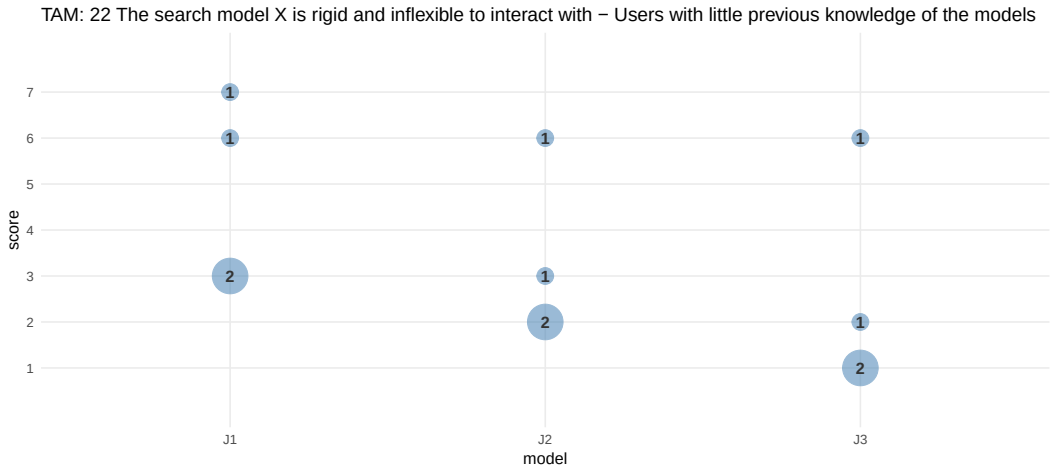



Figure D.43: Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models

### D.2.2.11
### TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models

Table D.43: Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 4.5 | 3.25 |
| J2 | 2.5 | 1.75 |
| J3 | 1.5 | 2.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 6.417$, df=4, p-value=0.1701).

PUC-Rio - Certificação Digital Nº 1721764/CA

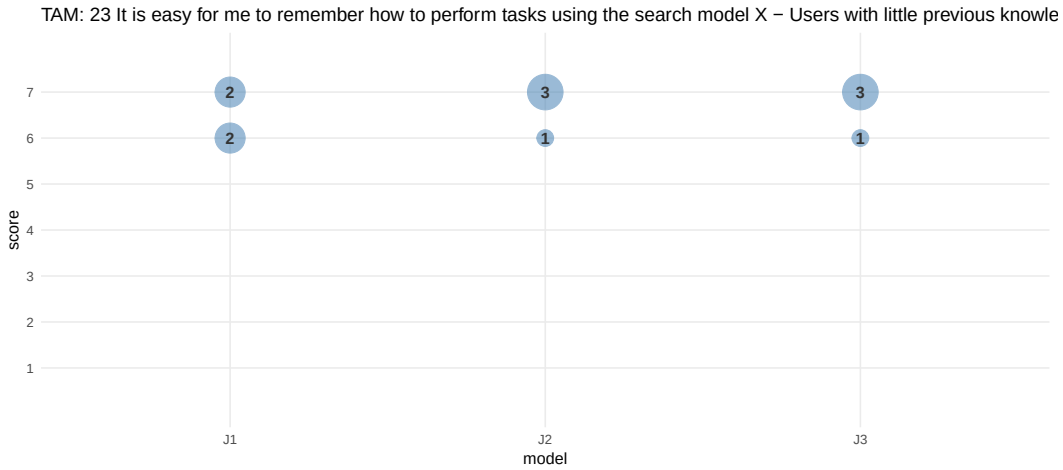



Figure D.44: Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models

### D.2.2.12
### TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models

Table D.44: Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.00 |
| J2 | 7.0 | 0.25 |
| J3 | 7.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.516$, df=1, p-value=0.4727).

PUC-Rio - Certificação Digital Nº 1721764/CA

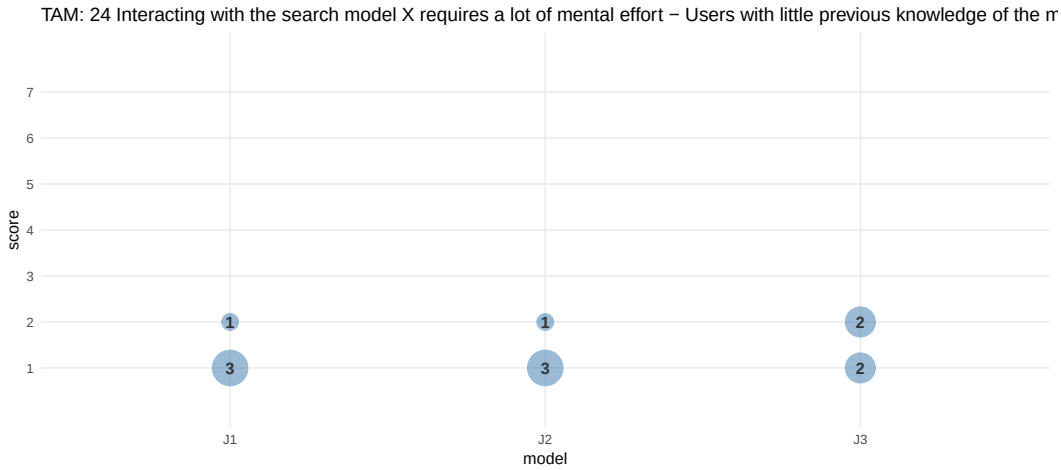



Figure D.45: Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models

### D.2.2.13
### TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models

Table D.45: Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 0.25 |
| J2 | 1.0 | 0.25 |
| J3 | 1.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.516$, df=1, p-value=0.4727).

PUC-Rio - Certificação Digital Nº 1721764/CA

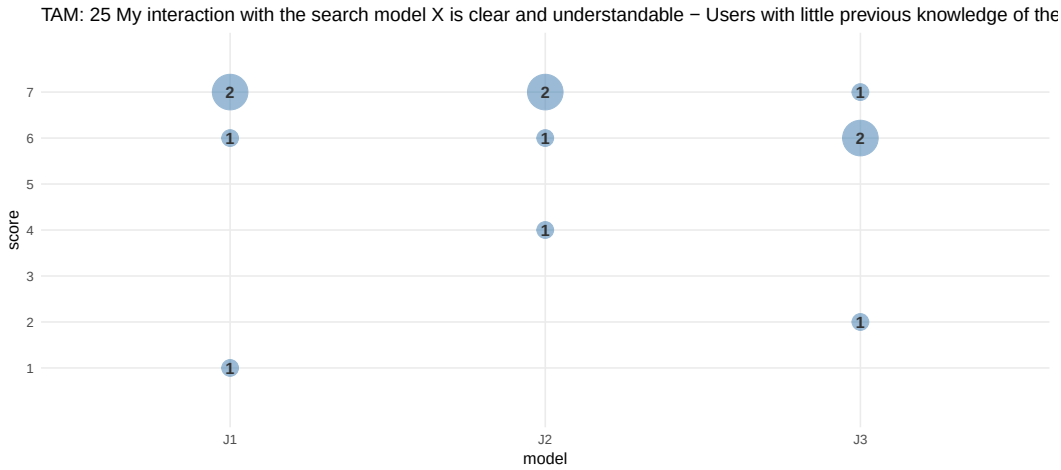



Figure D.46: Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models

### D.2.2.14
### TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models

Table D.46: Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 2.25 |
| J2 | 6.5 | 1.50 |
| J3 | 6.0 | 1.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.369$, df=4, p-value=0.4981).

PUC-Rio - Certificação Digital Nº 1721764/CA

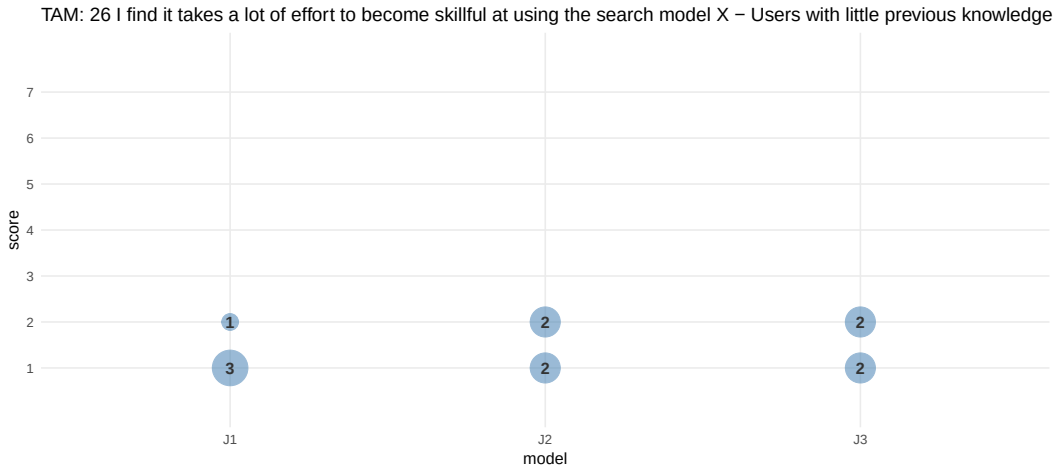



Figure D.47: Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models

### D.2.2.15
### TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models

Table D.47: Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 0.25 |
| J2 | 1.5 | 1.00 |
| J3 | 1.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.471$, df=1, p-value=0.4923).

