MASTER'S THESIS IN LIBRARY AND INFORMATION SCIENCE
FACULTY OF LIBRARIANSHIP, INFORMATION, EDUCATION AND IT

"You Have to Use Your Imagination"

Attitudes and Information Practices of Fans Searching the Archive of Our Own

RICKARD NILSFRED

HÖGSKOLAN I BORÅS
VETENSKAP FÖR PROFESSION

© Rickard Nilsfred
This work is licensed under CC BY-ND 4.0.
To view a copy of this licence, visit https://creativecommons.org/licences/by-nd/4.0/

English title: “You have to use your imagination”: Attitudes and information practices of fans searching the Archive of Our Own

Swedish title: “Man måste använda sin fantasi”: Attityder och informationspraktiker hos fans som söker på Archive of Our Own

Author: RICKARD NILSFRED

Completed: 2026

Abstract:

Several studies within the field of fan studies have examined fans’ attitudes to the curated folksonomy (tagging system) and search features of fanfiction archive Archive of Our Own (AO3). However, factors affecting these attitudes have not been studied extensively, and despite the conceptual proximity to the field of human–computer interaction (HCI), existing research employing HCI methods in fan studies is limited.

Using a mixed-methods approach, this exploratory study collects its data from a survey, two interviews, and three condensed contextual inquiry sessions. The survey uses a modified end-user satisfaction instrument to determine attitudes to AO3 search, comparing this to certain types of academic, professional, and fandom-related experience. Weak correlations are found between search satisfaction and familiarity with the tagging system, search satisfaction and experience using other fanwork platforms, and academic or professional experience in IT or LIS and confidence using the tagging system. Contextual and qualitative content analysis suggest fans both appreciate and are confused by AO3’s interface and that AO3 performs well on exploratory searches but falls short on specific lookups. Previous findings on fans’ attitudes to AO3’s tagging system and its effects on search are also corroborated.

This study adds to the literature about attitudes to searching and tagging on AO3 and demonstrates how HCI methods can benefit the fan studies field. Future research might use contextual inquiry or the search satisfaction scale developed here to study AO3 usage, or consider the meaning of these findings for questions in LIS more broadly.

Keywords: Information practices, fan studies, folksonomies, information retrieval, surveys, condensed contextual inquiry, end-user satisfaction, Archive of Our Own

Table of Contents

1 INTRODUCTION 1

1.1 Research context 1

1.2 Research problem 2

1.2.1 Research questions 3

1.3 Purpose and relevance 3

1.4 The structure of this thesis 4

2 BACKGROUND AND (MORE) TERMS 5

2.1 The Archive of Our Own 5

2.2 Curated and uncurated folksonomies 5

2.3 AO3's search systems and interfaces 6

2.3.1 Search syntax and retrieval model 8

2.4 User satisfaction and contextual design 8

3 PREVIOUS RESEARCH 10

3.1 Fan studies: a brief history 10

3.2 FanLIS and fan information practices 11

3.3 Usability, user satisfaction 13

3.4 Usability methods in LIS 14

4 METHODS 16

4.1 Survey 17

4.1.1 Survey design 17

4.1.2 Pilot testing 19

4.1.3 Full-scale survey: recruitment and sampling 24

4.2 Condensed contextual inquiry and interviews 25

4.2.1 Recruitment and sampling 26

4.2.2 Format of the sessions 27

4.2.3 On autoethnography 28

4.3 Data analysis methods 29

4.3.1 Statistical analysis 29

4.3.2 Qualitative content analysis 30

4.3.3 'Condensed' contextual analysis 31

4.4 Research ethics 34

4.4.1 Privacy and pseudonymity 34

4.4.2 Lexical and linguistic considerations 35

5 RESULTS AND ANALYSIS 36

5.1 Quantitative survey results 36

5.1.1 Likert sets: internal reliability and item analysis 36

5.1.2 Summary of responses 39

5.1.3 Correlations 46

5.2 Qualitative content analysis 50

5.3 Quantitative CCI results 55

5.4 Condensed contextual analysis 57

5.4.1 Finding a work by its title 57

5.4.2 Finding a work by description ('lost fic') ... *60*
5.4.3 Finding something (new) to read or recommend ... *60*

6 DISCUSSION ... **62**

6.1 FANS' ATTITUDES TO SEARCH AND TAGGING ... 62
6.1.1 Search satisfaction ... *62*
6.1.2 Perceived strengths of and appreciation for AO3 search ... *62*
6.1.3 Perceived weaknesses of AO3 search ... *64*
6.1.4 The tagging system ... *65*
6.1.5 Tagging confidence ... *67*
6.1.6 Tagging practice, tagging culture ... *68*
6.2 DIFFERENCES BETWEEN ATTITUDES OF 'SPECIALISTS' AND 'NON-SPECIALISTS' ... 69
6.2.1 Weak tagging confidence–specialist experience correlation ... *70*
6.2.2 Possible causes of the weakness of the correlations ... *70*
6.3 SEARCH BEHAVIOUR AND ITS INFLUENCE ON ATTITUDES ... 72
6.3.1 Preferred search methods ... *72*
6.3.2 Search strategies ... *74*
6.3.3 Discoverability issues and breakdowns ... *75*
6.3.4 Effects on attitudes ... *77*
6.4 THE INFLUENCE OF OTHER EXPERIENCE FACTORS ... 78
6.4.1 Caveats to the correlations ... *78*
6.4.2 Possible meanings of the correlations ... *79*
6.5 THE APPLICABILITY OF THE CCI METHOD ... 80
6.6 FANDOM AS A COMMUNITY ... 82

7 CONCLUSIONS ... **85**

7.1 FURTHER RESEARCH ... 86
7.2 LIMITATIONS ... 87

REFERENCES ... **88**

APPENDIX A ADDITIONAL SCREENSHOTS

APPENDIX B SURVEY

APPENDIX C SURVEY INVITATION MESSAGE

APPENDIX D CCI SESSION PLAN

APPENDIX E REACTION CARDS

APPENDIX F BREAKDOWNS

List of Tables

Table 1 *Changes to the Search Satisfaction Likert Scale Items* 22
Table 2 *Changes to the Tagging Confidence Likert Scale Items* 23
Table 3 *Search Satisfaction, Item–Total Correlations* 37
Table 4 *Tagging Confidence, Item–Total Correlations* 38
Table 5 *Tagging Familiarity, Item–Total Correlations* 38
Table 6 *Tagging Confidence (Question 8) Total Scores in Percent* 42
Table 7 *Correlation Table (ρ) for Specialist Experience, Search Satisfaction, Tagging Confidence, and Tagging Familiarity* 47
Table 8 *Correlations (ρ) Between Search Satisfaction and Length of Time Using AO3, Frequency of Using AO3, and Other Platforms Used* 48
Table 9 *Categories and Themes Emerging* 50
Table 10 *Lookup Task Counts* 56
Table 11 *Number of Iterations per Lookup Task* 56
Table 12 *Average Iterations and Perceived Success by Query Type* 57

List of Figures

Figure 2.1 *AO3's Main Menu and Search Bar* 7
Figure 4.1 *Coarse-Grained Illustration of the Steps of Contextual Design* 32
Figure 5.1 *Question 1, Search Methods Used by Participants* 39
Figure 5.2 *Question 2, Other Search Methods Described* 40
Figure 5.3 *Question 3, Preferred Search Method* 40
Figure 5.4 *Distribution of Search Satisfaction (Question 4) Total Scores* 41
Figure 5.5 *Distribution of Tagging Confidence (Question 8) Total Scores* 42
Figure 5.6 *Tagging Familiarity (Question 9), Distribution of Summed Totals* 43
Figure 5.7 *Tagging Familiarity (Question 9), Individual 'Knowledge Items'* 44
Figure 5.8 *Question 11, "When Did You First Start Using AO3?"* 45
Figure 5.9 *Question 12, "How Often Would You Say You Use AO3?"* 45
Figure 5.10 *Question 5, "Have You Used Any Other Platforms to Search for Fanworks?"* 45
Figure 5.11 *Answers to Specialist Experience Matrix (Question 13)* 46
Figure 5.12 *Average Score Totals by Search Method Preference* 49
Figure 5.13 *Median Score Totals by Search Method Preference* 49
Figure 5.14 *Frequency of Main Themes Among Survey Respondents* 55
Figure 5.15 *Sequence Model: Finding a Work by Title* 58
Figure 5.16 *Sequence Model: Finding a Work by Description* 59
Figure 5.17 *Sequence Model: Finding Something to Read or Recommend* 61
Figure A1 *Search Result Page on AO3*
Figure A2 *The Work Search Page on AO3*
Figure A3 *Browse Interface on AO3, Showing the Filter Sidebar*
Figure A4 *Tag Search Result Page on AO3*
Figure A5 *Tag Page for a Canonical Tag on AO3*

Acknowledgements

Some time in the late noughties, my flatmate gave me a book. Fifteen-odd years later, I found this group on Facebook where people talked about fanfiction, and this place called AO3.

I owe a debt of gratitude to the friends I've made since then for their invaluable lessons about writing, fact-finding, and how to survive the ordeal of getting feedback. Special thanks to my pilot testers, who kindly answered all my questions, including the bad ones. And to Dr Rene Nickel-Moore, my silliest serious friend, for brainstorming with me across time zones, reminding me to eat and sleep, and never once suggesting I might be in over my head.

I'd also like to thank my thesis supervisor, Wout Dillen, for his level-headed advice and his patience with me when the data didn't say what I expected and I might have very slightly panicked.

Lastly, I must acknowledge the fine folks of the FanLIS Symposium's Discord server, who have been generous with their advice and encouragement, and who I hope will think I've made a worthy contribution to their (our?) field. I can't believe I get to talk to my academic idols via emoji.

Glossary*

Fan studies The study of fans, fandoms, fanworks, and fan behaviour.

FanLIS The subfield of fan studies where it intersects with library and information science.

Fan An enthusiastic follower of some media, group of people, or famous person. Sometimes restricted to active participants in fan communities. To describe something pertaining to fans, the adjective **fannish** may be used.

Fannish librarian A fan who is also a student or researcher of LIS or works in a library or information profession.

Fandom A community of fans centred around a particular piece of media, fictional universe, character(s), or person(s).

Fanworks Transformative works based on the characters, worlds, or stories of one or more fandoms. Examples include **fanart** and **fanfiction**.

Archive of Our Own (AO3) A large non-profit archive of fanworks, primarily fanfiction, owned by the non-profit Organization for Transformative Works, OTW. Allows tagging of and search for fanworks using a hierarchical, curated folksonomic tagging system.

Canonical tags The name for preferred terms in AO3's curated folksonomy.

Tag wrangler Volunteer tag moderators on AO3, who organise user-entered tags by semantic relationship into a hierarchical system of canonical tags and their synonyms, hypernyms, and hyponyms.

Curated/hybrid folksonomy A compromise between a controlled vocabulary and a free folksonomy, in which users can enter tags freely and human curators or moderators link commonly-used tags with synonymous or related meanings together into a semi-controlled system.

User satisfaction A concept in usability research which measures users' subjective perception of overall success of an information system. Simply put, user satisfaction tells us whether or not users think a system is "good."

Contextual inquiry (CI) A naturalistic method in user studies, whereby users are observed in their normal work or home environment and interviewed about how they solve tasks. When done in short sessions (at most 1-2 hours), possibly remotely, it is called **condensed contextual inquiry (CCI).**

Contextual analysis A set of qualitative methods used to analyse CI data.

---

* All terms are explained more fully in chapters 1 and 2.

1 Introduction

> Anyone who has ever fantasized about an alternate ending to a favorite book or imagined the back story of a minor character in a favorite film has engaged in creating a form of fan fiction.
>
> (Hellekson & Busse, 2014, p. 1)

This is a study about fans. It is also a study about a digital archive: the Archive of Our Own (AO3),[1] an enormous, searchable online collection of fanfiction and other fan-made creative works. How do fans look for information on AO3? Do its search and discovery features work like its users expect? And can we, by looking at prior experiences and observing actual use, explain when or why they do not?

1.1 Research context

In the interdisciplinary field of **fan studies,** a **fan** may be broadly defined as an enthusiastic follower of some media (such as a TV show, book, or film), a music group or sports team, or a famous person (Click & Scott, 2017; Price & Robinson, 2017). Some authors prefer a more participatory definition, restricting the term to only those who actively participate in fan communities (Click & Scott, 2017) – although, as observed by Hellekson and Busse in the quote above, even this is a rather broad definition.

Fan studies also studies **fandoms**, communities of fans interested in particular fictional universes (such as the Marvel Cinematic Universe or J. R. R. Tolkien's Middle Earth), characters (like Captain Kirk or Wonder Woman), or people (such as Taylor Swift). Fandoms are participatory communities where members do not merely consume media but are often also active producers (Hellekson & Busse, 2014; Price & Robinson, 2017), with many creating **fanworks**, transformative works based on the worlds, people, or characters of their fandom. Fanworks can, for instance, take the form of art – **fanart** – or written, fictional works – **fanfiction** (or **fanfic**).

As modern fan communities are in large part an online phenomenon, many fanworks are shared on online platforms. One popular such platform[2] is the non-profit **Archive of Our Own**, with over 10 million registered users and 17 million posted works (Archive of Our Own [AO3], 2026b). An important feature of AO3 is its social tagging system, which has been described as a **hybrid** or **curated folksonomy** (Bullard, 2018; Johnson, 2014). Works on

[1] https://archiveofourown.org

[2] To be clear, AO3 was not designed to be a platform; it was "only" meant to be an archive with limited social features. However, as Stanfill (2024) argues, it can with good reason be treated as a platform in practice.

AO3 are tagged by the users uploading them, and although, like on many other popular platforms, uploaders are able to type and create their own tags freely, in the background, these user tags are organised by a volunteer team of **tag wranglers** into a hierarchical system of **canonical tags** (Bullard, 2018; Gursoy et al., 2018), combining some of the advantages of free tagging with some of the structure and searchability of controlled vocabularies (Fiesler et al., 2016; Johnson, 2014). These tag categories, as well as other metadata, can then be used when searching for fanworks on AO3, making its search and browse features potentially very powerful.

1.2 Research problem

AO3's curated folksonomy, its effectiveness in aiding search in the archive, and fans' attitudes towards it have been studied fairly extensively within the subdivision of fan studies informally known as **fanLIS**, that is, the intersection of fan studies and library and information science. Several studies (e.g. Dalton, 2012; Fiesler et al., 2016; Miller, 2022; Price & Robinson, 2017) indicate that fans' attitudes to AO3's search and tagging systems are, overall, positive; however, this author's experiences as a participant in online fan communities indicate that some users of AO3 remain confused about how to effectively search the archive, or even annoyed at its perceived complexity.

Some existing research focuses, in one way or another, on fans with various kinds of special expertise: Koven Matasy (2013), for instance, studies fans with experience working or studying in LIS-related fields – a group which she calls "**fannish librarians**" – and the participants in Price and Robinson's (2017, 2021) studies are what might be termed "serious leisure experts" – that is, fans who are experienced and knowledgeable in their role as fans, such as AO3 tag wranglers, but not necessarily academic or professional experts. One might expect the first group, fannish librarians, to quite naturally have a better grasp and appreciation for knowledge organisation systems in general than the broader population, and perhaps a higher likelihood of reporting positive experiences with such systems. Similarly, tag wranglers will, of course, have gained a deeper understanding of AO3's tagging system working within it. Another group which might be better equipped to understand an information system of any kind, but which is not mentioned at all in this context in the fan studies literature, consists of those with experience in IT or computer science.