PUC-Rio - Certificação Digital Nº 1721764/CA

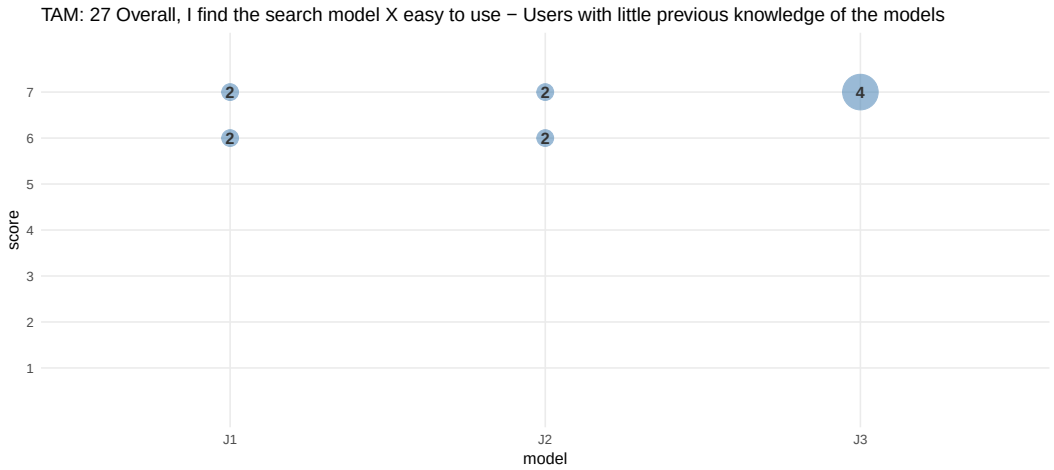



Figure D.48: Response to survey - TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models

### D.2.2.16
### TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models

Table D.48: Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1 |
| J2 | 6.5 | 1 |
| J3 | 7.0 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.063$, df=1, p-value=0.1510).

PUC-Rio - Certificação Digital Nº 1721764/CA

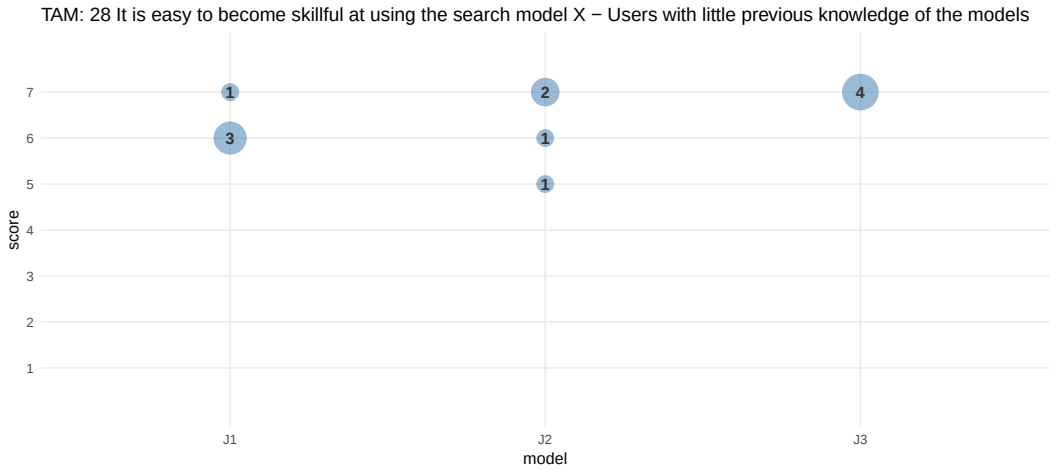



Figure D.49: Response to survey - TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models

### D.2.2.17
### TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models

Table D.49: Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.0 | 0.25 |
| J2 | 6.5 | 1.25 |
| J3 | 7.0 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.862$, df=2, p-value=0.0880).

PUC-Rio - Certificação Digital Nº 1721764/CA

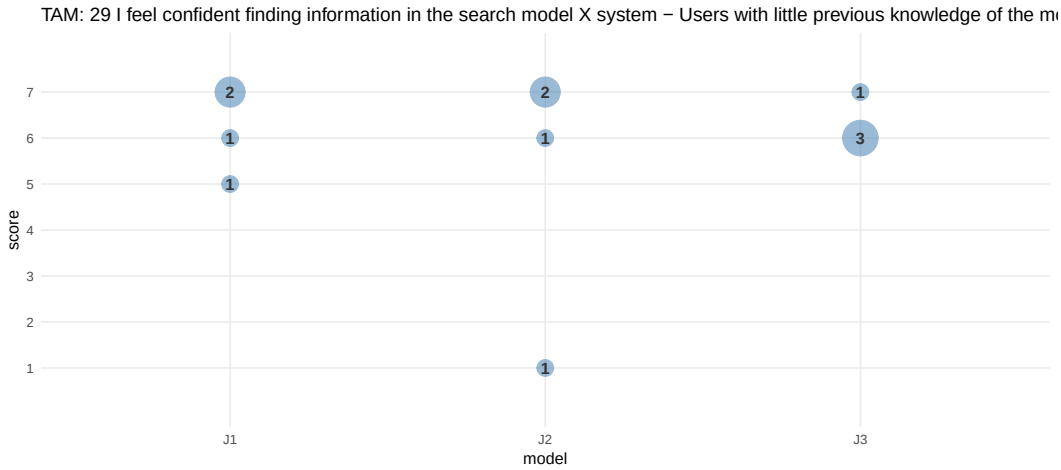



Figure D.50: Response to survey - TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models

### D.2.2.18
### TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models

Table D.50: Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.25 |
| J2 | 6.5 | 2.25 |
| J3 | 6.0 | 0.25 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.750$, df=3, p-value=0.4318).

PUC-Rio - Certificação Digital Nº 1721764/CA

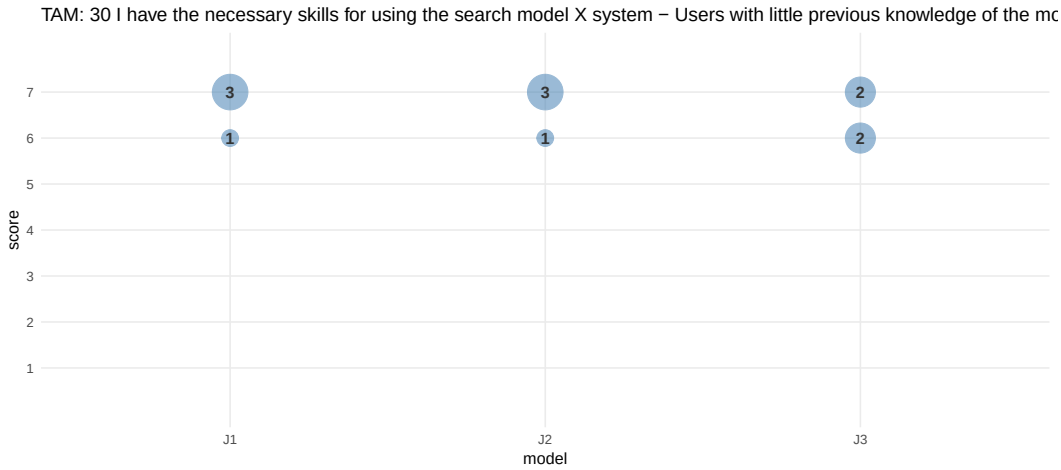



Figure D.51: Response to survey - TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models

### D.2.2.19
### TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models

Table D.51: Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7.0 | 0.25 |
| J2 | 7.0 | 0.25 |
| J3 | 6.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.516$, df=1, p-value=0.4727).

PUC-Rio - Certificação Digital Nº 1721764/CA

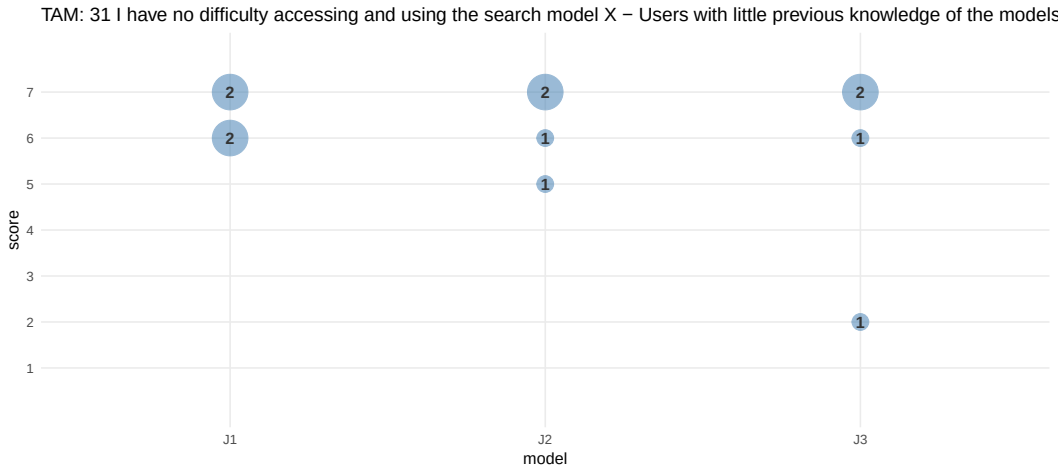



Figure D.52: Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models

### D.2.2.20
### TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models

Table D.52: Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - Users with little previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.00 |
| J2 | 6.5 | 1.25 |
| J3 | 6.5 | 2.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.719$, df=3, p-value=0.6328).

PUC-Rio - Certificação Digital Nº 1721764/CA

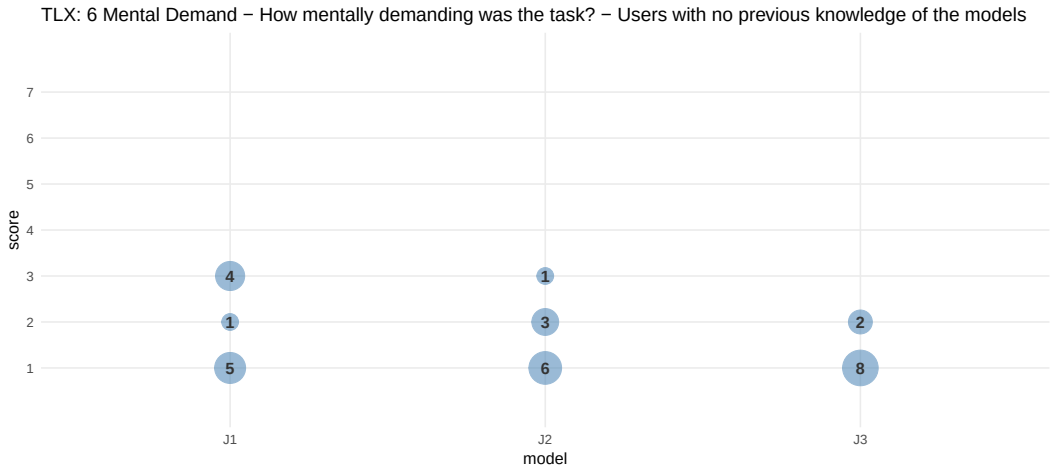



Figure D.53: Response to survey - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models

## D.3
## Hypothesis tests - Users with no knowledge of model J3

### D.3.1
### TLX

#### D.3.1.1
#### TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models

Table D.53: Median and interquartile range of scores per model - TLX: 6 Mental Demand - How mentally demanding was the task? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.5 | 2 |
| J2 | 1.0 | 1 |
| J3 | 1.0 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.569$, df=2, p-value=0.0618).

PUC-Rio - Certificação Digital Nº 1721764/CA

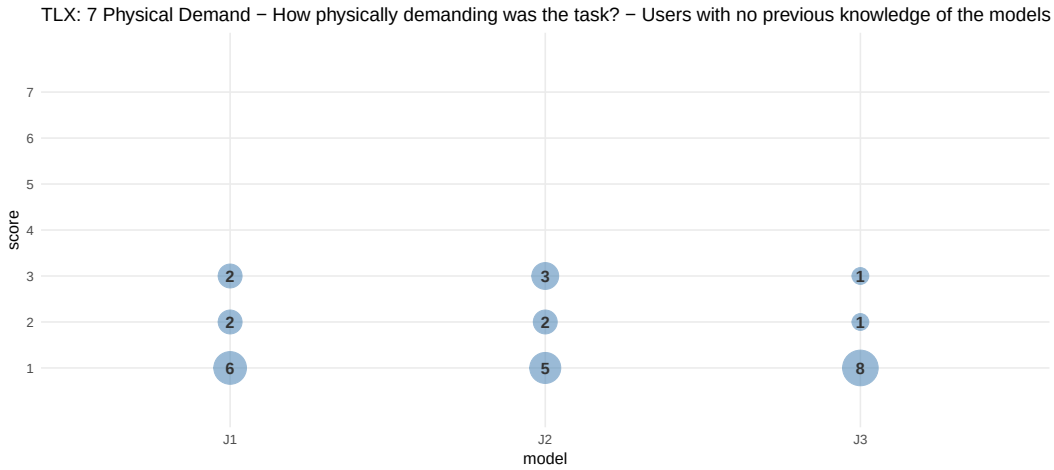



Figure D.54: Response to survey - TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models

### D.3.1.2
### TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models

Table D.54: Median and interquartile range of scores per model - TLX: 7 Physical Demand - How physically demanding was the task? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 1.00 |
| J2 | 1.5 | 1.75 |
| J3 | 1.0 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.837$, df=2, p-value=0.6581).

PUC-Rio - Certificação Digital Nº 1721764/CA

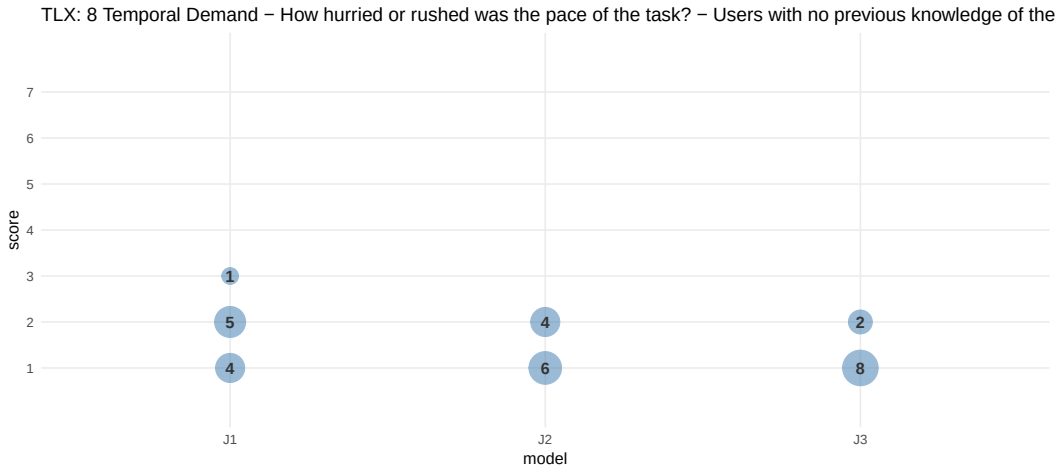



Figure D.55: Response to survey - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models

### D.3.1.3
### TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models

Table D.55: Median and interquartile range of scores per model - TLX: 8 Temporal Demand - How hurried or rushed was the pace of the task? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 2 | 1 |
| J2 | 1 | 1 |
| J3 | 1 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 3.925$, df=2, p-value=0.1405).

PUC-Rio - Certificação Digital Nº 1721764/CA

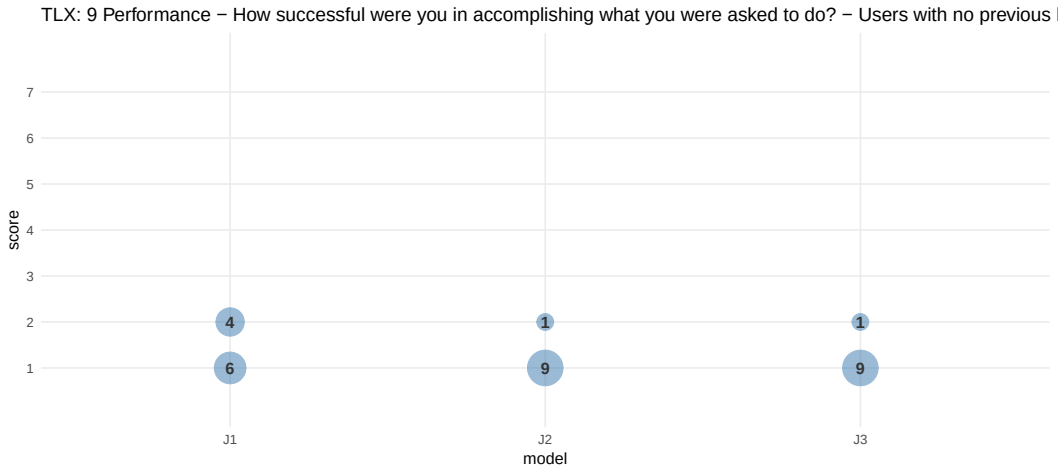



Figure D.56: Response to survey - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models

### D.3.1.4
### TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models

Table D.56: Median and interquartile range of scores per model - TLX: 9 Performance - How successful were you in accomplishing what you were asked to do? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 1 |
| J2 | 1 | 0 |
| J3 | 1 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.719$, df=1, p-value=0.0992).