There are studies which do study fans without any specialist experience (see e.g. Miller, 2022; Waugh, 2019); however, they do not focus specifically on that group. Dalton's (2012) study has a mixture of fannish librarians, tag wranglers, and 'non-specialist' fans, and does attempt to compare the groups; however, her 'specialist' groups are fairly small and are only compared 'expertise by expertise' (such as tag wranglers versus non-tag-wranglers), never in aggregate. The existing literature on how attitudes to and usage of

AO3's search and tagging systems may relate to fans' specialist experiences is, therefore, as yet limited.

In addition, there appears to be a general lack of fanLIS research which applies methods from the field of HCI (human–computer interaction) to problems in fan studies, or which investigates fans' actual usage of technological tools from a usability perspective. This is despite the fact that much of fanLIS research concerns fans' usage of and attitudes to platforms.

1.2.1 Research questions

The focus of this study, then, is on fans' attitudes to the search and tagging features of the Archive of Our Own, how fans actually use these features in practice, and what factors might affect this. Specifically, we ask:

RQ1. What are fans' attitudes to the search and tagging systems of AO3?

RQ2. Are there any significant differences between fans with and without specialist experiences (that is, fannish librarians, tag wranglers, and fans with professional or academic IT experience) in their attitudes to these systems?

RQ3. How do fans search AO3 in practice, and does this influence their attitudes to searching on AO3?

RQ4. Do any other experience factors (outlined below) significantly influence fans' attitudes?

In addition, the study also considers in what ways methods from HCI can help study fans' search behaviour on AO3 and shed further light on RQs 1 and 3.

To limit the scope, the other experience factors explicitly studied are limited to:

- Familiarity with AO3's tagging system
- Amount of experience using AO3 (that is, how often and for how long the respondent has used the site), and
- Experience searching for fanworks on other platforms.

As certain parts of the study are qualitative and inductive in nature (see chapter 4), this list is not entirely exhaustive; however, it sets the bounds for the questions asked during data collection and serves as a focus for the analysis.

1.3 Purpose and relevance

The Archive of Our Own is only one website, albeit a very popular one with an interesting and somewhat unique approach to knowledge organisation.[3] In this sense, the result of the present study might be relevant for AO3 itself: usability

[3] That said, it is hardly the only example of a curated folksonomy in the wild. For instance, Bullard's (2018) comparison of AO3 to two other hybrid folksonomies, about which more is said in chapter 3, finds both differences and similarities between the three implementations.

data for certain user groups should be of interest to any web site's maintainers. Broadening our view, however, there is also potential for practical applications in other information retrieval and knowledge organisation systems, including library catalogues (see Holley, 2010; Noorhidawati et al., 2013; and Peterson, 2009; for examples where folksonomic tagging in a library catalogue was explored in practice) and general and specialist search engines.

Furthermore, a better understanding of fans' information practices is relevant to the research in classification and information systems more broadly (Bullard, 2018; Johnson, 2014) as well as to other researchers and professionals in other fields, like education, policy making, and the media industry (Price & Robinson, 2017). Finally, there are potential new avenues that could be opened by introducing new (to the field) methods into the discipline of fan studies, and this, of course, has value in its own right.

1.4 The structure of this thesis

Beyond this introductory chapter, where the research context, questions, and purpose have been introduced, this thesis is structured as follows.

Chapter 2 elaborates on the research context and background by giving an outline of the history behind and current functionality of AO3, as well as defining relevant concepts from the field of human–computer interaction.

Chapter 3 then summarises the existing research relevant to this study, providing a brief history of the field of fan studies and the literature on fans' information practices, user satisfaction, and the use of HCI methods in LIS.

Chapter 4 describes the present study's research methods: an online survey and a set of interviews and condensed contextual inquiry sessions, analysed using statistical methods, qualitative content analysis, and tools from contextual analysis. It concludes with a discussion about some ethical considerations, particularly as they relate to studying fandom.

The results of the data collection and analysis are presented in chapter 5 and contextualised in chapter 6, where each research question is answered by reference to the empirical results, and the relationship of those results to the wider fan studies field are discussed. Chapter 6 then concludes by arguing for the usefulness of condensed contextual inquiry as a fanLIS method.

Finally, chapter 7 sums up the results and discussion and suggests directions for further research.

2 Background and (more) terms

2.1 The Archive of Our Own

The Archive of Our Own (AO3) was launched in 2008 by the American non-profit Organization for Transformative Works (OTW; Fiesler et al., 2016). There were, and still are, several other fanfiction archives, many of which are specific to individual fandoms, and some of which are general. Today, however, AO3 is the single largest fanwork platform, especially among white and Anglophone users (Rouse & Stanfill, 2025; Stanfill, 2024), and it has a central enough place in online fandom that it may be considered essential fandom infrastructure (Stanfill, 2024).

The OTW itself was founded in 2007 in response to a number of controversies in and around fandom and transformative fanworks. Most immediately, commercial fanfiction archive FanLib had been launched in May 2007 in an effort to profit from the free labour of fan-writers, generating intense backlash from fans (Fiesler et al., 2016; Flegel & Roth, 2014; Hellekson, 2009). Soon thereafter, the popular blogging site LiveJournal, which was an important platform for online fandom in the early 2000s (Busse & Hellekson, 2006; Fiesler et al., 2016), implemented a new policy against certain categories of sexual content and, without warning, removed a large number of fan accounts and fanworks in an event that became known as Strike-through[1] (Flegel & Roth, 2014). Fanfiction.net, another large fanfiction archive, had also made unexpected policy changes, and did not allow any adult content at all (Fiesler et al., 2016). These and other similar events sparked a debate about the legal status of fanworks and fans' reliance on commercial platforms and led, ultimately, to the establishment of OTW as a non-profit run by fans, for fans (Organization for Transformative Works, n.d.). With the stated goals to protect the legality of transformative works and to preserve fan culture (Fiesler et al., 2016; Flegel & Roth, 2014), OTW developed AO3 to be an archive belonging to fans themselves – an archive of their own – accepting of essentially any genre of fanworks, including those forbidden on other platforms (Rouse & Stanfill, 2025; Stanfill, 2024), requiring only that it is legal and that certain content warnings and appropriate age ratings are applied (Fiesler et al., 2016). This is important, as fan creators may write about unpopular topics, including violence or taboo sexual practices, within the safety of online pseudonymity (Busse, 2017; Dym & Fiesler, 2020).

2.2 Curated and uncurated folksonomies

The tagging system of AO3 is, as has already been stated, somewhat unique. When a user publishes a work on AO3, they are encouraged to mark it up with

[1] So named because the affected blogs were marked with a line through their names – a strike-through – in links across the site.

tags in seven different categories (some of which are required), including fandoms; characters appearing in the work; central relationship categories; rating; and themes (AO3, n.d., 'Different types of tags' section). Four of these categories allow for the free creation of new tags, while three have short, fixed lists of options to choose from.

Folksonomies, in their most basic form, are indexing systems ungoverned by controlled vocabularies, where index terms are assigned by individual users (Bawden & Robinson, 2022, Chapter 8). Folksonomic or social tagging has become very popular with the rise of the internet, and is now an integral feature on many social media platforms, such as Instagram, X (formerly Twitter), and Tumblr.

The chief strengths of social tagging are its accessibility and flexibility. To effectively use a controlled vocabulary, one first needs to become familiar with the controlled vocabulary in question; a high threshold for many users (Buckland, 2017, p. 122). On the other hand, it is relatively easy for anyone to think of words to describe a known and familiar piece of content: a user of Instagram might declare that his post is about cats, that her picture was taken in Paris, or that the video clip contains sponsored content. A social tagging system also allows its users to describe concepts that the designer of the system did not or could not have anticipated (Bullard, 2018).

However, description is not the only purpose of metadata. The foremost weakness of folksonomic tagging systems becomes apparent when considering the second use of metadata as an aid for search and discovery (Buckland, 2017, Chapter 6). Taggers might use different synonyms for the same concept, use the same words to mean different things, choose terms with different specificity – such as "golden retriever" instead of "dog" – or simply spell words in different ways (Bullard, 2018), which may cause a tag search to leave out items that would have been relevant (weakened recall), include items which are not (weakened precision), or both. These are the trade-offs of social tagging when compared to controlled vocabularies: social tagging decreases difficulty for taggers, but at the expense of findability (Bullard, 2018).

To combine the advantages of controlled vocabularies with the convenience and simplicity of completely uncontrolled folksonomies, then, AO3's volunteer tag wranglers link commonly used tags with similar or related meanings together, creating a hierarchical system of preferred terms – the canonical tags. When users select tags for new works or use AO3's browse interface or advanced work search, canonical tags are shown as auto-complete suggestions in the relevant fields.

2.3 AO3's search systems and interfaces

There are two main systems to search for fanworks on AO3 – **work search** and

browsing – and several ways to find and use each of them. This section briefly describes these systems, with additional screenshots provided in Appendix A.

AO3's **search bar** is placed in the main menu bar at the top of every page (see Figure 2.1). Clicking the 'Search' button brings the user to a search result page like the one shown in Figure A1.

From the search result page, users can click 'Edit Your Search' to go to the advanced search, or work search, page, where many more search options, including sort order, can be found (see Figure A2). The work search page can also be reached by choosing 'Works' in the Search menu (see Figure 2.1). A very similar result page is shown when searching for user-created bookmarks ('Bookmarks' in the Search menu).

AO3's browse feature, meanwhile, allows users to browse through all fanworks which are marked with a given tag from any of its seven tag categories (see section 2.2). There are numerous ways to reach the browse interface, including:

- Choosing 'Tags' in the Browse menu and clicking a tag of interest
- Selecting a fandom category (such as "Books & Literature," "Video Games," or "Theater") in the Fandoms menu, then clicking on a fandom tag of interest in the list shown
- Choosing 'Tags' in the Search menu, locating a tag of interest, and clicking on 'Works' at the top of its tag page
- Clicking on any tag in the metadata of any work anywhere on the site
- Going to any user's page and selecting 'Works' or 'Bookmarks.'

The browse 'result page' – which is, technically, the entire browse interface – includes a **filter panel**, placed in a sidebar on large screens and hidden behind a toggle button on small screens, where users can include or exclude tags, specify sort options, and more. The filter panel also contains a search field which functions, syntax-wise, like the 'main' search bar. An example browse result page is shown in Figure A3.

Figure 2.1

AO3's Main Menu and Search Bar, With the Search Menu Open and a Search Tip Shown

2.3.1 Search syntax and retrieval model

The search bars in the main menu and the filter panel, as well as the Any Field text box on the work search page, all accept Boolean AND, OR, and NOT operators, wildcards, phrase search using quotation marks, and other advanced syntax. When the 'main' search bar has focus, a little pop-up with a randomly chosen search syntax tip is shown below it (see Figure 2.1). A page with fairly detailed search help[2] is also provided, as well as context-sensitive help which can be accessed from the work search interface.

AO3 searches exclusively in the metadata of works, not their full text, and its retrieval model is straight-forwardly Boolean (Ceri et al., 2013, Section 3.2), in contrast to the semantic and personalised search of most generalist search engines today. When searching, browsing, or filtering in tag fields, however, the system includes works tagged with sub-tags of and tags synonymous with any canonical tags which appear in the search terms (AO3, n.d., 'Listings for tags I didn't use' section) – that is, something of a partial semanticity.

This leads to some subtlety in how search terms are handled with respect to different search methods and types of metadata. For instance, a keyword search for "Office Party" – a canonical tag – will find works tagged "Office Party" as well as works with the text string "office party" in their titles, summaries or other metadata fields; it will also find works tagged with the synonymous tag "Workplace Party," but *not* works whose titles or summaries contain the text string "workplace party" (unless they also contain the string or tag "Office Party"). Conversely, a search for "Workplace Party" will *not* find works tagged "Office Party" – but *browsing* for the tag "Workplace Party" will automatically yield results for the tag "Office Party" and all its sub-tags and synonyms. The 'partial semanticity' of AO3's search systems, then, is almost entirely one-way.

2.4 User satisfaction and contextual design

In human–computer interaction, the area of computing science concerned with the interplay of human behaviour and technical systems, the term "**usability**" refers to the quality of an object – for example, a piece of software – of having been designed with its users' psychology and physiology in mind such that it is as easy, efficient, and satisfying as possible for them to use in the contexts within which they use it (Hartson & Pyla, 2012, pp. 6, 17). The term "user-friendliness" is sometimes treated as synonymous, although there have been objections to this term since at least the 1980s (e.g. Rutkowski, 1982, p. 303).

An important concept in HCI is **user satisfaction**, sometimes referred to as user information satisfaction (UIS), computer user satisfaction (CUS), or end-user computing satisfaction (EUCS), among several other synonyms. User satisfaction is a measurement of users' subjective perception of the overall

[2] https://archiveofourown.org/faq/search-and-browse

success of an information system, and instruments measuring user satisfaction usually take the form of Likert-type questionnaires, yielding straight-forward, quantitative data of users' attitudes towards a system (see section 3.3 for more detail on this). This makes user satisfaction instruments popular and convenient tools for quickly assessing (an approximation of) system usability.

Closely related to usability is the notion of **user-centred design** (D. Norman, 1988). User-centred design uses various different methods and principles to ensure that products, including computer systems, are designed with users' needs as the top priority. One of these principles is **contextual design**.

In contrast to many 'traditional' methods for design, development, and user testing, which involve first planning and designing a product, then having users perform specified tasks with that product in a lab setting and evaluating how well they solve problems, how fast they complete tasks, how many mistakes they make along the way, and so on (Nielsen, 1993, Chapter 6), contextual design uses ethnographic methods, observing users in as natural an environment as possible while they perform normal tasks, then using qualitative analysis methods on the observation data, and iterating the process as needed (Beyer & Holtzblatt, 1998; Holtzblatt & Beyer, 2016; Wixon et al., 1990). The central observation method in contextual design, called **contextual inquiry (CI)**, is a specific form of participant observation in which the researcher and participant take on roles such as apprentice and master or novice and expert, with the participant showing the researcher how to do the work the participant does every day, as if instructing a new colleague (Holtzblatt & Beyer, 2016, Chapter 3; Kuniavsky, 2003, Chapter 8), and the researcher occasionally asks the participant for clarifications or explanations, just like an interested apprentice would. Contextual inquiry observations are done over several hours, so a simplified version, **condensed contextual inquiry (CCI)**, has also been developed (Kantner et al., 2003). CCI sessions are limited to 1–2 hours and can be conducted over video call, making it an attractive method for user studies done remotely on IT systems that a participant would have easy access to at home, such as an existing web site, while still allowing the researcher to observe users in the field.

Contextual analysis is the analytic method – or rather, methods – of contextual design. It consists of a 'tool box' of methodological tools, some of which are applied in every contextual analysis process, while others can be used or disregarded depending on the needs of the project (Holtzblatt & Beyer, 2016, Chapter 8). Contextual analysis is a process in multiple steps, starting with inductive thematic analysis of observation data obtained via contextual inquiry, the results of which are then used to construct models of users and their behaviours, which lead to design concepts, storyboards, sketches, and eventually to finished designs for new products – the typical end goal of contextual design. (Contextual analysis will be described further in chapter 4.)

3 Previous research

3.1 Fan studies: a brief history

Fanworks, if one uses the term broadly, have existed for as long as there have been stories to tell further stories about. More narrowly, we may talk about fandoms at least as far back as the Jane Austen and Sherlock Holmes societies in the 1920s and 30s (Hale, 2005; Pugh, 2004), and perhaps even of roots in 15th century continuations of Chaucer's *Canterbury Tales* (Derecho, 2006).