PUC-Rio - Certificação Digital Nº 1721764/CA

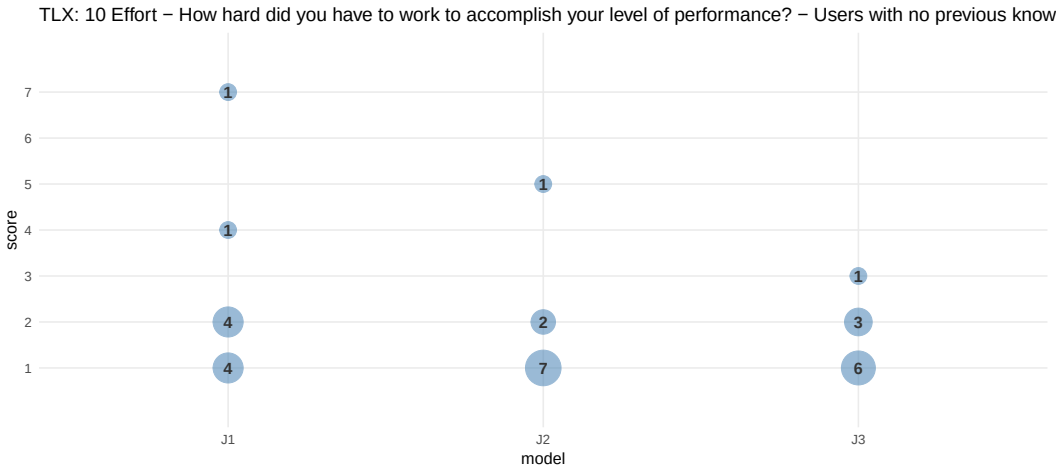



Figure D.57: Response to survey - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models

### D.3.1.5
### TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models

Table D.57: Median and interquartile range of scores per model - TLX: 10 Effort - How hard did you have to work to accomplish your level of performance? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 2 | 1.00 |
| J2 | 1 | 0.75 |
| J3 | 1 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.852$, df=5, p-value=0.4342).

PUC-Rio - Certificação Digital Nº 1721764/CA

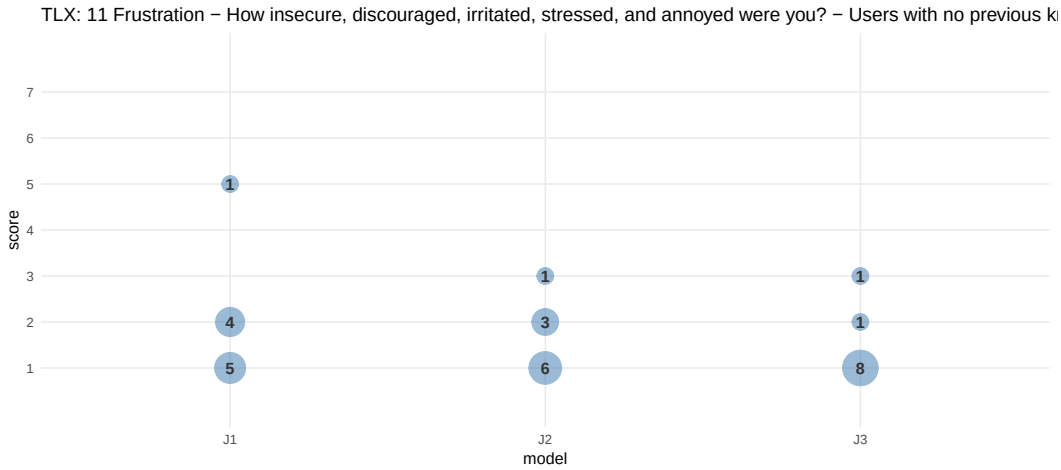



Figure D.58: Response to survey - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models

### D.3.1.6
### TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models

Table D.58: Median and interquartile range of scores per model - TLX: 11 Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you? - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.5 | 1 |
| J2 | 1.0 | 1 |
| J3 | 1.0 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.493$, df=3, p-value=0.2129).

PUC-Rio - Certificação Digital Nº 1721764/CA

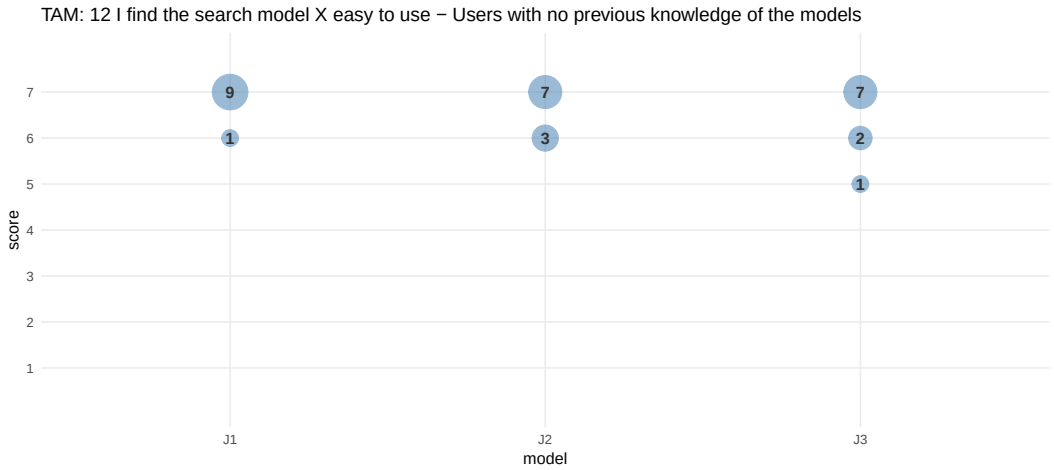



Figure D.59: Response to survey - TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models

## D.3.2 TAM

### D.3.2.1 TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models

Table D.59: Median and interquartile range of scores per model - TAM: 12 I find the search model X easy to use - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.75 |
| J3 | 7 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.944$, df=2, p-value=0.3784).

PUC-Rio - Certificação Digital Nº 1721764/CA

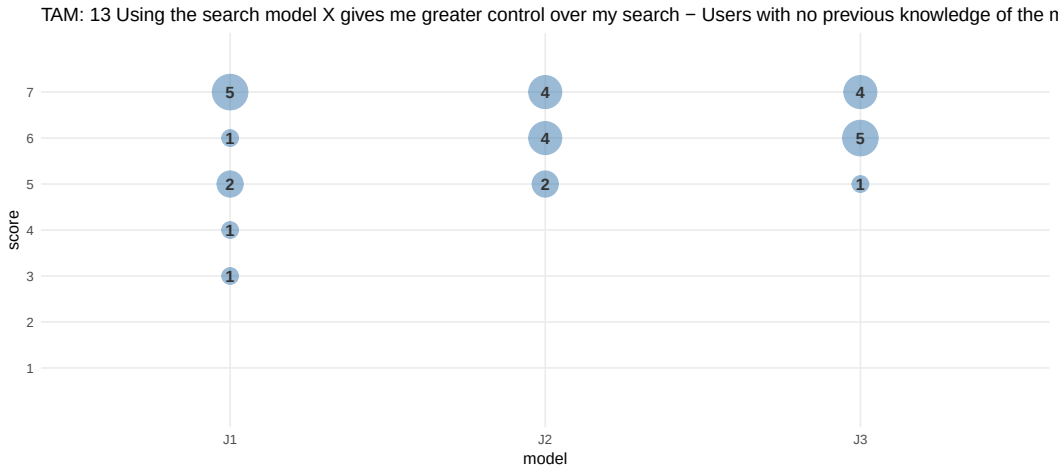



Figure D.60: Response to survey - TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models

### D.3.2.2
### TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models

Table D.60: Median and interquartile range of scores per model - TAM: 13 Using the search model X gives me greater control over my search - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 2 |
| J2 | 6.0 | 1 |
| J3 | 6.0 | 1 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.622$, df=4, p-value=0.2293).

PUC-Rio - Certificação Digital Nº 1721764/CA

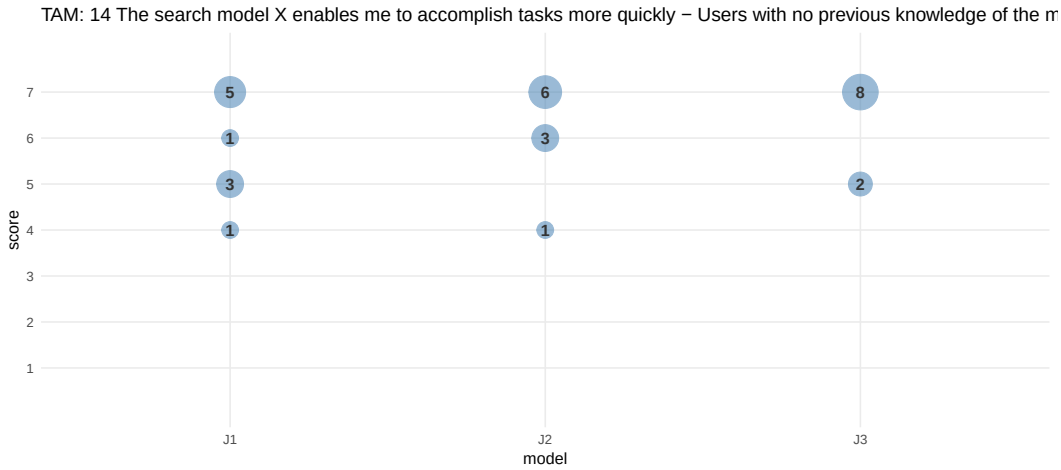



Figure D.61: Response to survey - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models

### D.3.2.3
### TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models

Table D.61: Median and interquartile range of scores per model - TAM: 14 The search model X enables me to accomplish tasks more quickly - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 2 |
| J2 | 7.0 | 1 |
| J3 | 7.0 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.064$, df=3, p-value=0.5592).

PUC-Rio - Certificação Digital Nº 1721764/CA

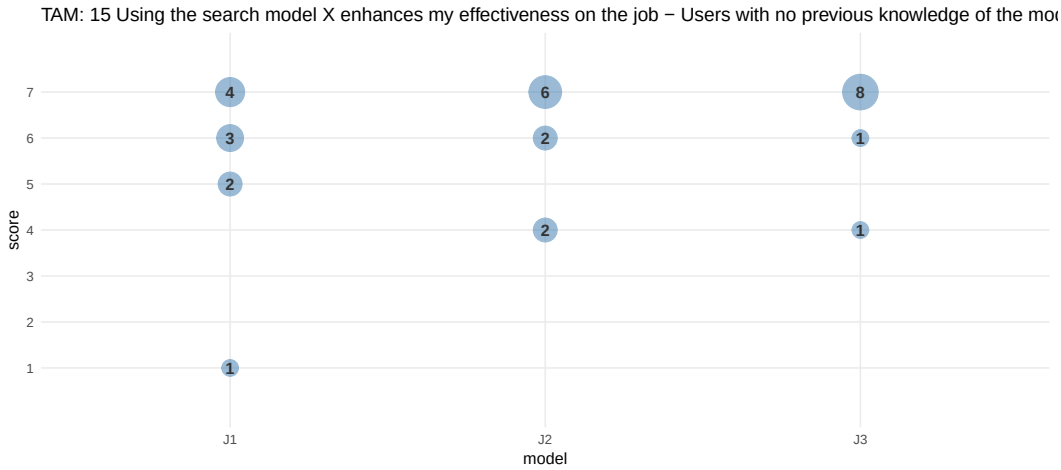



Figure D.62: Response to survey - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models

### D.3.2.4
### TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models

Table D.62: Median and interquartile range of scores per model - TAM: 15 Using the search model X enhances my effectiveness on the job - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6 | 1.75 |
| J2 | 7 | 1.00 |
| J3 | 7 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 7.089$, df=4, p-value=0.1313).

PUC-Rio - Certificação Digital Nº 1721764/CA

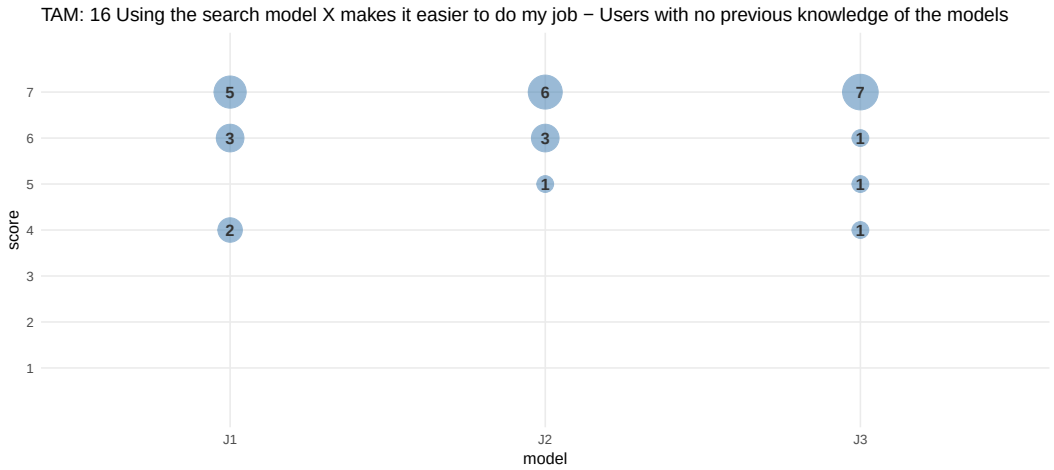



Figure D.63: Response to survey - TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models

### D.3.2.5
### TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models

Table D.63: Median and interquartile range of scores per model - TAM: 16 Using the search model X makes it easier to do my job - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.00 |
| J2 | 7.0 | 1.00 |
| J3 | 7.0 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.359$, df=3, p-value=0.5013).

PUC-Rio - Certificação Digital Nº 1721764/CA

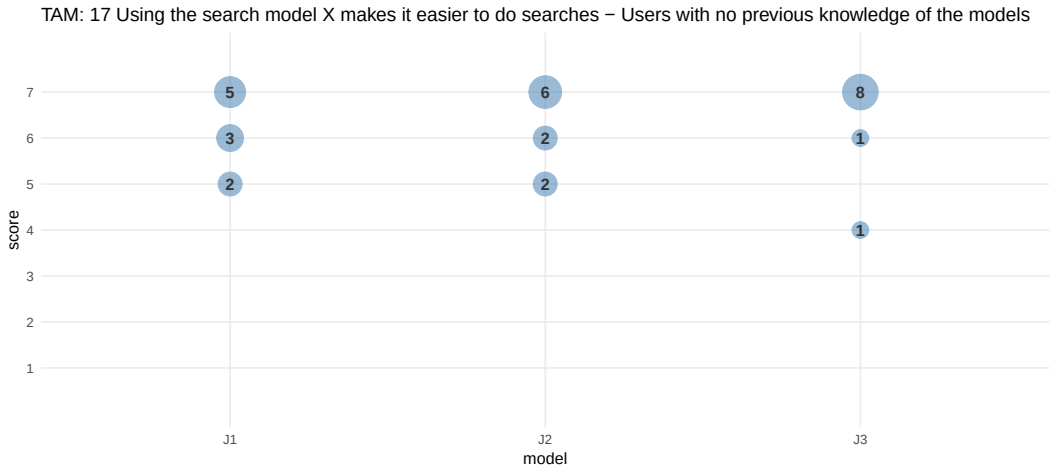



Figure D.64: Response to survey - TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models

### D.3.2.6
### TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models

Table D.64: Median and interquartile range of scores per model - TAM: 17 Using the search model X makes it easier to do searches - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1 |
| J2 | 7.0 | 1 |
| J3 | 7.0 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.554$, df=3, p-value=0.2076).

PUC-Rio - Certificação Digital Nº 1721764/CA

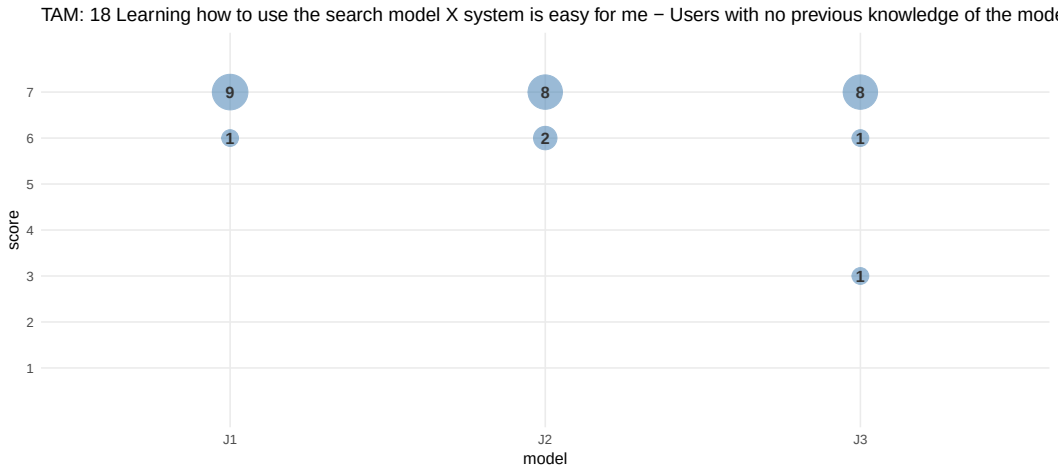



Figure D.65: Response to survey - TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models

### D.3.2.7
### TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models

Table D.65: Median and interquartile range of scores per model - TAM: 18 Learning how to use the search model X system is easy for me - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0 |
| J2 | 7 | 0 |
| J3 | 7 | 0 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.508$, df=2, p-value=0.4705).