The discipline of fan studies, however, is only a few decades old. When Lichtenberg et al. (1975) wrote about the enduring fandom of the TV series *Star Trek*, it was the first serious[1] English-language treatment of media fandom (Benecchi et al., 2022). Later, in the 1980s, Lamb and Veith (1986/2014) and Russ (1985/2014) discussed the preponderance of women in media fandom and the popularity of fanfiction about male characters in homosexual relationships (so-called "slash fiction") from different feminist perspectives. The conclusions they drew were different: Lamb and Veith chose to focus on the romance of these stories, and how heterosexual women might enjoy imagining male romantic pairings because it removes the patriarchal power aspect and gender roles often inherent in heterosexual romances, while Russ argued that the pornographic aspect was deserving of serious discussion on its own merits, and that women might simply be writing these stories because they enjoy the 'smut.' These essays have been called the earliest examples of what has come to be referred to as "acafandom" – the practice of fan studies research by academics who are themselves fans (*Acafan*, 2025). The feminist, and eventually queer, angle on fan studies inherent in these treatments remains an important part of the discipline; for a very small number of recent examples, see Fiesler et al. (2016), who discuss why the interface design of AO3 can be called feminist; Joyce (2025), who uses feminist theory to study queer men's relationship with the fandoms of Japanese homoerotic media; and Duggan's (2023) review of the literature on trans fans and gender in fandom.

The moment that many authors (e.g. Benecchi et al., 2022; Hellekson & Busse, 2014; Price & Robinson, 2017) point to as the start of the field as an academic discipline was the publications of Henry Jenkins's *Textual Poachers* (1992), Camille Bacon-Smith's *Enterprising Women* (1992), and *The Adoring Audience*, edited by Lisa Lewis (1992). These works were influential in that they did not treat fandom as abnormal, pathological, or childish; rather, they described media fans as a community with particular ways of consuming and interpreting media, social traditions, norms, and artistic creation. The authors

[1] Although this is perhaps not the most respectful term. After all, it might not be entirely fair to claim that the decades of fanzines which Lichtenberg et al. built upon (Coppa, 2006) did not constitute serious practice to its practitioners, simply because it was done as a hobby.

of these texts also continued the autoethnographic tradition – participants in fandom, studying fandom – which had begun with Lamb and Veith (1986/2014) and Russ (1985/2014) and continues to be common to this day (as asserted by Busse, 2017; Jacobs, 2020; and Popova, 2020).

In the late 1990s and the new millennium, fandom increasingly moved from printed fanzines into Usenet groups and online forums and became more globally interconnected (Coppa, 2006). This was also when fan studies grew from its beginnings in cultural and media studies into more of an interdisciplinary field, with scholars like Pugh (2005), Kaplan (2006), and Sandvoss (2007/2014) analysing fan texts themselves as literature, Derecho (2006) attempting to redefine fanfiction without the reference to ownership implicit in terms like "derivative" and "appropriative," and conflicts between fans and copyright holders in the early 2000s (Pugh, 2004; Flegel & Roth, 2014) leading Tushnet (1997, 2007) and Pugh (2005) to discuss the immaterial-rights dimension of fan culture. Meanwhile, Hart et al. (1999) and Kem (2005) analysed the structure of fan information from a LIS perspective and Rebaza (2009), Hellekson (2009), and others began to look more closely at fan information behaviour.

3.2 FanLIS and fan information practices

Information science, then, has been a part of fan studies since shortly after the latter came into being, even though the term "fanLIS" itself was only coined with the eponymous symposium series at City, University of London, which began in 2020 (Robinson & Price, 2025). Arguably, fandom, as Price and Robinson (2017, p. 651) assert, "is, and always has been, rooted in information practices," with even the first academic treatments spending at least some time on fans' particular ways of dealing with information.

A fair amount has been written on the topic of social tagging and other metadata in online fandom, both on Archive of Our Own and other platforms. Johnson (2014) outlines the advantages of social tagging versus other knowledge organisation systems and the history of fanfiction metadata, and then describes three main models of metadata creation and usage on archives and other platforms where fanfiction was shared in the early 2010s. On the one hand, she explains, the blogging platform LiveJournal's tagging system is an uncontrolled folksonomy, although some efforts had, at the time of her writing, been made to implement a rough kind of structure on fan information on the site with, for instance, guidelines about fanfiction tagging best practices. On the other extreme, Johnson names a number of fanfiction archives, such as general archive FanFiction.net, moderated *Doctor Who* archive A Teaspoon and An Open Mind, and *Harry Potter* fanfiction archive Ink Stained Fingers, which all employed controlled vocabularies with various levels of granularity.

In the middle, she places AO3, describing it as a combination of the advantages of both extremes.

Julia Bullard (2018) also compares different types of metadata creation, focusing specifically on hybrid folksonomies where human curators, as opposed to automated systems, review and organise folksonomies. She looks at AO3 as well as LibraryThing, a book database where any user can contribute to the semantic grouping of user tags, and StackOverflow, a question-and-answer forum for programmers, where users who reach a certain threshold of 'reputation' are allowed to participate in tag curation, and finds several differences as well as important similarities between the three. AO3, for example, is unique among the three in that it provides some training for its tag wranglers, and in that its tag structure is hierarchical (Bullard, 2018, pp. 647–648).

Other authors have studied tagging *practices*: Gursoy et al. (2018) use grounded theory techniques to study the meanings of tags on AO3, finding three overall categories: *tags that identify, tags that reflect,* and *tags that express*. Price and Robinson (2021) use a type of network analysis as well as semi-structured interviews to understand tag use on three prominent web sites (AO3, Tumblr, and Etsy) within one particular fandom, and propose a general taxonomy of fan tags, which Hill (2024) tests on a dataset from a different fandom and expands on somewhat. Bourlai (2018) and Kennedy (2024) look into one particular class of tag, namely, commentary-style (or, in the terminology of Gursoy et al., "expressive") tags, which originated on Tumblr and are often used to describe a user's emotional response to some content or simply to include side-note jokes with a post. Bourlai, who studies tag data from Tumblr, finds that fans on that platform use this type of tag frequently (Bourlai, 2018, Section 6.1), whereas Kennedy, who analyses online discussion about how commentary-style tags are used on AO3, notes some fairly heated controversy, with certain users enjoying the practice, while others are annoyed by it and worry that it will negatively affect AO3's tag wranglers and the tagging system as a whole (Kennedy, 2024, para. 4.5, 5.5).

Studies of information practices in fandom more broadly have been made by Kem (2005), Koven Matasy (2013), Hill and Pecoskie (2017), Price and Robinson (2017), Waugh (2019), and Silberstein-Bamford (2024), among many others. Of these, Koven Matasy and Kem specifically study fannish librarians and their perceived place and roles in fandom (in fact, Koven Matasy coins the term), while Waugh focuses on the information behaviours of teenaged fans. Although methodological details vary, most of these studies use some form of survey or interview to collect data; Hill and Pecoskie being a notable exception, instead performing a qualitative inductive analysis on fanfiction sites themselves.

Silberstein-Bamford (2024) describes three AO3 reading paradigms which

emerge from her data: *informed reading* – that is, readers using tags on works as a sort of content declaration, so that they might know immediately whether they will enjoy the work and whether it is worth engaging with – *rhizomatic reading* – that is, using the hyperlinked metadata on one work to find other works – and *reading in multiple*, the particular habit of fans of actively engaging with multiple fanworks at once, switching back and forth between stories which are, perhaps, being published serially, one chapter at a time (Silberstein-Bamford, 2024, pp. 137–142). The two former are of course relevant concepts to a study into how fans search for fanfiction. Also relevant for the present study are the survey study by Dalton (2012) and the pilot study by Miller (2022), both of whom directly investigate searching on AO3 and to what extent users feel that it is aided or affected by AO3's tagging system, and both of whom find that users are, in general, appreciative, especially when comparing it to searching for fanfiction or other reading materials elsewhere (Dalton, 2012, p. 38; Miller, 2022, para. 5.3–6.2).

3.3 Usability, user satisfaction

The concept of usability has been used for a long time to describe and analyse user interfaces. Foundational works in the field include Donald Norman's *The Design of Everyday Things* (originally titled "The *Psychology* of Everyday Things" (1988)) and Jakob Nielsen's *Usability Engineering* (1993).

User satisfaction has been measured and studied since the 1970s (e.g. Maish, 1979; Swanson, 1974), with a number of measuring instruments, most taking the form of Likert-type questionnaires, being developed and tested in the 1980s. One influential user satisfaction measure was that by Bailey and Pearson (1983), which asked respondents to rank 39 qualities in four groups, giving a total of 156 responses per user. Ives et al. (1983) tested Bailey and Pearson's measure for internal reliability and argued for the content validity, predictive validity, and construct validity of the instrument, then, reasoning that the large number of items may result in respondent fatigue, developed a shorter instrument of only 13 items, which they controlled for sufficient correlation with the original measure. The final measure was not included with their report, but is reported in Ryker et al. (1997), who use it to study correlations between users' *expectations* on a system with user satisfaction.

Doll and Torkzadeh (1988), noting that the measure by Ives et al. had been used widely in the intervening years, expressed concerns about its test-retest reliability and discriminant validity, and observed that it was better suited to measuring managers' attitudes to information systems than testing end user satisfaction with interfaces (Doll & Torkzadeh, 1988, pp. 262–263). Against this background, Doll and Torkzadeh developed a 40-item questionnaire, which they tested and analysed for validity and internal reliability and reduced to a 12-item scale for end-user satisfaction.

Many other attempts to measure user satisfaction have been made, including several different scales developed at IBM (J. R. Lewis, 1995) and the intentionally simple, 10-item System Usability Scale (SUS), which was whittled down by Brooke (1996) from a set of 50 candidate items. Microsoft took a somewhat different approach, developing a "desirability toolkit" consisting of two methods – one where testers selected reaction cards labelled with adjectives like "accessible," "trustworthy," or "stressful," and another where testers compared their feelings towards a system with pictures of human faces showing different emotions (Benedek & Miner, 2002). In addition to employing rather different methods, Microsoft's toolkit stands out in that it generates qualitative rather than quantitative satisfaction data.

Tullis and Stetson (2004) provide a brief overview of four user satisfaction questionnaires as well as one of their own design, comparing their results in a usability test of two web sites and finding that results are mostly similar; however, SUS and IBM's CSUQ showed higher reliability at smaller sample sizes than other questionnaires, and Tullis and Stetson's 'quantified' version of Microsoft's Product Reaction Cards was the test which gave the largest variability in responses (Tullis & Stetson, 2004, pp. 6–7).

3.4 Usability methods in LIS

Wixon, Holtzblatt, and Knox (1990) first presented the principles of contextual design in response to a perceived lack of qualitative methods in computing science, then a highly quantitative discipline, and Kantner et al. (2003) proposed the condensed field version of contextual inquiry, CCI, in order to be able to more quickly collect data from more participants who may be spread out geographically.[2] The methods of contextual design have been used many times in the field of library and information science since then: as early as 2004, for example, Neville (2004) documented the design of a new acquisitions system using contextual inquiry, and Large et al. (2004) used contextual inquiry as one of three design frameworks to create web portal prototypes in a team consisting of adults and elementary school students, observing that these methods allow children as young as age 8–9 to collaborate fairly effectively with adults and finding a number of different information needs that children expected web portals to fulfil (Large et al., 2004, pp. 1152–1153). Smith (2016) performed a CI study on a rather different user group: so-called ethical hackers, IT experts who test the security of systems using the methods of cyber-criminals. Smith's data shows several patterns in tools and strategies, as well as attitudes to the morally grey areas in which ethical hackers operate.

LIS studies incorporating other usability methods include Dinneen et al. (2018), who test different simplified LCSH (Library of Congress Subject

[2] The advantages of this approach became all the more obvious with the COVID-19 pandemic; see for example Donnelly et al. (2021).

Headings) hierarchies comparatively and show that simplifying the LCSH tree in these ways improved users' performance on various quantitative metrics; Noorhidawati et al. (2013), who use both a survey and a standardised usability inventory to evaluate the usability of a social tagging system in a digital library; and Bar-Ilan et al. (2012), who measure recall and user satisfaction in four different information retrieval interfaces, and find, firstly, that expert and regular-user assessment of what results are considered relevant may differ; and secondly, that user satisfaction is not necessarily highest in the interface with the best recall (Bar-Ilan et al., 2012, pp. 748, 752).

In fan studies, meanwhile, user studies are few and far between. Despite being a field so centred on the information practices of a specific group, current fanLIS practice is to study systems and platforms using interviews (e.g. Fiesler et al., 2016), surveys (Rouse & Stanfill, 2025), and sometimes data mining and content analysis techniques (e.g. Bourlai, 2018; Yin et al., 2017). Wagenaar's (2024) study comes closer to the mark in that it uses observational methods to study information behaviours on a specific platform (Discord), but it falls short of contextual inquiry and does not purport to describe users' relationship with the actual interface.

4 Methods

As seen in the previous sections, existing research into fan studies and fans' attitudes to and usage of information employs a number of different methods. Interviews and surveys are common (Benecchi et al., 2022); Koven Matasy (2013), Miller (2022), and Price & Robinson (2017) being a few examples of the former, and Dalton (2012), Silberstein-Bamford (2024), Waugh (2019) and Koven Matasy of the latter. These methods are also common within the wider field of information practices research (Given et al., 2023, Section 5.3; for examples, see e.g. Kasapović, 2025; Pilerot & Lindberg, 2024).

Surveys and interviews are, of course, quite natural choices for studies seeking to determine attitudes and investigate the reasons behind them. Closed-question surveys and structured interviews are appropriate for gathering quantitative data which allows comparisons between groups, and protocols with more open-ended questions allow for richer, qualitative data from a possibly smaller sample (Hank et al., 2017). Consequently, the present study uses a mixed-methods approach and bases its data collection chiefly on a self-administered online survey containing both closed-ended and open-ended questions. A small number of condensed contextual inquiry (CCI) sessions were conducted after the survey was closed in order to collect further qualitative data and explore this HCI method in a fan studies/fanLIS research context. The data thus collected was analysed using statistical methods, conventional content analysis, and an abridged version of contextual analysis.

One possible, but rejected, avenue of investigation into differences between groups of fans might have been to replicate an existing study, which explores the attitudes of one group (say, fannish librarians or other 'information specialists'), on a sample of subjects from the other. However, many potentially suitable studies pose problems with this approach. Price and Robinson's (2017) study, for example, uses modified Delphi interviews, a method which assumes all participants are, in some sense, experts (Luo & Wildemuth, 2017). Their definition of 'expert' also only covers a small part of what this study means by "specialist experience." Miller (2022), meanwhile, focuses on fans between the ages of 18 and 23; thus, while a replication could perhaps have used her sample as the 'non-specialist' group, it is not impossible that their ages might have confounding effects. Her sample was also very small: only ten subjects. Finally, Dalton's (2012) study, although perhaps the closest contender, is now fourteen years old, which means that the information systems it relates to may have changed considerably since it was published, making comparisons with any new sample difficult.

In the following sections, the data collection and analysis methods summarised above will be described in more detail, starting with the two data collection methods – the survey and CCI sessions – and continuing with the three data

analysis methods – statistical analysis of the quantitative data, content analysis of the open-ended survey questions and interview transcripts, and contextual analysis of the CCI sessions. Finally, some ethical concerns particular to fan studies will be discussed.

4.1 Survey

4.1.1 Survey design

The survey in this study was administered through Survey&Report hosted by Sunet Survey, and looked to the questions in Dalton (2012) as a starting point. However, not all of Dalton's questions were relevant for our research problem, and as Dalton noted a measure of respondent fatigue towards the end of her questionnaire (Dalton, 2012, p. 53), decreasing the number of questions, as well as introducing a greater variety of question formats, was desirable (Bryman et al., 2021, Chapter 10; Hank et al., 2017).

Search satisfaction scale

To measure how happy respondents are with searching on AO3, a search satisfaction Likert scale was designed. The intent of this scale was to be a single, overall measurement of how satisfied respondents are with AO3's search features, which could easily be compared to other numerical variables.