PUC-Rio - Certificação Digital Nº 1721764/CA

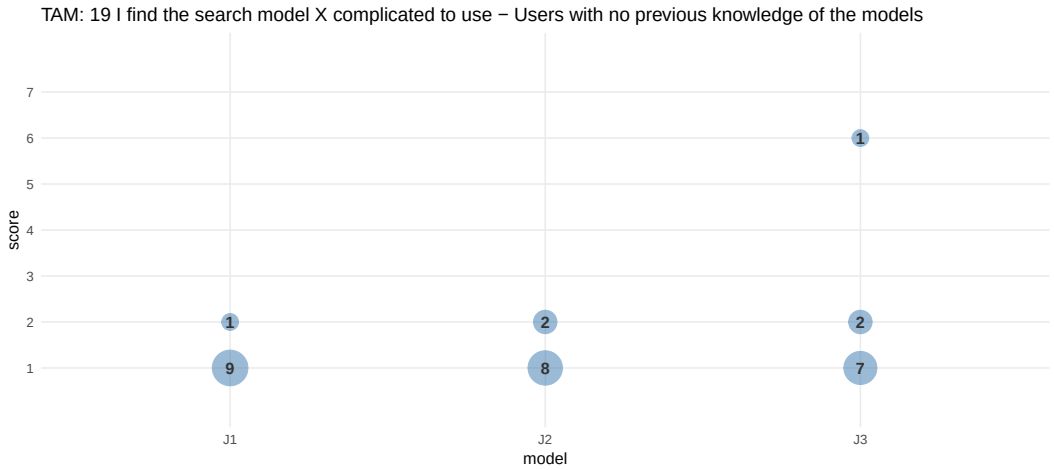



Figure D.66: Response to survey - TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models

### D.3.2.8
### TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models

Table D.66: Median and interquartile range of scores per model - TAM: 19 I find the search model X complicated to use - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.00 |
| J2 | 1 | 0.00 |
| J3 | 1 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.982$, df=2, p-value=0.3713).

PUC-Rio - Certificação Digital Nº 1721764/CA

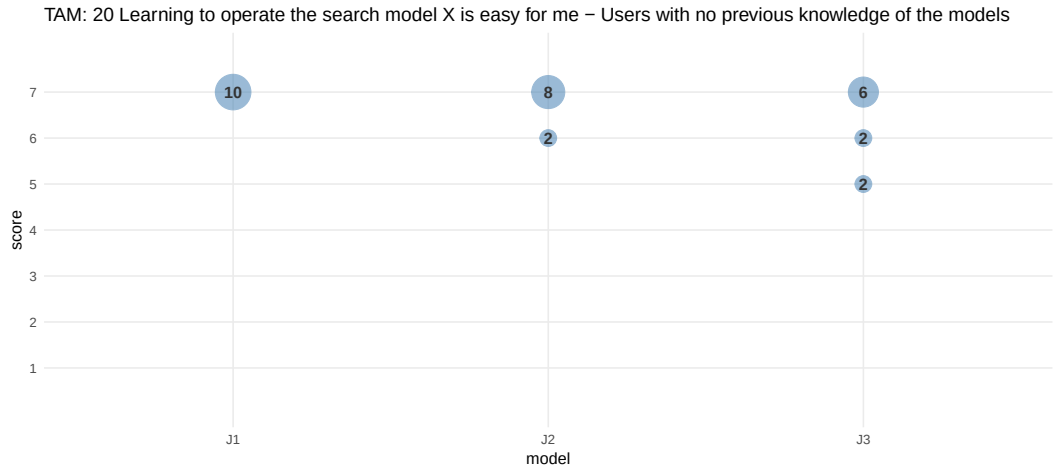



Figure D.67: Response to survey - TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models

### D.3.2.9
### TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models

Table D.67: Median and interquartile range of scores per model - TAM: 20 Learning to operate the search model X is easy for me - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0 |
| J2 | 7 | 0 |
| J3 | 7 | 1 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.317$, df=2, p-value=0.0701).

PUC-Rio - Certificação Digital Nº 1721764/CA

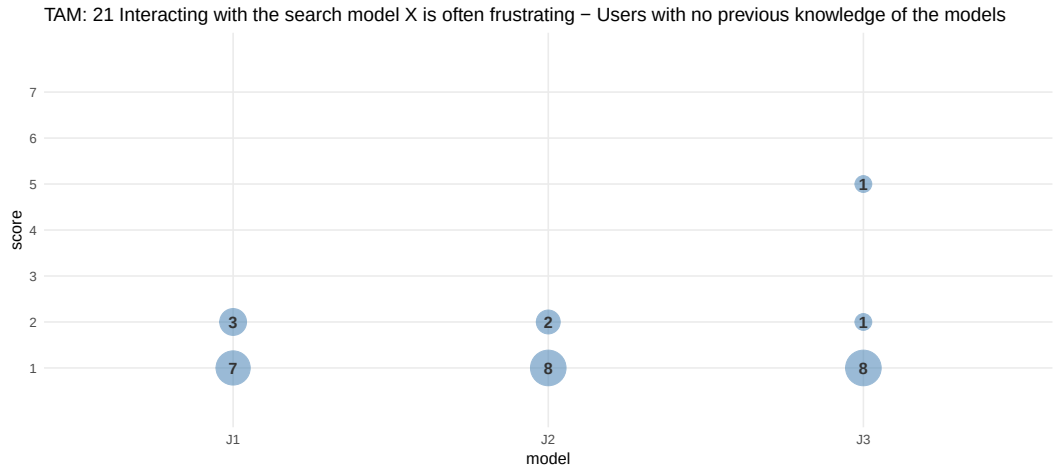



Figure D.68: Response to survey - TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models

### D.3.2.10
### TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models

Table D.68: Median and interquartile range of scores per model - TAM: 21 Interacting with the search model X is often frustrating - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.75 |
| J2 | 1 | 0.00 |
| J3 | 1 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.480$, df=2, p-value=0.2894).

PUC-Rio - Certificação Digital Nº 1721764/CA

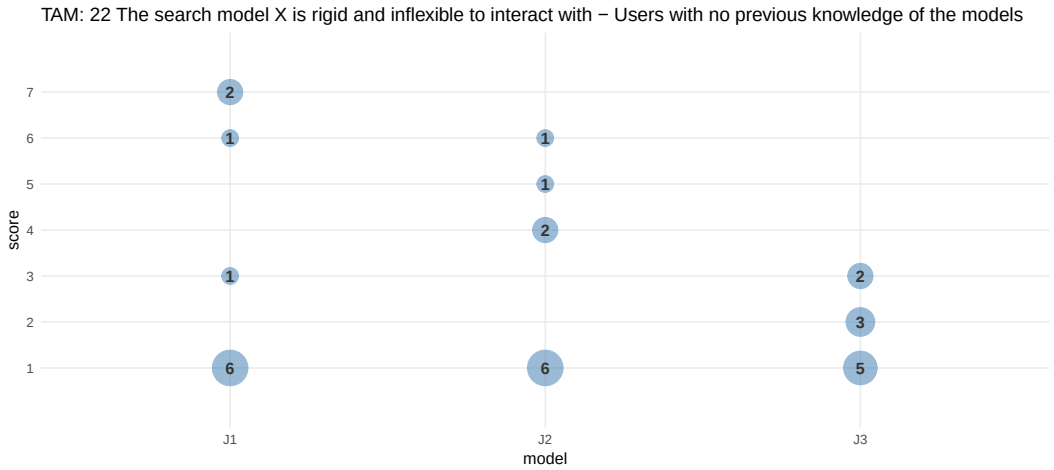



Figure D.69: Response to survey - TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models

### D.3.2.11
### TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models

Table D.69: Median and interquartile range of scores per model - TAM: 22 The search model X is rigid and inflexible to interact with - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1.0 | 4.25 |
| J2 | 1.0 | 3.00 |
| J3 | 1.5 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 8.544$, df=6, p-value=0.2009).