Several existing user satisfaction measures were considered as a basis for the scale, and while none was a perfect fit, the System Usability Scale, SUS (Brooke, 1996) was chosen because of its straight-forwardness and small number of items. At the outset, an attempt was made to change the items as little as possible, only substituting "the search interface" for the original's "the system" as well as minimal rewording to fit the case where the respondent has already used the system many times[1] – for instance, changing item 1 of SUS, "I think that I would like to use this system frequently," to "I like to use the search interface frequently." However, it was quickly found that some items, such as item 7 of SUS, "I found the various functions in this system were well integrated," were difficult to translate to our situation and might be confusing to respondents. Furthermore, pilot testing (see below) indicated that some items were perceived as ambiguous, irrelevant, or even irritating, and that questions on the relevance and/or precision of the search results, a dimension not covered by SUS, were needed. Therefore, for the final search satisfaction scale, some SUS items were removed and new items, inspired by the *Information products* dimension of the short-form measure by Ives et al. (1983), as described in Ryker et al. (1997), were added. The result was a set of 14 items.

On the matter of validity, the advantage of using an existing survey protocol is, of course, that it may already be well-tested and validated (Bryman et al., 2021, Chapter 11; Given et al., 2023, Section 5.2.3; Hank et al., 2017). As has

[1] SUS was designed to evaluate a system after having tried it out for the first time.

already been mentioned, the *Search satisfaction* items were based on scale items from existing, generic end user satisfaction scales, which have been tested and validated in a number of different studies (see section 3.3 for examples of this). Thus, it is reasonable to conclude that the final *Search satisfaction* scale does measure aspects of user satisfaction related to the system in focus, implying a promising construct validity. Face validity was chiefly assessed by the present researcher as the scale was constructed; however, feedback from pilot testers and the changes to which it led (see section 4.1.2) are also indications that the scale items actually measure what they were intended to measure.

Tagging confidence and familiarity

To measure respondents' confidence with and knowledge of AO3's tagging system, two smaller sets of Likert items were included. The *Tagging confidence* set, which was similar to the *Search satisfaction* scale in form, included eight items and was intended to focus on the respondent's level of confidence and comfort with applying tags to works on AO3. Since it was not intended to measure user satisfaction, it was not heavily based on a user satisfaction instrument, although questions for SUS were used as inspiration. This Likert set was only shown if the participant answered "yes" to the question, "Have you posted any works to AO3 yourself?"

The question of validity is a little more complicated for *Tagging confidence* than for *Search satisfaction*, as the former does not have the latter's strong connection to existing measures. Again, face validity of the final *Tagging confidence* set (see Results of the pilot test, section 4.1.2) was deemed fair by the researcher, and the inclusion of a pilot test meant that some initial validity 'risks,' such as confusing wording and items uncorrelated to the rest of the set, could be mitigated. However, to get a reliable (!) assessment of the validity of this Likert set, more testing than could be performed in the present study would likely have been required.

General knowledge about the tagging system was roughly approximated using a set of four items, which respondents were asked to rate on a four-point scale with the options, *I don't know this at all*, *I've heard of it*, *I think I understand this reasonably well*, and *I am very familiar*. The topics referred to in these items were, respectively,

1. The hierarchy of tags on AO3
2. The symbols used in so-called 'relationship tags'
3. What canonical and non-canonical tags are, and
4. The popular genres and tropes of the respondent's fandom.

The tag hierarchy and the concept of canonical tags were explained in sections 2.2 and 2.3. Relationship tags are a category of AO3 tags which describe all those relationships between characters appearing in a work which the author

deems relevant. They are made up of the names of two or more characters, separated by a symbol denoting the nature of their relationship in the tagged work, with the symbol "&" denoting a friendship, familial relationship, or other platonic relationship, and "/" denoting a romantic or sexual relationship.[2]

Items 1 through 3 cover important features of how AO3's tagging system works in general, and items 2 and 4 cover the semantic meanings of a large portion of the tags which might be encountered in a given fandom. Therefore, these four topics were determined to be a suitable estimate of the type of overall tagging system familiarity which might affect a fan's use of AO3's search features.

Specialist experience

To assess whether respondents had any potentially-relevant specialist knowledge, a set of questions on academic and professional experience was included. Initially, these questions only considered whether the respondent was a fannish librarian or had ever been a tag wrangler; however, it was expanded after pilot testing to also include questions on work, studies, or research in IT and related fields.

Finally, questions on participants' experience using AO3 and their preferred search methods were included, as well as two open-ended questions where respondents could develop or nuance their answers about searching and tagging, respectively. No demographic questions (age, gender, and so on) were included, as this was beyond the scope of the present study.

When the survey had been closed, responses were downloaded in .xlsx format for analysis in Microsoft Excel.

4.1.2 Pilot testing

In order to identify any gaps, biases, or confusing language (Bryman et al., 2021, Chapter 11; Choemprayong & Wildemuth, 2017; Hank et al., 2017; Ruel et al., 2016), the survey was pilot tested on a convenience sample of eight online fans familiar to the researcher, who all use AO3. The pilot survey was opened on March 1, 2026, and closed on March 5 after all testers had responded. Before opening the survey, testers were instructed to report the time it took them to complete it, as well as any issues with wording, apparent bias, or the general layout of the survey web page, and whether, at any point, they perceived the questions as confusing or uninteresting. Reported response times were then averaged in order to be able to include an approximate 'time to complete' in the survey invitation, and responses and other feedback analysed.

[2] Popular relationship tags from well-known fandoms include "James T. Kirk/Spock" for a romantic or sexual relationship between the characters Kirk and Spock of Star Trek, and "Peter Parker & Tony Stark" for a friendly or familial relationship between Spiderman and Iron Man of the Marvel Cinematic Universe. A work can have zero, one, or several relationship tags, and relationships is one of the tag categories in which users can create new tags freely.

Internal reliability of the Likert sets

When designing Likert scales, it is good practice to statistically test the results for internal reliability – that is, the ability of the set of items to all 'measure the same thing.' This is done by, on the one hand, calculating a measure of the overall internal consistency of the item set over this response set, and on the other, by correlating each item's responses against the rest of the item set (Bikos, 2024, Chapter 7; Choemprayong & Wildemuth, 2017). If an item's correlation with the other items is weak, one can reasonably question whether this item measures the same concept as the others. If the overall internal consistency of the item set is low or many items correlate weakly with the total set, the response set as a whole might not be measuring a single concept, and so, it may not be justified to use the Likert set as a combined Likert *scale*.

To compare each item to the rest of the set, the most straight-forward solution is to correlate the item's score against a total sum of *all* item scores. However, this sum includes the score for the item itself, which may inflate the correlation coefficient, as an item will always correlate perfectly with itself. Therefore, the method used here considers the total sum *with the item's contribution removed*. This ensures that when we check item–total correlations, we are checking items against *only* the rest of the set.

To assess the internal reliability of our Likert sets, the procedure outlined by Lynette Bikos (2024, Chapter 7) was used as a guide, with the modification that calculations were made in Microsoft Excel instead of R. This procedure, as applied to our case, is, in short:

1. Reverse-code any negatively-worded scale items.[3]
2. Find Cronbach's alpha (α) for this normalised set of item scores.
3. Calculate the total score for each respondent by summing up the normalised item scores.
4. For each item, calculate the total score with the item's contribution removed. (This is simply the total from point 3 minus the item's score.)
5. Calculate the correlations between each scale item score and the total score with the scale item's contribution removed, using Pearson's *r*.
6. Given the α and *r* values obtained, consider removing or changing scale items with low *r* values and repeating steps 1–5.

Cronbach's α and Pearson's *r* both yield numbers between –1 and 1 inclusive,[4] where values closer to ±1 are better (stronger internal consistency or correlation). Bikos does not provide a firm cut-off for what is considered an acceptable item–total correlation, but implies (Bikos, 2024, Section 7.4.2) that

[3] This was done following the instructions for SUS (Brooke, 1996), that is, positive items were given a score of (*response value* – 1) and negative items a score of (5 – *response value*).
[4] Cronbach's α is often claimed to range from 0 to 1, however, it may yield negative numbers as well; for example, when negatively-worded items are not reverse-coded (Cho & Kim, 2015).

items with *r*-values on the order of .4 are acceptable. Some authors accept even lower values (Chin et al., 2002, for example, use .3).

As for Cronbach's α, opinions differ on what constitutes a sufficiently high value, with many authors advocating for .7 as a cut-off (Schmitt, 1996), while others require at least .8 or accept values as low as .5 or .6 (Cho & Kim, 2015; Choemprayong & Wildemuth, 2017, p. 297). That said, Schmitt (1996) and Cho and Kim (2015) emphasise that there is no empirically proven "magic number" to aim for, and recommend that authors take into account factors beyond the numerical when judging the internal reliability of Likert scales.

The use of parametric statistics

It is worth noting that Bikos suggests the use of Pearson's *r*, which is normally only advised for interval or ratio-level data (Bryman et al., 2021, Chapter 15; Wildemuth, 2017a), on Likert item scores, which are ordinal variables.[5] This is a fairly common, albeit controversial, practice (see e.g. Carifio & Perla, 2008; Jamieson, 2004; G. Norman, 2010; Sullivan & Artino, 2013). While we will not go into the full mathematical argument here, Pearson's *r* is a parametric statistical method, which means that it makes some statistical assumptions on the data that make it unsuitable for non-interval variables. Instead, for ordinal data, methods like Spearman's rho (ρ) and Kendall's tau (τ) are recommended (Bryman et al., 2021, Chapter 15; Wildemuth, 2017a).

However, authors such as Geoff Norman (2010) have reviewed studies 'misusing' parametric statistics in this way and found parametric tests to be very robust, giving overall accurate answers even when statistical assumptions on the data are violated. In the interest of rigour, in this study, both Pearson's *r* and Spearman's ρ were calculated for the item–total correlations with the item's contribution *included* and the correlation coefficients were compared. For most items, the differences were very small, so during the pilot test phase, Pearson's *r* was deemed sufficient for the item–total correlations with the item's contribution removed. As Pearson's *r* requires fewer steps to calculate than Spearman's ρ in Microsoft Excel, this simplified the statistical calculations somewhat.

Results of the pilot test

The *Search satisfaction* Likert matrix was found to have an α of .66, which is considered questionable or, at the most generous end, just about satisfactory (Cho & Kim, 2015, p. 217). Item–total correlations were found to be very low ($\leq .1$) on items 2 and 10 and low ($< .2$) on items 4 and 6. Item 10 also showed very low variability, with seven out of eight respondents answering "Disagree." With items 2 and 10 removed, the α value improved to .71. As outlined above, this would be deemed satisfactory by many, if not most, authors.

[5] Although there is debate on this point (Bryman et al., 2021, Section 15.4; G. Norman, 2010).

A similar analysis performed on the *Tagging confidence* Likert matrix gave even lower values: α = .27 and an extremely low (near zero) item–total correlation for item 1. With item 1 removed, the α improved to .34, which is still well below acceptable. Now, Schmitt (1996), Rickards et al. (2012), and Choemprayong and Wildemuth (2017) point out that Cronbach's α is sensitive to the number of items in a scale and may give a lower result simply because the scale items are few. However, even taking this into account, .34 is a very low number. Furthermore, items 4 and 6 showed very little variability. This indicated that the *Tagging confidence* set needed to be reworked.

Table 1

Changes to the Search Satisfaction Likert Scale Items

Pilot Test Version of Item	Final Version of Item
I enjoy using the search interface.	I enjoy using the search interface.
I do not feel like I am using the search interface to its fullest potential.	
	The search interface is unnecessarily complex.
The search interface is easy to use.	The search interface is easy to use.
I would need (or have needed) technical support to learn how to use the search interface.	I needed (or would need) to read the FAQ or ask a friend before I could understand the search interface.
The various functions in the search interface suit my needs very well.	The various functions in the search interface suit my needs very well.
I don't understand what some of the functions in the search interface are for.	It is unclear what some of the functions in the search interface are for.
I would imagine that most people would learn to use the search interface very quickly.	I think that most people would learn to use the search interface quickly.
The search interface is very cumbersome to use.	The search interface is very cumbersome to use.
I feel very confident using the search interface.	I feel confident using the search interface.
I needed to learn a lot of things before I could get going using the search interface.	
	I'm not sure the search interface was made with people like me in mind.
	The search results are usually precise and relevant.
	I think there is too much inconsistency in the search interface.
	It's hard to get the search results I need.
	Overall, I am satisfied with the search interface.

Pilot tester feedback indicated, as a whole, that testers found the questions sufficiently varied and few to avoid survey fatigue, the survey interface easy to use, and the instructions helpful. However, feedback also suggested an apparent bias in the wording of the *Search satisfaction* and *Tagging confidence* sets. Most concerningly, one tester had the impression that the survey was not looking for their subjective feelings of difficulty with the system. Another tester summarised their impression as, "the creator of the survey [thinks] this search function is easy," and "if I don't agree […] it's because I don't know what I'm doing." Several testers also expressed frustration with AO3's search results, about which there were no questions at the time. This indicated an even stronger need to rework both of these Likert scales.

The changes to the items of both Likert scales are shown in Table 1, above, and Table 2. Greyed-out items in the tables were left entirely or essentially unchanged.

Table 2

Changes to the Tagging Confidence Likert Scale Items

Pilot Test Version of Item	Final Version of Item
Feel that you are pretty good at tagging so that readers can find your works?	I feel confident at choosing tags for my works.
Understand the purposes of the different tag types?	I feel like I understand the purposes of the different tag types (fandom, rating, relationships, and so on).
Find that deciding what tags to use is very confusing?	Deciding what tags to use is difficult.
Think that putting tags on works is kind of pointless?	I'm not sure how important tags are to get my works seen.
Know where to go if you need more information about tagging?	I know where to go if I need more information about tagging.
Think that you would need more instruction in order to tag works well?	I still think I need more practice or instruction in order to tag works well.
	There is a high threshold for new users to learn how to tag.
	Overall, I like the tagging system.

Note. In the pilot, the prompt for this Likert set read, "When you choose tags for your works, do you…" For the full-scale survey, this was changed to, "[Considering when you choose tags for your works,] how do you feel about the following statements?"

The *Tagging familiarity* set (see section 4.1.1) was also added at this time.

Given the amount of changes prompted by the pilot test, a second round of testing to evaluate the modified survey instrument would have been warranted (Choemprayong & Wildemuth, 2017; Rickards et al., 2012), especially to improve the *Tagging confidence* set, evaluate the *Tagging familiarity* questions, and to further attempt to assess validity. However, due to the limited time available, this step was omitted.

The final set of questions in the full-scale survey is included in Appendix B.

4.1.3 Full-scale survey: recruitment and sampling

The full-scale survey was open from March 9 to March 19, 2026. Participants were recruited from fandom forums on Reddit and Facebook where the author is a member. Initially, the invitation (see Appendix C) was posted in one subreddit (Reddit forum). When the rate of new responses became very low, rather than repeat the invitation in the same forum and risk being accused of 'spamming,' invitations were posted in another subreddit and a Facebook group as well. All three forums are centred on the Good Omens fandom.[6]

It is often recommended that pilot testers are excluded from the full-scale survey due to the risk of test-retest effects (Bryman et al., 2021, Chapter 11; Hank et al., 2017). However, the strength of these effects varies depending on the nature of the questions and, of course, to what extent the survey is changed after the pilot test. It is also important to note that this survey was completely anonymous and open to the public, and thus, even if pilot testers had been asked not to take the modified survey, there would have been no way to reliably ensure that they all followed this instruction. Given this fact as well as the nature of the questions (mainly focusing on attitudes and experiences, rather than assessing knowledge or testing skill), the amount of changes made after the pilot test, and the fairly small pool from which respondents were recruited (Ruel et al., 2016), pilot testers were allowed to take the full-scale survey, if they wished. The risks mentioned above were clearly explained to the pilot testers when they were informed of this.

Sample size and representativity

The recruitment pool described above is somewhat narrow, and this naturally raises questions about homogeneity of the sample, selection bias, and representativity (Bryman et al., 2021, Chapter 8). AO3 has over ten million registered users (Archive of Our Own, 2026b), and most of its features – including, relevantly, the search features – are open to unregistered visitors as well, with official statistics showing over a hundred million total daily page views in 2025 (Archive of Our Own, 2026a). If the total population of a study

[6] The fandom surrounding the novel *Good Omens: The Nice and Accurate Prophecies of Agnes Nutter, Witch* by Neil Gaiman and Terry Pratchett, first published in 1990; as well as the 2019 TV series *Good Omens*, the 2014 BBC radio series of the same name, and other adaptations.

like the present one were to be defined as all users of AO3, this would, thus, make a truly representative sample large enough to be unmanageable, at least for a project with the present one's limitations on time and resources; not to mention the complications involved in recruiting participants from a sufficient variety of demographics.

Restricting samples to a single fandom with which the researcher is familiar is, on the other hand, not unusual in fan studies (Rouse & Stanfill, 2025); see, for example, Price and Robinson (2021), who studied only a fairly small subset of the X-Men fandom. Unfortunately, finding reliable numbers for the size of the segment of AO3's users, registered and unregistered, who are active in a given fandom, is not trivial. AO3's people search feature[7] allows for listing users who have posted works tagged with a given fandom, but this does not, of course, include readers who do not themselves post works. A people search for users who have posted works with a Good Omens fandom tag[8] yielded well over 20 000 results as of 21 April 2026.

As a final attempt to define the total population, one might take the subreddits and Facebook group where invitations were posted. Even determining the size of this population, however, is difficult: Reddit only displays approximate numbers for "weekly visitors" and "weekly contributions" to subreddits, and the subreddits this study recruited from are both open to the public, so finding the total number of recurring readers or posters is practically impossible. This makes it very difficult to determine how large a representative sample of any of these populations would need to be.

However, as a master's thesis project, it is perhaps both natural and acceptable for this study to lean towards the exploratory rather than the fully generalisable. Thus, the sample used here, while small and something of a convenience sample, may still shed some light on the research questions and point to areas of potential further study, even though it is not representative.

4.2 Condensed contextual inquiry and interviews

One problem with survey-based studies is that of recall bias. Survey questions about specific behaviours run the risk of yielding invalid data, simply because people do not always accurately remember their experiences (Donnelly et al., 2021; Hartson & Pyla, 2012, Chapter 3; Schwartz, 2019).

To mitigate this effect, especially in the investigation of RQ3, a version of condensed contextual inquiry (Kantner et al., 2003) was used. CCI was chosen

[7] https://archiveofourown.org/people/search

[8] There are several, including one for the original novel and one for each of five different adaptations (TV, radio, graphic novel, musical theatre, and an abandoned movie concept). The Fandom field of AO3's People Search uses Boolean AND, so to count all users posting in *one or more* of these 'sub-fandoms,' a combination of several searches must be performed.

over field versions of classic user studies because of the more naturalistic properties of contextual inquiry (Crystal & Wildemuth, 2017b; Hartson & Pyla, 2012, Chapter 3) and the ability of CI to capture problems of both usability and utility (Kronberg, 2020, Chapter 6), that is, problems with the interface as well as features which are actually missing (Hartson & Pyla, 2012, pp. 6, 16–17). While CCI is less suitable for measuring certain performance metrics (Rosenbaum & Kantner, 2007), these data are less relevant for the present study than the rich, qualitative data that a CCI study can provide.

4.2.1 Recruitment and sampling

Participants were recruited via a message at the end of the survey, which invited respondents to contact the researcher via e-mail. To reduce the effort required for this step, a simple message template was included in the invitation (see Appendix B), and some respondents made use of this. However, the additional steps involved in writing and sending an e-mail, as compared to entering one's e-mail address in a form, say, undoubtedly limited the pool of potential respondents. This risk was judged acceptable, since other solutions would have necessitated the collection of personal data (such as e-mail addresses) in the survey, which would have complicated participant recruitment and survey administration due to the University's policy on personal data management in student projects.

In total, six survey respondents reported their interest in the follow-up CCI study. Of these, two participants were interviewed and observed over video call. Sessions were recorded using Zoom's built-in recording functionality, and recordings were transcribed soon as practicable after each session was over using the automatic transcription feature of Microsoft Word 365, after which transcriptions were manually double-checked for errors.

To augment the sample, the researcher then performed the same lookup tasks (see section 4.2.2) as the participants, taking notes about the steps taken, any problems encountered, and so on (for further discussion about this choice, see section 4.2.3 below). This data was included in the contextual analysis described in section 4.3.3. Thus, this phase of the study had two interview participants and three CCI participants.

Sampling strategy and representativity

While quantitative studies tend to be concerned with representativity, CCI is a qualitative method, and as such, a purposive sampling method is arguably more appropriate. The ideal sampling method for the CCI phase would have been a maximum variation strategy (Bryman et al., 2021, Section 17.3), whereby participants are selected based on specialist and fandom experience and scores on *Search satisfaction* and *Tagging confidence* such that search strategies and other usage patterns could be compared between fans with high satisfaction and fans with low satisfaction, fans with more or less experience, and so on.

However, the time available limited the number of sessions possible, and the ability to select and group participants based on these factors was impeded by the lack of connection between CCI volunteers and their survey responses, which, as noted above, was a necessary consequence of the need to avoid the collection of any personal data in the survey. Thus, the sampling method for this phase was, in actuality, simply a random selection between the survey respondents who had volunteered to participate.

In any case, the data collected here is not intended to produce generalisable conclusions all on its own, but rather to act as a case study of the application of CCI to a fanLIS research problem, to augment and nuance the findings of the survey, and to suggest directions for further research.

4.2.2 Format of the sessions

The CCI sessions with the two 'external' participants were about an hour long and began with a short interview, wherein a small number of questions from the survey were repeated. This was followed by the main observation session, during which the participant was asked to show the researcher how to perform a number of lookup tasks while sharing their screen, and to explain the process as if the researcher had never used AO3 before.[9] Participants were encouraged to show lookups that they might normally be interested in, but suggestions for tasks were also provided by the researcher.

After this, any follow-up questions noted down during the observation session were discussed and a card-choosing task inspired by Microsoft's Product Reaction Cards (Benedek & Miner, 2002; Tullis & Stetson, 2004), was administered using an online card sorting tool.[10] During this task, the participant was presented with a large number of 'adjective cards' (see Appendix E) and invited to arrange as many of these as they wished into categories of their choosing in a way which described their feelings about the session, after which the chosen arrangement was explained to and discussed with the researcher. Finally, the participant was invited to share any feedback they may have, and the session was concluded.

A plan for these sessions (see Appendix D) was constructed using the guides in Kuniavsky (2003, Chapter 8) and Holtzblatt and Beyer (2016, Chapter 3), as well as taking advice from Hartson and Pyla (2012, Chapter 3), Kantner et al. (2003), Kronberg (2020, Chapter 4) and Rosenbaum and Kantner (2007). Interview questions were chosen to place participants in either the 'specialist experience' or the 'no specialist experience' group, and the follow-up questions as well as lookup tasks suggested by the interviewer were informed by the quantitative and qualitative results of the survey. In practice, the researcher offered both interviewees one lookup task related to finding a work

[9] That is, the sessions followed the *master–apprentice* model (Holtzblatt & Beyer, 2016, Chapter 3; Kuniavsky, 2003, Chapter 8).
[10] https://cardsort.userbit.com

with a known title, and one task where a specific work was sought, but only some information about its plot and characters was given.[11] Interviewees also spontaneously chose to perform lookup tasks related to finding works based on loose specifications, such as wanting a 'comfort read' or wishing to recommend something new to a friend.

The card-sorting task was intended as a cognitive aid in support of a more qualitatively-oriented discussion of user satisfaction, as the full fourteen-item Likert scale in the survey was judged to be too long and repetitive for an oral interview. Thus, the actual categories and card arrangements produced will not themselves be reported here; rather, the conversations around them form parts of the qualitative analysis of the interviews as a whole (see sections 4.3.2 and 4.3.3). In practice, the card-sorting task was also used by participants as an opportunity to give feedback about the interview experience itself.

A few days before their session, participants were given information about CCI and, in general terms, how the main observation session would work. Consent forms for personal data collection were also sent out at this time, and participants invited to ask any questions they may have about the process.

4.2.3 On autoethnography

As was mentioned in chapter 3, autoethnographic research is common, and to some extent even expected, in fan studies (Busse, 2017; Jacobs, 2020; Popova, 2020). While the use of oneself as a research subject sometimes comes with accusations of bias and weakened objectivity (Busse, 2017; Neustaedter & Sengers, 2012), autoethnography and autoexperimentation have rich histories, for example within medical research (see e.g. Chen et al., 2012; David et al., 2014; Poldrack et al., 2015) and, relevantly for this study, in the field of HCI (Garcia & Cifor, 2019; Marcengo et al., 2016; O'Kane et al., 2014), including user-centred design (Neustaedter & Sengers, 2012).

The chief risk of the autoethnographic strategy used in the present study is that of confirmation bias. Specifically, this would manifest as the researcher allowing his search strategy or his subjective reports of difficulty and success (see section 4.3.3) to be informed by what he wants or expects the results to show, perhaps by letting his observations of the interviewees inform his choices of search methods to increase or decrease success rate, or by choosing to perform certain types of lookup tasks simply because they might prove a particular point. In the case of the 'lost fic' lookup task described at the top of this page, the researcher could also have used knowledge of the work not shared with the interviewees – such as the work's title – in order to achieve success more quickly.

[11] This is a very realistic lookup task; in fact, questions of this type are asked frequently enough in fandom spaces that there are forums and social media accounts entirely dedicated to finding 'lost fics' – Hill and Pecoskie (2017, pp. 848–850) give some examples of this practice.

Now, the present researcher is, as has been noted, a member of the fandom studied here; he is also, relevantly, a user of AO3, and has been so for long enough that he has developed strong habits and preferences regarding search methods and strategies. The aim of the CCI phase of the present study was simply to find out *how fans search for fanworks*. The limitations on the sample described in 4.2.1 meant that no serious comparisons between participants with different attitudes or levels of experience would be possible, so the 'expected' or 'desired' results at this point were simply a set of examples of how users of AO3 use AO3 in practice. Knowing the aim of the study, then, is unlikely to have biased its results; in fact, both interviewees were given information about the study and its aims and methods (see section 4.2.2) before their sessions began. If anything, the researcher might have been 'biased' to use strategies that were as close to his usual habits as possible – which is exactly what contextual inquiry seeks to observe.

Watching others use a familiar tool may, of course, give an observer new ideas for how to better use it themselves. However, this is true for any user of any system; AO3 users certainly do not exist in a vacuum, and having learnt search tricks from other people would not have disqualified other participants. As it happens, however, most strategies employed by the interviewees that were different from those already favoured by the researcher were ones he had little interest in trying for himself. Thus, the actual influence of interviewees' search strategies on those of the researcher was low.[12]

Finally, making the 'lost fic' lookup task fair is as simple as limiting oneself to using only the exact clues given to other participants, and as the researcher chose lookup tasks equal or equivalent to those of the interviewees, the risk of bias in the choice of tasks was limited.

4.3 Data analysis methods

4.3.1 Statistical analysis

Statistical methods were employed to analyse the quantitatively-oriented data from the survey (responses to all closed-ended questions as well as one short-form, open-ended question). To quickly get an initial overview of the response set, the Reports feature of Survey&Report was used to create some quantitative summaries and simple graphs of responses. All further calculations were made in Microsoft Excel.

The *Search satisfaction* and *Tagging confidence* Likert sets were analysed for internal reliability similarly to the pilot test responses, that is, negatively-worded scale items were reverse-coded and Cronbach's α and item–total

[12] Note that the sheer number of ways in which one can search for fanworks on AO3 (see section 2.3 for a brief overview) might be an argument in itself: While two users of Google, say, have essentially one search method to pick from and very limited power over exactly how their search phrase will affect results, on AO3, the situation is very much the opposite.

correlations were calculated and analysed for both response sets. Both Pearson's *r* and Spearman's ρ were used, and correlations were calculated using scale totals with the scale item's contribution removed, as described in section 4.1.2. Calculations were repeated after removing the items with the lowest item–total correlations. Scale totals were then normalised to percentages of the maximum possible score.

The *Tagging familiarity* question set, while much smaller at only four items, was also checked for internal consistency using Cronbach's α and Spearman's ρ and added up to a total score. The *Specialist experience* question set was combined so that respondents with at least one "yes" answer were marked as 'specialists' and the remaining respondents, 'non-specialists'. Different variables in the resulting response set were then compared, pairwise, using suitable statistical methods. For most pairs of variables, Spearman's ρ was deemed the most appropriate, as it is suitable for comparing ordinal, interval, and ratio-level variables (Bryman et al., 2021, Chapter 15; Wildemuth, 2017a). However, since it can be argued that Likert *scale totals* may be treated as interval variables, even though Likert *items* might not (Carifio & Perla, 2008), Pearson's *r* was used to compare the total scores of *Search satisfaction* and *Tagging confidence*.

4.3.2 Qualitative content analysis

A conventional content analysis (Hsieh & Shannon, 2005) approach was used to analyse the free-form answers in the survey and the interview transcripts.

First, after the survey was completed but before interviews were held, all free-form survey answers were grouped by question and read as a list from start to finish without taking any notes. Then, answers were grouped by respondent, and the new list was read from start to finish in the same way to achieve further immersion. On a printed list of answers grouped by question, the researcher then highlighted key words and phrases and noted down terms to create initial codes, which were transferred back into the computer and sorted into rough categories. This initial, tentative coding scheme was then applied to the list of answers grouped by respondent, and any apparent gaps in the scheme, initial impressions, patterns, and connections to existing research noted down.

Similarly, after the end of each CCI session, the (manually double-checked) transcript was subjected to the same content analysis method, with the tentative coding scheme developed in the analysis of the survey results being modified, as needed, to accommodate any new codes and categories found.

The fact that the study was conducted by a single researcher with no prior experience of qualitative content analysis naturally places some limitations on the level of rigour possible in this step, posing some risks to the trustworthiness (credibility and dependability) of the results (Given et al., 2023, Section 5.2.3; Zhang & Wildemuth, 2017). Attempts were made to mitigate these risks by

repeatedly re-coding samples of the survey answers and checking whether this new coding matched previous attempts, but such risks can, of course, never be completely eliminated.

4.3.3 'Condensed' contextual analysis

Lastly, the CCI transcripts were analysed using contextual analysis methods, condensed and simplified for application by a single researcher.

Taking inspiration from the CCI case study in Kantner et al. (2003), some quantitative data were compiled after the conclusion of each session, namely:

- The number of lookups performed by each participant
- The number of iterations per search
- The type of each query, using Broder's categories of navigational, informational, and transactional queries (Ceri et al., 2013, Section 14.1)
- The search methods used by the participant (browse, work search, search bar, or other methods), and
- The participant's perception of task success.[13]

In a larger CCI study, this data might have been used to study patterns in variables such as query type and perceived difficulty, search methods and specialist experience, and so on (see e.g. Kantner et al., 2003, Section 3.1.2). Even in a study like the present, however, the data can still provide an *indication* of similarities and differences in search behaviour, and as specific, real-world examples of search behaviour, it can certainly be used to reinforce and enrich the results of the survey.