PUC-Rio - Certificação Digital Nº 1721764/CA

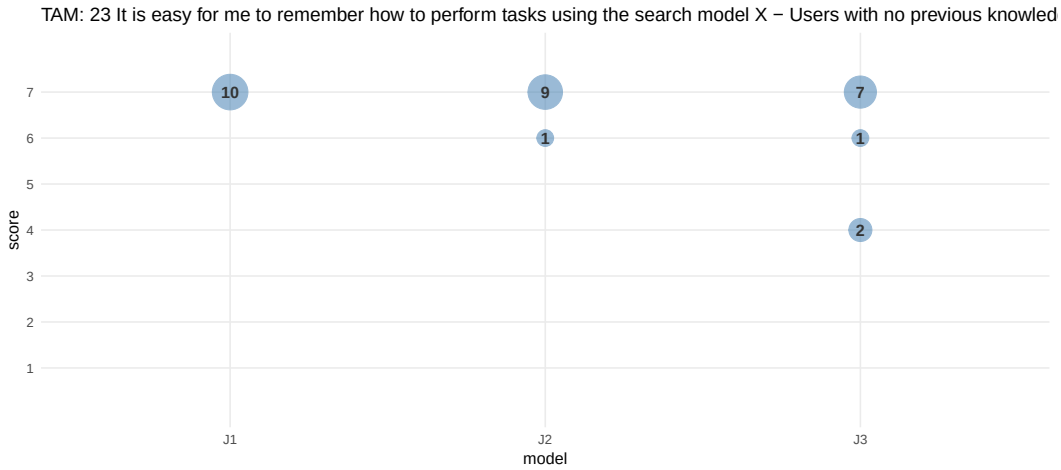



Figure D.70: Response to survey - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models

### D.3.2.12
### TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models

Table D.70: Median and interquartile range of scores per model - TAM: 23 It is easy for me to remember how to perform tasks using the search model X - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.00 |
| J3 | 7 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.127$, df=2, p-value=0.1270).

PUC-Rio - Certificação Digital Nº 1721764/CA

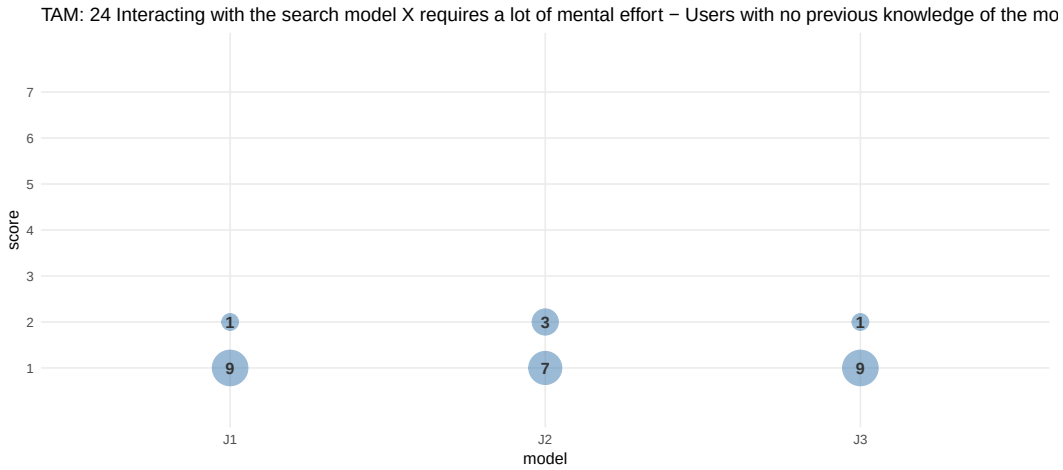



Figure D.71: Response to survey - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models

### D.3.2.13
### TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models

Table D.71: Median and interquartile range of scores per model - TAM: 24 Interacting with the search model X requires a lot of mental effort - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.00 |
| J2 | 1 | 0.75 |
| J3 | 1 | 0.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 0.000$, df=1, p-value=1.0000).

PUC-Rio - Certificação Digital Nº 1721764/CA

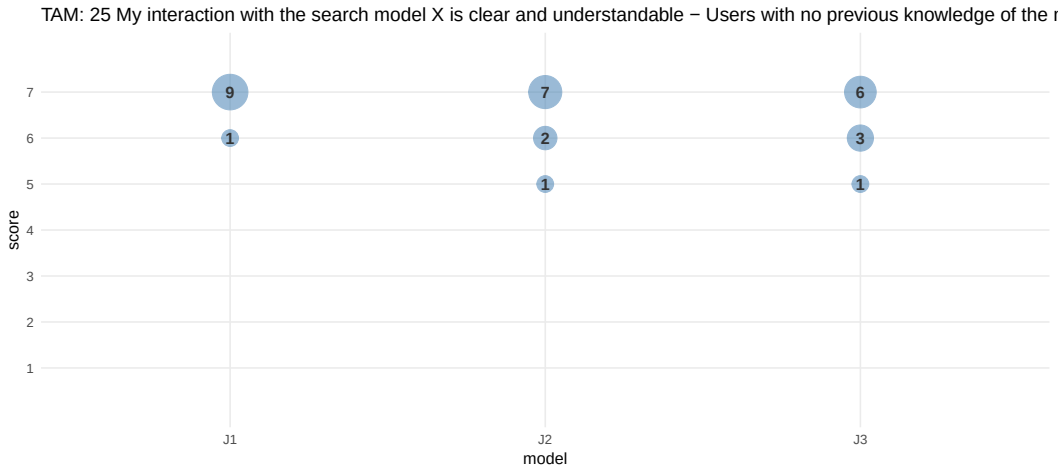



Figure D.72: Response to survey - TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models

### D.3.2.14
### TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models

Table D.72: Median and interquartile range of scores per model - TAM: 25 My interaction with the search model X is clear and understandable - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.75 |
| J3 | 7 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.285$, df=2, p-value=0.3190).

PUC-Rio - Certificação Digital Nº 1721764/CA

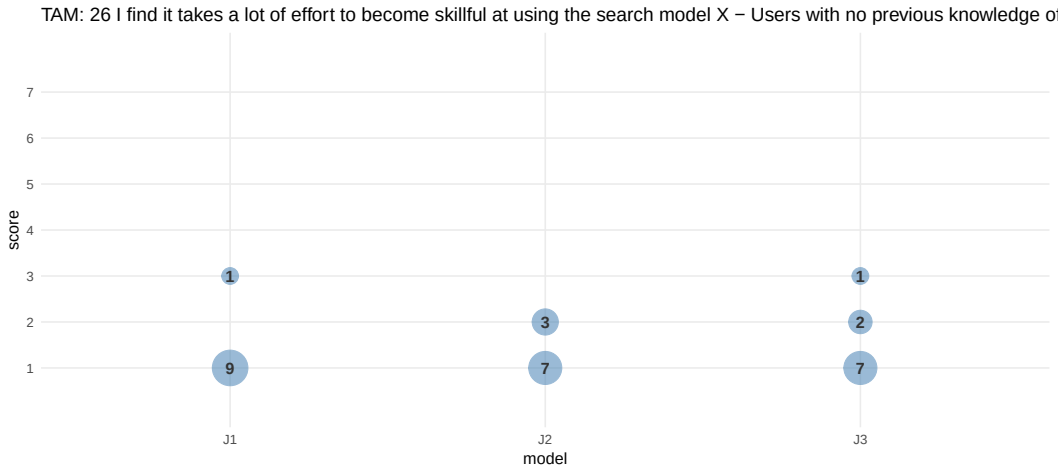



Figure D.73: Response to survey - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models

### D.3.2.15
### TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models

Table D.73: Median and interquartile range of scores per model - TAM: 26 I find it takes a lot of effort to become skillful at using the search model X - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 1 | 0.00 |
| J2 | 1 | 0.75 |
| J3 | 1 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.412$, df=2, p-value=0.4936).

PUC-Rio - Certificação Digital Nº 1721764/CA

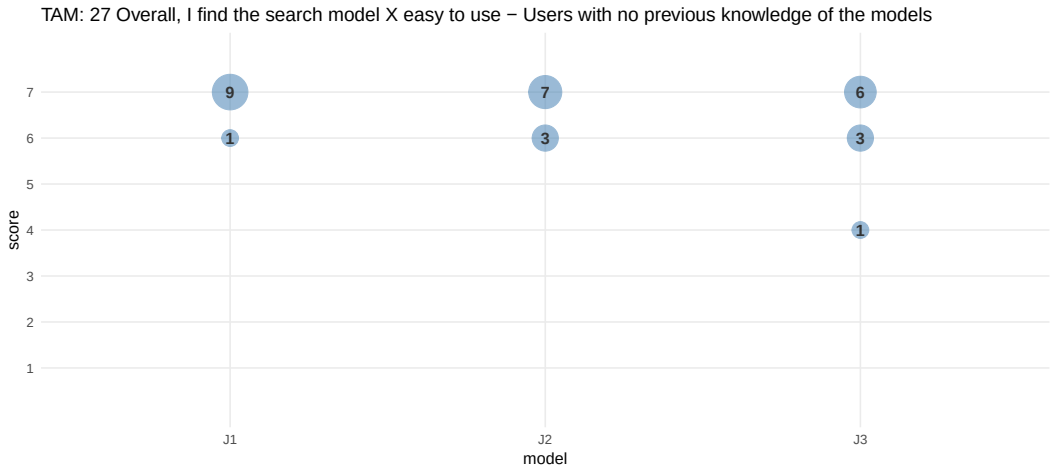



Figure D.74: Response to survey - TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models

### D.3.2.16
### TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models

Table D.74: Median and interquartile range of scores per model - TAM: 27 Overall, I find the search model X easy to use - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.75 |
| J3 | 7 | 1.00 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 2.872$, df=2, p-value=0.2379).