'Traditional' contextual analysis

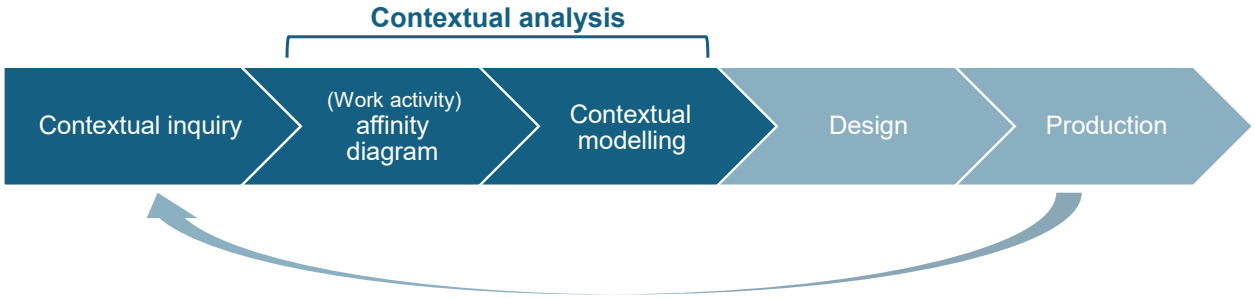
Normally, the goal of contextual analysis is to produce designs for products. This is done through a process in multiple steps, where results from each step are used as the 'raw data' in the next step and different analytical tools from the contextual analysis tool box are applied. First, observation data from contextual inquiry sessions is compiled into a large number of notes, which are arranged in a hierarchical graph of topics or themes, called a *work activity affinity diagram* (Hartson & Pyla, 2012, Chapter 4; Holtzblatt & Beyer, 2016, Chapter 6; Kuniavsky, 2003, Chapter 8). The themes from the affinity diagram are then used to construct one or more models of user behaviour, user attitudes, decision-making processes, and more (Hartson & Pyla, 2012, Chapter 6; Holtzblatt & Beyer, 2016, Chapters 7–8; Kuniavsky, 2003, Chapter 8), and these are used as a basis for creating design concepts and eventually prototypes. The process can then be repeated to evaluate existing products and improve upon them.

A simplified overview of the contextual design process is shown in Figure 4.1.

[13] In the cases where a lookup task was deemed unsuccessful, the decision to stop trying and move on to the next task was made by the participant.

Figure 4.1

Coarse-Grained Illustration of the Steps of Contextual Design



Note. As Design and Production are irrelevant to the present study, they have been greyed out. Diagram constructed by the author based on descriptions of the CD process in Hartson and Pyla (2012, Chapter 2) and Beyer and Holtzblatt (1998, Chapter 1).

A typical contextual analysis process is conducted over a number of days by a team of several people (Beyer & Holtzblatt, 1998, Chapter 7; Holtzblatt & Beyer, 2016, Chapters 3, 6). Many parts of the process, as described in the literature, involve splitting the team into smaller groups, whose insights are then synthesised into a combined view representing the whole team.

The present study

First and foremost, this study is not concerned with designing products, so the analytic process was halted at the point where analysis turns from (inductively) seeking to understand users into (deductively) generating specifications. Thus, the end goal of the analysis is the development of contextual models, rather than designs. Secondly, the analysis in this study was performed by a single person. Fortunately, many parts of the contextual analysis process can be adapted to suit smaller teams or even individuals – Kuniavsky (2003, p. 180) discusses this in the context of affinity diagrams, and Hartson and Pyla (2012; see e.g. Sections 4.5, 6.11) discuss ways to abridge and simplify various parts of the contextual analysis process.

The affinity diagram process can be approximately described as a qualitative content analysis method, specified in great detail[14] for use by professionals (e.g. user interface designers, software developers, and business managers) who are not familiar with qualitative content analysis, which leads to a set of categories and themes presented in diagrammatic form. As these results are, conceptually speaking, similar to the codes, categories, and themes of qualitative content analysis methods common in the social sciences, the present study could re-use parts of the content analysis already performed (see section 4.3.2) for this step.

[14] Holtzblatt and Beyer (2016, Chapter 6) go so far as to specify the colours of post-it notes and types of painter's tape to use on the office wall where the affinity diagram is to be constructed.

The literature describes many different contextual models which can be used to turn the affinity diagram into insights about users and their needs and behaviours. Hartson and Pyla (2012, Chapter 6), for example, outline ten models; Beyer and Holtzblatt (1998, Chapter 6) describe five. Not all of these will be relevant to every project, and practitioners are encouraged to pick and choose whichever models suit their situation best (Hartson & Pyla, 2012, Section 6.11; Holtzblatt & Beyer, 2016, Chapter 5).

In this study, we restrict ourselves to a single model, the *sequence model* (Beyer & Holtzblatt, 1998, Chapter 6; Holtzblatt & Beyer, 2016, Chapter 8). This model is intended to describe the sequences of steps users take to perform tasks, and can thus help us understand in detail how the participants in this study carry out searches on AO3 in practice, as well as how and when they run into problems. The latter is done by identifying so-called *barriers* (Hartson & Pyla, 2012, Section 6.3.4) or *breakdowns* (Beyer & Holtzblatt, 1998, p. 99), points where a user's work is interrupted or impeded for any reason. Sequence models were constructed for each individual lookup task undertaken during the sessions, and were then consolidated to find similar and different intentions and breakdowns across tasks and participants. A full list of all breakdowns was also compiled.

Breakdowns

What exactly constitutes a breakdown varies, but in this study, we use the following criteria, inspired by Kronberg (2020, Section 4.1.3):

a. The user expresses frustration
b. The user makes a mistake
c. The user needs (or believes they need) to use a work-around
d. The user is unable to complete the task

If a situation satisfies at least one of these criteria, it constitutes a breakdown.

It is important to note here that the word "mistake" does *not*, in general, imply incompetence or failure on the user's part. A central principle of user-centred design is that systems should be designed to help users do what they need; thus, it might fairly reasonably be argued that many cases of 'user error' are in fact symptoms of a poorly designed system.[15] In contextual analysis, we are interested in breakdowns because of what they tell us about the system, not the individual user. This is, of course, doubly true when multiple users share similar 'mistakes.'

The condensed and simplified contextual analysis described above was conducted after all CCI sessions and the content analysis of their transcripts were concluded. While contextual analysis is often started while sessions are

[15] That said, item b does also cover accidental misclicks, misspellings, and other innocent everyday blunders.

still on-going, or even during the first session itself (Holtzblatt & Beyer, 2016, Chapter 3), waiting in this way allowed the results of the content analysis to simplify the CA process and inform the choice of model. As there were only three sessions, waiting did not lead to an overwhelming amount of data to analyse at once, which could have been the case in a larger study.

Finally, as we are interested in the method itself, participants' attitudes to and feedback on the sessions will also be (briefly) reported.

4.4 Research ethics

4.4.1 Privacy and pseudonymity

Studying fandom gives rise to particular ethical concerns. Fans often use pseudonyms online and may have legitimate worries about the effect on their private and professional lives, should their interest in taboo topics (such as sexuality or LGBTQ+ issues) or participation in a stigmatised subculture become known (Bennett, 2017; Dym & Fiesler, 2020; Marsh & Lerner, 2023). The grey area in which transformative works exist from an immaterial rights standpoint may also be cause for anxiety, should one's identity as a fan-creator be revealed (Marsh & Lerner, 2023).

The delicate nature of some of the information that fans share within the safety of pseudonymity raises special concerns with respect to data protection under the General Data Protection Regulation, GDPR. Survey questions needed to be designed so as not to accidentally collect personal information and CCI participants instructed to avoid revealing sensitive personal information during interviews. Furthermore, as the University's policy for GDPR in student theses required that participants sign consent forms with their legal name, survey respondents needed to be advised not to use a fandom e-mail account when contacting the researcher unless they wanted the connection between their pseudonym and legal name to become known to the researcher. Participants also had to be given the option to perform the lookup tasks during the CCI sessions logged out of their AO3 account, if they wished, as their user name would otherwise be visible on the recording. Assuring participants of the safety of their personal information was also particularly important, as online fans tend to be protective of their privacy (Marsh & Lerner, 2023) and can be suspicious of external observers, including researchers (Dym & Fiesler, 2020), due to outsiders' sometimes biased perceptions of fandom.

When researchers are themselves participants in fandom, this poses its own concerns of bias, conflicts of interest, and more (Busse, 2017). Furthermore, concerns regarding participant anonymity may apply to the researcher-as-a-fan as well. While the author's full legal name is, of course, present on the cover of this thesis and was used in all communication surrounding the study, the reluctance to connect fandom aliases to 'government names' – 'real' identities,

if you will[16] – certainly applies to this author as well, which meant that all pseudonyms the author has used in fandom needed to be withheld from participants. The result was a curious 'anonymous deanonymisation,' whereby it was quite possible that the researcher and any given participant had known each other for some time, but due to the usage of legal names instead of pseudonyms, neither of them knew whether this was the case. As transparency is crucial to maintain trust in research (Swedish Research Council, 2025, Chapter 2 and Section 7.2), this desire for separation of 'identities' on the researcher's part and the justifications for it needed to be addressed in an honest and transparent way when communicating with participants.

4.4.2 Lexical and linguistic considerations

Recognising Dym and Fiesler's (2020) assertion that fans are sometimes uncomfortable with being studied academically, it is, arguably, particularly important that the text of this thesis treats participants as fans and individuals with agency rather than faceless research subjects. While anonymity precludes this to some extent, an attempt is made to provide a semblance by using words such as "participant," "respondent," and "interviewee" rather than "subject." In a similar vein, we will generally use words such as "fans" and "participants" rather than "users," even when considering fans *as* users of the web site AO3, because although AO3 is viewed, in some parts of this study, as primarily a system for information retrieval, this is, in reality, only a very small aspect of the full story of fans' engagement with AO3.

Finally, online fandom is a global phenomenon, and most members of the communities where invitations were posted do not speak Swedish. Thus, in the interest of open communication (Swedish Research Council, 2025, Chapter 7), the survey, related information texts, and this thesis were all written in English.

[16] Although what 'real' means in terms of personal identity is a question best left to the field of philosophy.

5 Results and analysis

A total of 62 respondents completed the online survey. The survey was opened on March 9, 2026, and set to automatically close on March 19 with the option to keep it open longer if needed; however, after the first week had passed, the rate of new responses slowed down to less than one per day, so the survey was allowed to close as planned. One-on-one CCI interviews with two survey respondents were subsequently held in April and May, 2026.

5.1 Quantitative survey results

5.1.1 Likert sets: internal reliability and item analysis

Using the methods described in section 4.1.2 to test internal reliability, the 14-item *Search satisfaction* Likert set was found to have a Cronbach α of .93, which is an excellent level of internal consistency. Item–total correlations varied between .40 and .79 using Spearman's ρ, with an average value of .63. Values were somewhat higher using Pearson's *r*, with an average item–total correlation of .68.

Removing the item with the lowest item–total correlation, item 4, gave a slightly improved average ρ of .65 and an average *r* of .70, with item–total Spearman correlations ranging from .44 to .81. The Cronbach α increased very slightly to .94. All item–total correlations are shown in Table 3.

According to Bikos (2024, Section 7.4.2), none of these correlation coefficients are concerningly low, so even if one includes the somewhat lower item 4, the *Search satisfaction* set has good enough internal consistency and item–total correlations that it can justifiably be treated as a combined scale.

The Cronbach α of the full eight-item *Tagging confidence* set was .53, which, as detailed in section 4.1.2, is too low to be deemed acceptable. Item–total correlations (see Table 4) showed extremely low values for items 3 and 6, $\rho = -.01$ and $\rho = -.06$, respectively, and the average of all item–total correlations was .32, which is below the cut-off suggested by Bikos (2024, Section 7.4.2), but just about enough for Chin et al. (2002).

With item 3 removed, the Cronbach α improved to .57. With item 3 kept and item 6 removed, $\alpha = .60$ and the average item–total correlation using Spearman's ρ improved to .38. Finally, with items 3 and 6 both removed, α was .63; a questionable but passable level of internal reliability. However, this Likert set now only had six items, meaning that the Cronbach α may have been lower than it might have been, had the scale been more extensive. It is also worth noting that this set only had 25 responses, as the question was only shown to participants who said that they had posted works on AO3. Overall, this set may just about be argued to work as a combined scale, albeit somewhat

less convincingly than the *Search satisfaction* set.

The *Tagging familiarity* matrix, while much smaller at only four items, had a Cronbach α of .60. Item–total correlations ranged between ρ = .27 and .47 with an average of .36, which rose to an average ρ = .42 when the lowest-scoring item, item 1, was excluded. This set, thus, is also debatable or just about acceptable as a combined scale. All correlations are shown in Table 5.

Table 3

Search Satisfaction, Item–Total Correlations (ρ) on Full Set vs. Set Excluding Item 4

	Item	Full set	Ex. 4
1.	I enjoy using the search interface.	.77	.80
2.	The search interface is unnecessarily complex.	.62	.63
3.	The search interface is easy to use.	.78	.79
4.	I needed (or would need) to read the FAQ or ask a friend before I could understand the search interface.	.40	
5.	The various functions in the search interface suit my needs well.	.72	.74
6.	It's unclear what some of the functions in the search interface are for.	.52	.48
7.	I think that most people would learn to use the search interface quickly.	.45	.44
8.	The search interface is very cumbersome to use.	.79	.81
9.	I feel confident using the search interface.	.61	.60
10.	I'm not sure the search interface was made with people like me in mind.	.71	.68
11.	The search results are usually precise and relevant.	.62	.64
12.	I think there is too much inconsistency in the search interface.	.52	.51
13.	It's hard to get the search results I need.	.73	.72
14.	Overall, I am satisfied with the search interface.	.64	.65
	Average item–total correlation	.63	.65
	Cronbach's α	.93	.94

Note. Item–total correlations calculated using total scores with item's contribution removed.

Table 4

Tagging Confidence, Item–Total Correlations (ρ) on Full Set vs. Set Excluding Items

Item	Full set	Ex. 3	Ex. 6	Ex. 3, 6
1. I feel like I understand the purposes of the different tag types (fandoms, rating, relationships, and so on).	.30	.21	.28	.25
2. I am not sure how important tags are to get my works seen.	.35	.22	.34	.23
3. I know where to go if I need more information about tags and tagging.	–.01		.11	
4. Deciding what tags to use is difficult.	.33	.41	.36	.47
5. I feel confident at choosing tags for my works.	.62	.51	.69	.63
6. There is a high threshold for new users to learn how to tag.	–.06	.05		
7. Overall, I like the tagging system.	.48	.46	.43	.48
8. I still think I need more practice or instruction in order to tag works well.	.52	.56	.45	.55
Average item–total correlation	.32	.35	.38	.43
Cronbach's α	.53	.57	.60	.63

Note. Item–total correlations calculated using total scores with item's contribution removed.