PUC-Rio - Certificação Digital Nº 1721764/CA

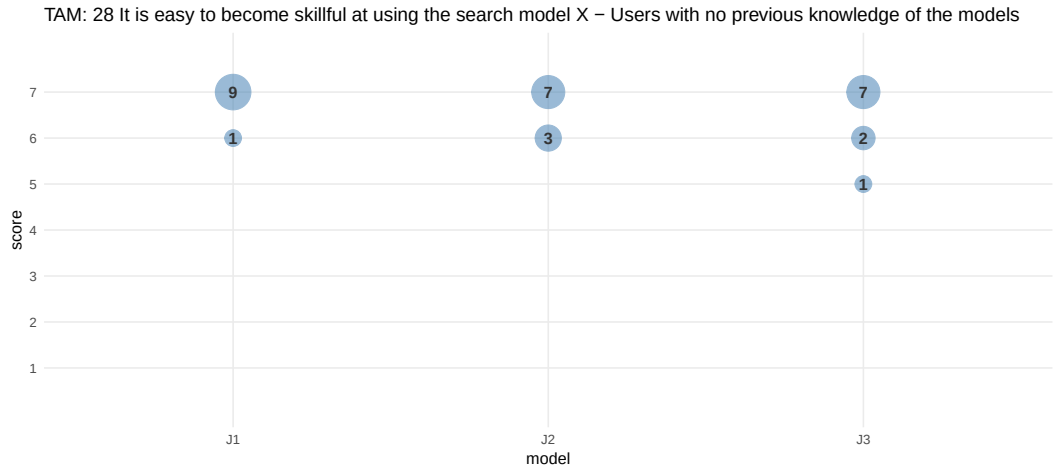



Figure D.75: Response to survey - TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models

### D.3.2.17
### TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models

Table D.75: Median and interquartile range of scores per model - TAM: 28 It is easy to become skillful at using the search model X - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.75 |
| J3 | 7 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.944$, df=2, p-value=0.3784).

PUC-Rio - Certificação Digital Nº 1721764/CA

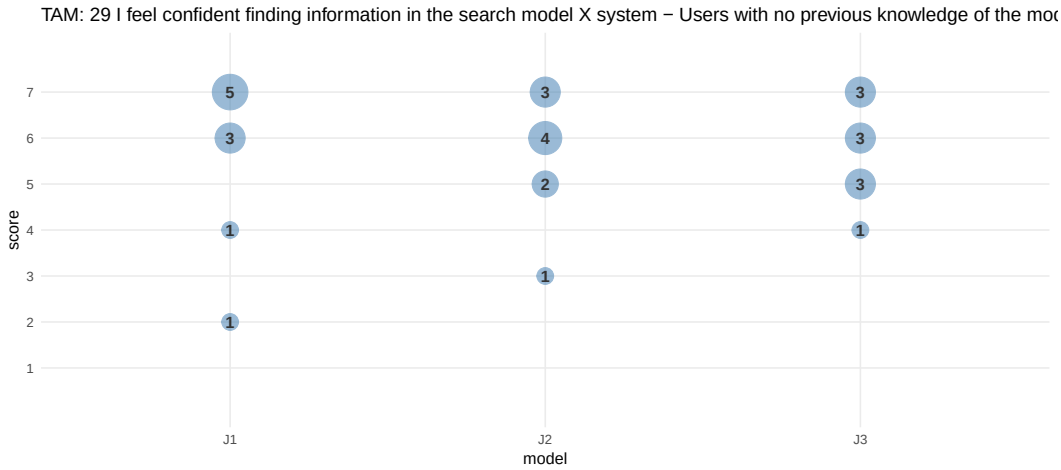



Figure D.76: Response to survey - TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models

### D.3.2.18
### TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models

Table D.76: Median and interquartile range of scores per model - TAM: 29 I feel confident finding information in the search model X system - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 6.5 | 1.00 |
| J2 | 6.0 | 1.50 |
| J3 | 6.0 | 1.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 4.587$, df=5, p-value=0.4683).

PUC-Rio - Certificação Digital Nº 1721764/CA

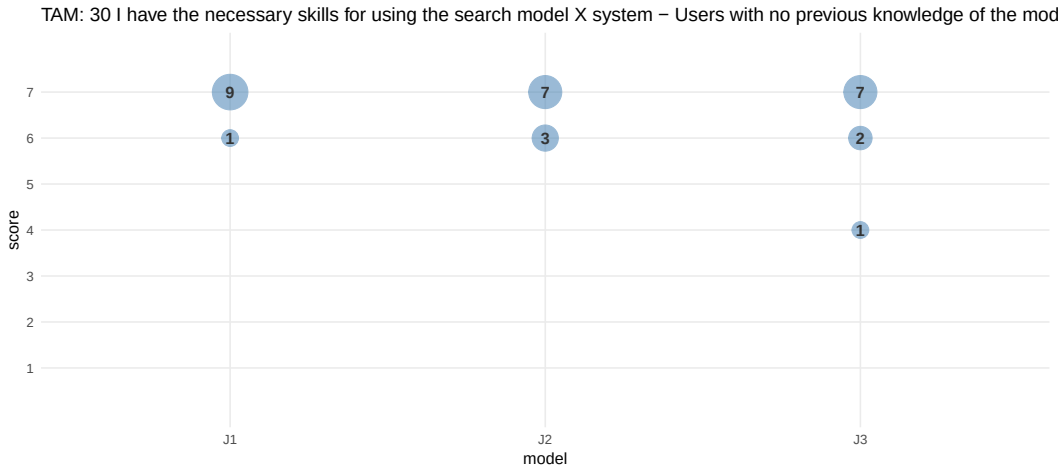



Figure D.77: Response to survey - TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models

### D.3.2.19
### TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models

Table D.77: Median and interquartile range of scores per model - TAM: 30 I have the necessary skills for using the search model X system - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.00 |
| J2 | 7 | 0.75 |
| J3 | 7 | 0.75 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 1.944$, df=2, p-value=0.3784).

PUC-Rio - Certificação Digital Nº 1721764/CA

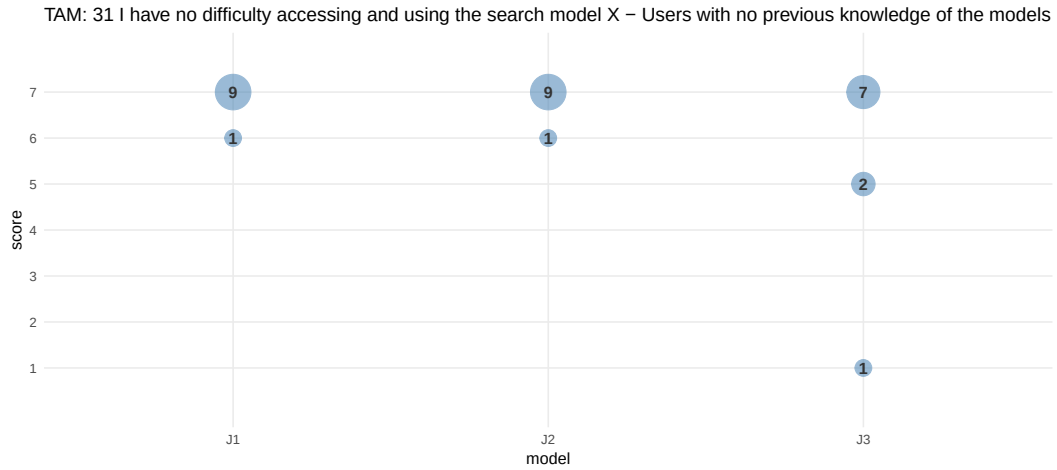



Figure D.78: Response to survey - TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models

### D.3.2.20
### TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models

Table D.78: Median and interquartile range of scores per model - TAM: 31 I have no difficulty accessing and using the search model X - Users with no previous knowledge of the models

| model | median | IQR |
|---|---|---|
| J1 | 7 | 0.0 |
| J2 | 7 | 0.0 |
| J3 | 7 | 1.5 |

We ran a Kruskal-Wallis rank sum test, which showed NO significant difference at $\alpha = 0.05$ level ($\chi^2 = 5.307$, df=3, p-value=0.1506).

PUC-Rio - Certificação Digital Nº 1721764/CA