Table 5

Tagging Familiarity, Item–Total Correlations (ρ) on Full Set vs. Set Excluding Item 1

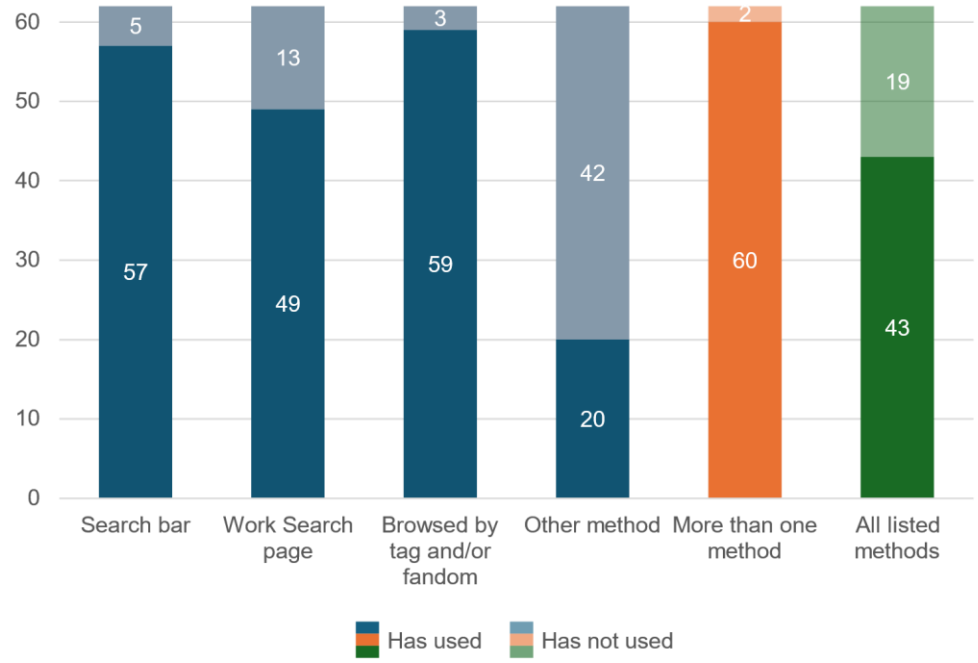
Item	Full set	Ex. 1
1. The hierarchy of tags on AO3	.27	
2. The meaning of "&" and "/" in relationship tags	.47	.45
3. What canonical and non-canonical tags are	.29	.41
4. The popular genres and tropes of your fandom	.39	.39
Average item–total correlation	.36	.42
Cronbach's α	.60	.58[a]

Note. Item–total correlations calculated using total scores with item's contribution removed.

[a] This decrease in α may well have been caused by the very small number of items.

Figure 5.1

Question 1, Search Methods Used by Participants



5.1.2 Summary of responses

Search methods (questions 1–3)

Question 1 began by reminding respondents of the three main ways of searching for works on AO3: searching using a search bar at the top of every page; the work search (or advanced search) page; and the browse feature. Respondents were then asked whether they had ever used any of these methods or some other method not listed to search for fanworks on AO3. Respondents who said they had used some other method were asked in question 2 to describe that method.

Out of the 62 respondents, 57 had used the search bar, 49 had used the work search page, 59 had at some point used the browse feature, and 21 stated that they had used some other method (see Figure 5.1 above).

The responses from the free-text question 2 fell into six different categories. Seven participants said that they sometimes search AO3 using an external search engine. Six participants described going to the profile page of the author of a work they liked and looking through that author's other works. Six said that they browse other AO3 users' bookmarks, and two stated that they read recommendation lists on other platforms. Three described clicking links in the metadata of works, such as tags or the names of collections that a work is part of, to find similar works – borrowing a term from Silberstein-Bamford (2024, pp. 139–141), we will refer to this practice as "rhizomatic reading." Finally, two respondents described the browse feature using different words. As both of

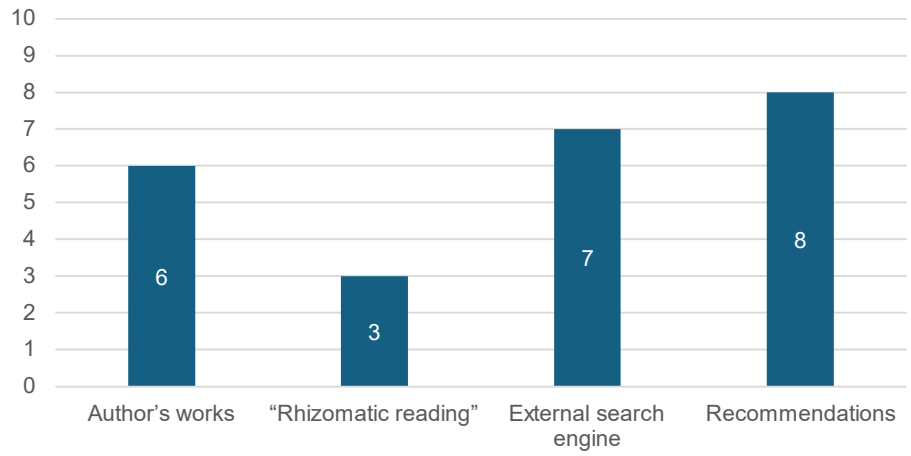
these had already answered, in question 1, that they have used the browse feature, and one of them did not describe any other method beyond it, the total number of respondents who have used another method than the three listed is in fact 20.

Almost all (60 respondents) had used more than one method, and 43, or just over 69%, had used all three of the methods listed in question 1.

Question 3 asked which search method respondents preferred. Browsing was preferred by a majority – 41 respondents – while three people preferred the search bar, 13 preferred the work search page, and five had no overall preference. No respondent said that they prefer another method described in question 2.

Figure 5.2

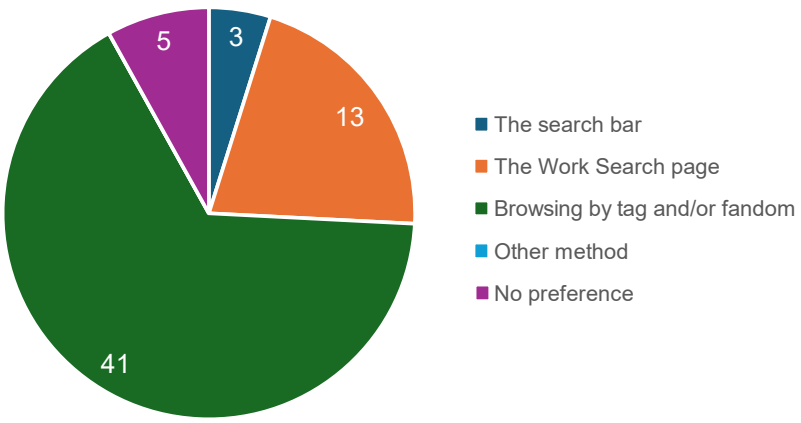
Question 2, Other Search Methods Described



Note. "Other users' bookmarks" and "recommendation lists off-site" have been combined into a single category, "Recommendations."

Figure 5.3

Question 3, Preferred Search Method



Search satisfaction (question 4)

Question 4 was the *Search satisfaction* Likert scale. Respondents rated their agreement or disagreement with each item while considering a typical time when they had searched for fanworks. Total scores, expressed in percents, ranged from 23.2% to 96.4%, with a mean of 65.4% and a median of 70.5%. Excluding item 4, the mean dropped to 64.7%, while the median rose to 71.2%. The range also grew by a few percentage points. The distribution of *Search satisfaction* total scores, with and without item 4, are illustrated in Figure 5.4.

Tagging confidence and familiarity (questions 7–9)

Question 7 asked whether participants had ever posted any works to AO3 themselves. This question was used as a filter for question 8, the *Tagging confidence* Likert set, which was only shown to the 25 respondents, or about 40% of the full sample, who said that they had posted works.

Total scores were overall low, with the maximum score being 56.3% for the full set, 53.6% for the full set excluding item 6, and 54.2% for the set excluding items 3 and 6. Average and median scores sat at about or just under 35% for all three 'versions' of the scale (see Figure 5.5 and Table 6).

Figure 5.4

Distribution of Search Satisfaction (Question 4) Total Scores in Percent

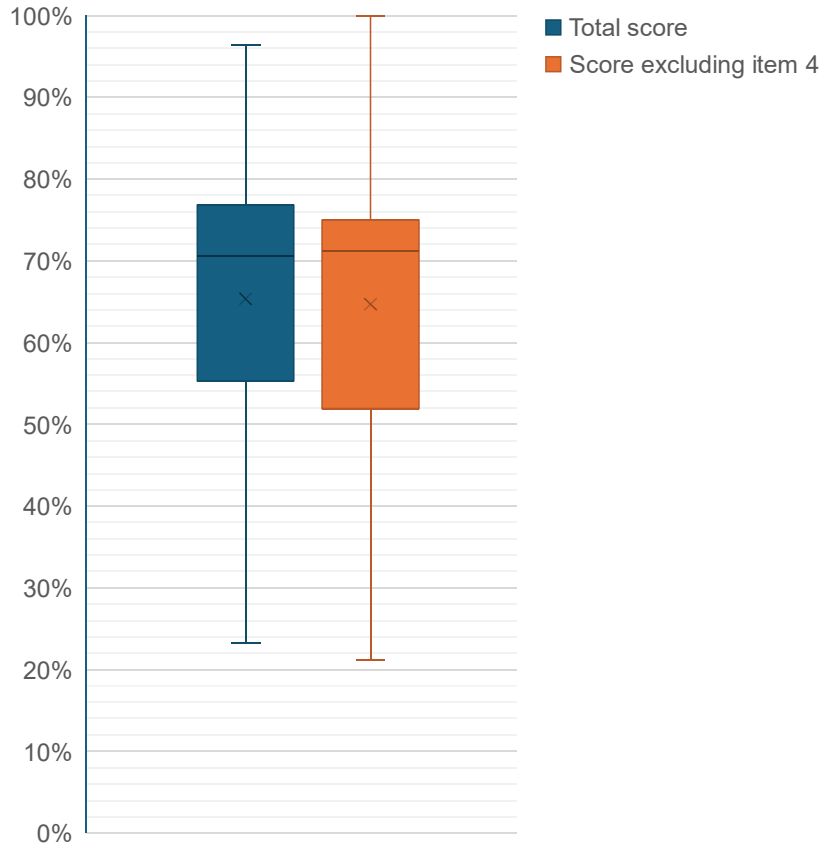


Figure 5.5

Distribution of Tagging Confidence (Question 8) Total Scores in Percent

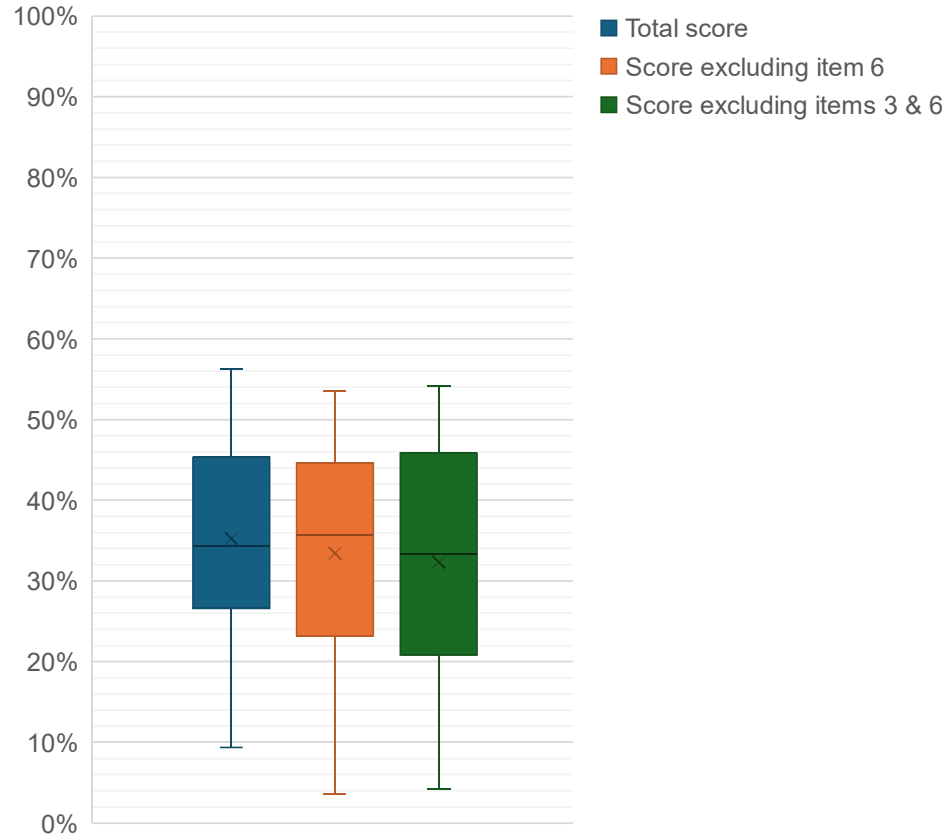


Table 6

Tagging Confidence (Question 8) Total Scores in Percent: Minima, Maxima, Means, Medians, Standard Deviations, and Modes

%	All Items	Excluding 6	Excluding 3, 6
Min	9.4	3.6	4.2
Max	56.3	53.6	54.2
Mean	35.3	33.4	32.3
Median	34.4	35.7	33.3
Standard deviation	11.8	13.1	14.0
Mode(s)	34.4	35.7 and 50.0	45.8

Question 9, *Tagging familiarity,* listed four aspects of AO3's tagging system (see section 4.1.1), and respondents rated their familiarity with each of these aspects on a four-point scale. All respondents, regardless of whether they had posted works or not, were shown this question.

As might perhaps be expected for questions about self-reported knowledge without verification, most users reported high familiarity ("understand reasonably well" or "very familiar") with most items, and summed totals (where answers were given a score of 0 = *don't know at all* through 3 = *very familiar*) were high, with a median score of 9 out of a maximum of 12. Figure 5.6 shows the distribution of *Tagging familiarity* totals, and answers to the individual 'knowledge items' are illustrated in Figure 5.7.

The first 'knowledge item' in question 9, concerning the tag hierarchy, was a little less familiar to respondents, with only about half being reasonably or very familiar, while the item about popular genres was reasonably familiar or very familiar to over 98% of respondents.

Platform experience (questions 5, 11, 12)

These questions ask how long respondents have used AO3 (question 11), how frequently they use AO3 at the moment (question 12), and whether they have ever used any platforms other than AO3 to search for fanworks (question 5).

Almost all respondents in this survey had used AO3 for a year or more, with the largest proportion – 48%, or 30 respondents – having first begun to use it between one and five years ago (see Figure 5.8). Most respondents are also frequent site users: over three quarters answered that they use AO3 several times per week, with only four respondents stating that they use it less frequently than about every week (see Figure 5.9).

Figure 5.6

Tagging Familiarity (Question 9), Distribution of Summed Totals

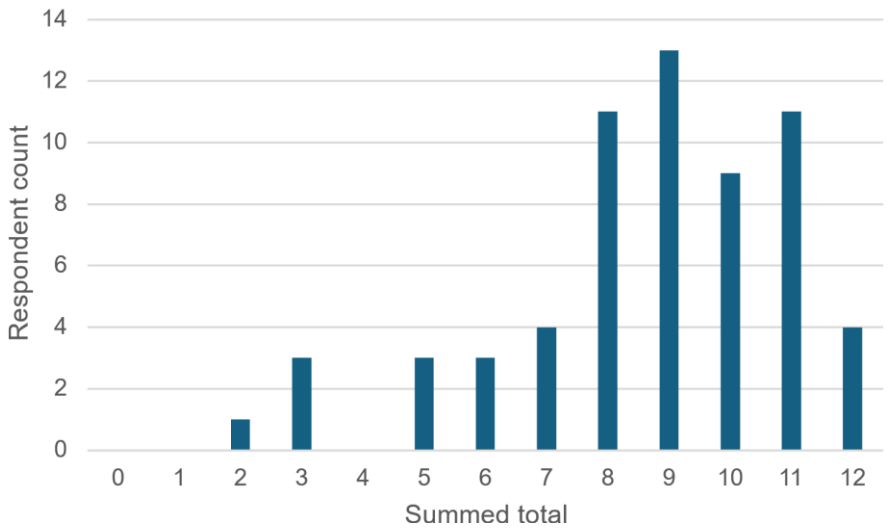
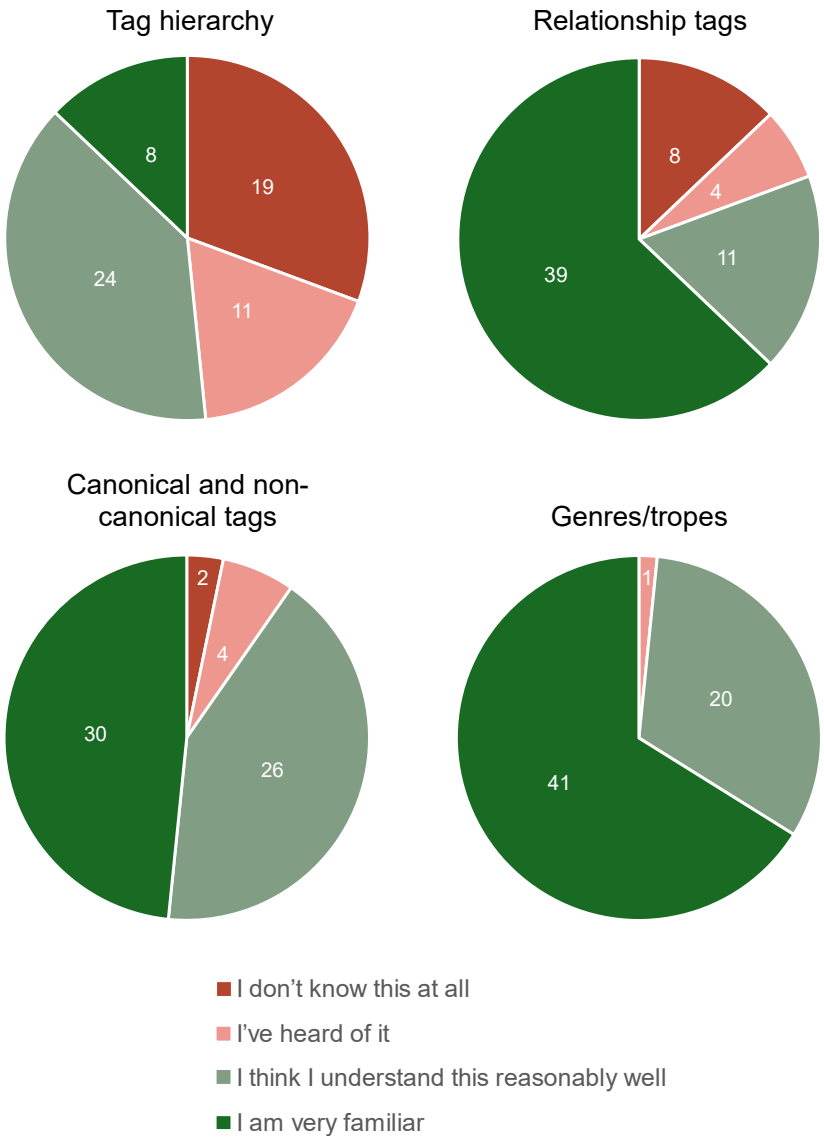


Figure 5.7

Tagging Familiarity (Question 9), Individual 'Knowledge Items'



When it came to other platforms, the group who said that they have never used another platform to search for fanworks was almost as large as the group who had used one or two other platforms. Only a handful of people had used more than two other platforms to search for fanworks (see Figure 5.10).

Specialist experience (question 13)

Question 13 concerns the specialist experiences mentioned in section 1.2. Participants were asked about their academic and professional experience in library and information science and IT, as well as whether they have ever volunteered for AO3 as tag wranglers. No participant in this survey had ever been a tag wrangler, but 22 respondents had experience in at least one of the other four things listed. 14 of these respondents were fannish librarians.

Figure 5.8

Question 11, "When Did You First Start Using AO3?"

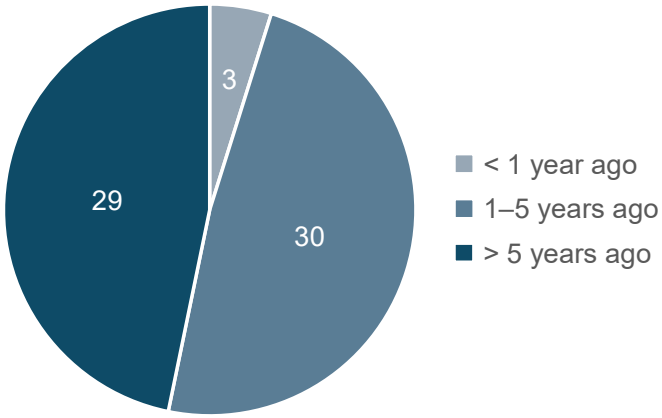


Figure 5.9

Question 12, "How Often Would You Say You Use AO3?"

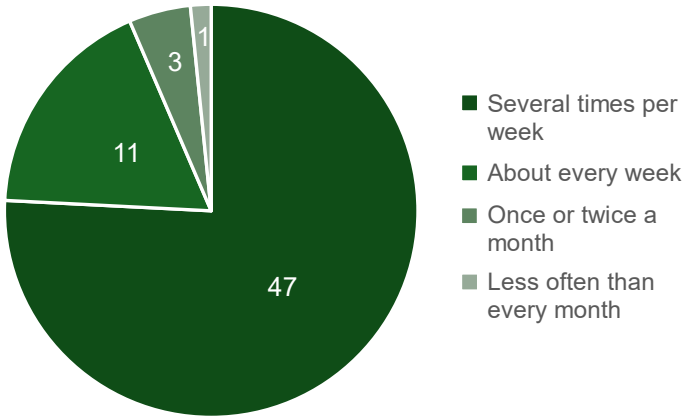


Figure 5.10

Question 5, "Have You Used Any Other Platforms to Search for Fanworks?"

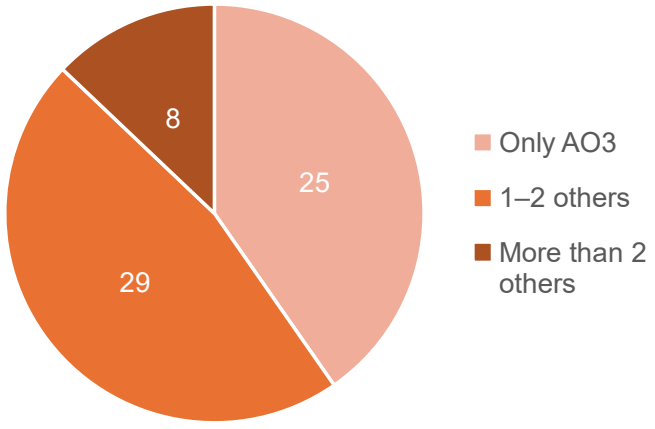
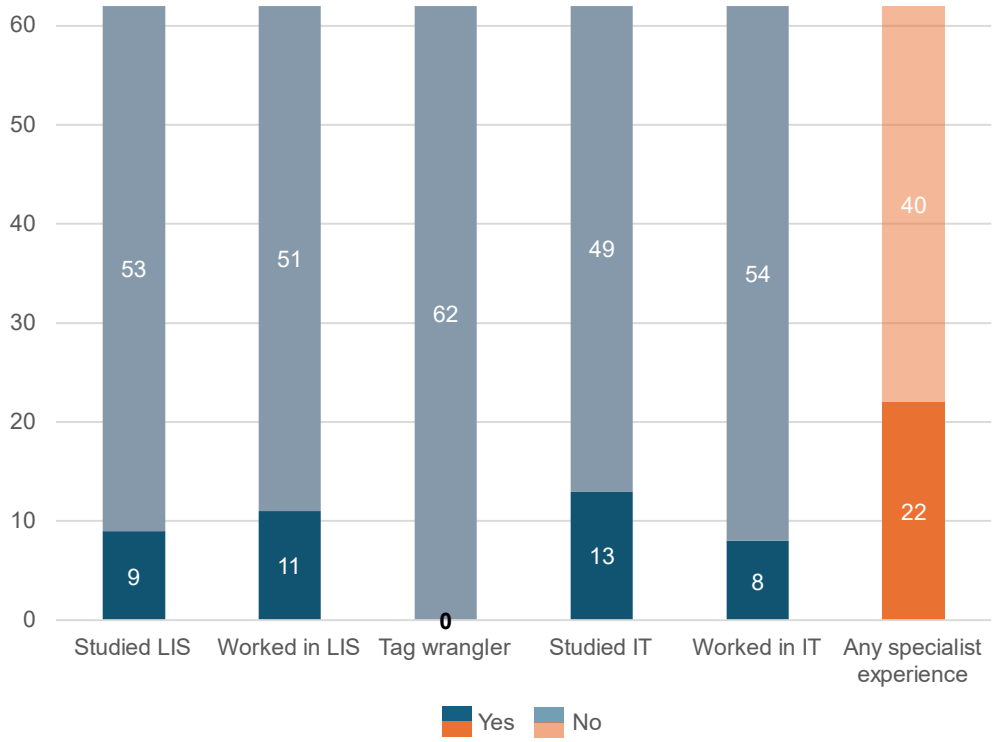


Figure 5.11

Answers to Specialist Experience Matrix (Question 13)



Note. A respondent is counted in the "Any specialist experience" column if they have answered "yes" to at least one question in the experience matrix.

5.1.3 Correlations

To determine whether the answers to any of our questions – that is, our variables – have a relationship with each other, we compare variables pairwise using statistical methods appropriate for the variable types. Most of our variables are ordinal, that is, their possible values have an order to them, but the distances between the possible values may not be equal and it might not make sense to imagine 'in between' values (Wildemuth, 2017b).[1] As was discussed in section 4.3, Spearman's ρ is a suitable correlation statistic for this type of variable. For interval variables, which do have uniform distances between their possible values, Pearson's *r* may also be used (Wildemuth, 2017a). For both Pearson and Spearman correlations, values will range from –1 to 1 inclusive, with values far from zero implying stronger correlations, and negative values implying that an increase in one variable is connected to a decrease in the other. Correlations between –.5 and .5 are deemed 'weak.'

[1] To take the *Tagging familiarity* set as an example, it makes little sense to say a respondent "doesn't know 'and a half.'"

Specialist experience and Search satisfaction

Comparing the total score for *Search satisfaction* to the dichotomous (two-valued, or binary) variable "Any specialist experience" gives a Spearman correlation of ρ = –.01. Comparing *Search satisfaction* instead with the number of "yes" answers in the *Specialist experience* set gives a ρ of –.03. These correlation coefficients are very close to zero, and thus, the correlation between these variables is negligible.

Specialist experience and Tagging confidence

For *Tagging confidence*, a judgement needs to be made on which 'version' of the scale to use. Excluding items makes an already small item set even smaller, however, the item–total correlations for items 3 and 6 were found to be so low that it is difficult to argue that these items measure the same thing as the rest of the scale (see section 5.1.1). Thus, the scale excluding items 3 and 6 will be used here.

Comparing *Tagging confidence* scores to "Any specialist experience" yields a Spearman correlation of ρ = –.41, a weak negative correlation. Using the number of "yes" answers to *Specialist experience* instead of "Any specialist experience" gives ρ = –.40. Both of these correlations are rather weak, so it is hard to say anything conclusive about the connection between these variables.

Specialist experience and Tagging familiarity

Comparing the *Tagging familiarity* total with "Any specialist experience" and the number of "yes" answers to *Specialist experience* gives a ρ of –.03 and –.10, respectively. *Specialist experience*, therefore, does not correlate with *Tagging familiarity* in this data set.

Table 7

Correlation Table (ρ) for Specialist Experience, Search Satisfaction, Tagging Confidence, and Tagging Familiarity

	1	2	3	4	5
1. Specialist experience (any)	1				
2. Specialist experience (number of yesses)		1			
3. Search satisfaction score	–.01	–.03	1		
4. Tagging confidence score[a]	–.41	–.40	–.12[b]	1	
5. Tagging familiarity total	–.03	–.10	.48	–.25	1

[a] Excluding items 3 and 6. [b] Calculated using Pearson's *r*.

Table 8

Correlations (ρ) Between Search Satisfaction and Length of Time Using AO3, Frequency of Using AO3, and Other Platforms Used

	AO3 since	AO3 freq.	Other platforms
Search satisfaction score	.39	–.11	.41

Search satisfaction and Tagging confidence

Search satisfaction and *Tagging confidence* are both Likert scale scores, and may therefore be treated as interval variables (Carifio & Perla, 2008). Thus, we use Pearson's *r* to compare them. The correlation coefficient between these variables is $r = –.12$, which is a very weak correlation.

Search satisfaction and Tagging familiarity

The correlation between the *Search satisfaction* score and the total sum from *Tagging familiarity* is $\rho = .48$, a weakly positive correlation.

Search satisfaction and platform experience

Spearman correlations between *Search satisfaction* and the three *Platform experience* variables are shown in Table 8. There is a weak correlation, $\rho = .41$, between *Search satisfaction* and the number of other platforms used, an even weaker correlation, $\rho = .39$, between *Search satisfaction* and the length of time during which the respondent has used AO3, and a very weak correlation between *Search satisfaction* and the frequency of AO3 use.

Search method preference versus Search satisfaction, Tagging confidence, and Tagging familiarity

Search method preference is a categorical variable, that is, there is no reasonable way to assign a numerical 'order' to its possible values.[2] We cannot, therefore, use Spearman or Pearson correlations to compare it to other variables. (Wildemuth, 2017a). The usual recommendation for categorical variables is to use contingency tables and the Chi-square (χ^2) statistic; however, χ^2 is unsuitable for data sets where the sample size is small and one variable has many possible values.[3] This is, unfortunately, the case for *Search satisfaction, Tagging confidence*, and *Tagging familiarity*.

[2] The possible values of the search method preference variable are, to reiterate: the search bar, the Work Search page, browsing by tag and/or fandom, another method, or no preference.

[3] The reason, although it will not be explained in full detail here, is that if, when one calculates the expected values necessary to find χ^2, one finds that too many of these values are too low, χ^2 will not yield a valid result (Wildemuth, 2017c). In the three cases here, the *majority* of the expected values are, in fact, too low.

Figure 5.12 and Figure 5.13 below show visually the differences in mean and median scores for *Search satisfaction, Tagging familiarity,* and *Tagging confidence* over different search method preferences.

Note that *Tagging familiarity* bars are sized by percentage of the maximum score (12), but labelled with actual scores.

Figure 5.12

Average Score Totals by Search Method Preference

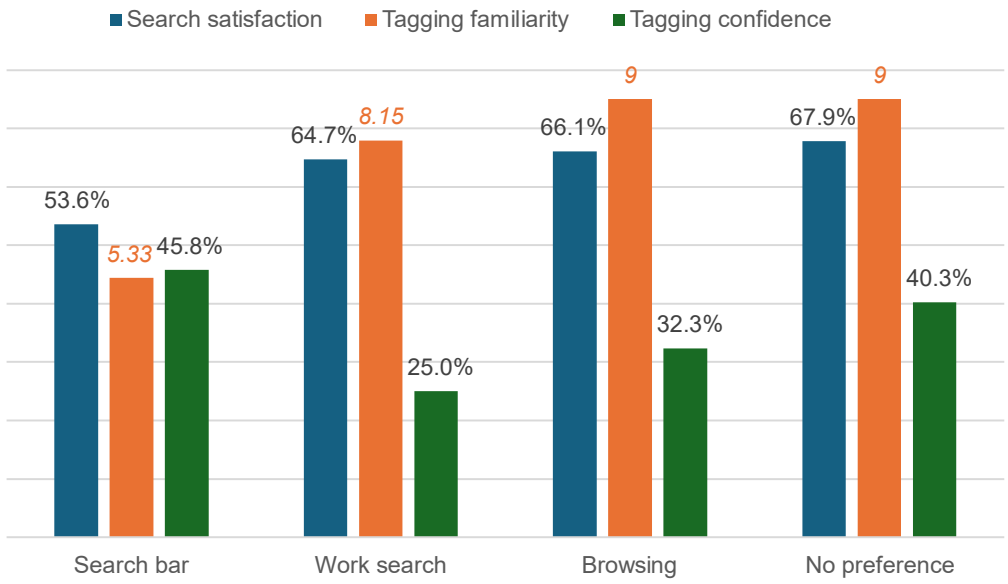
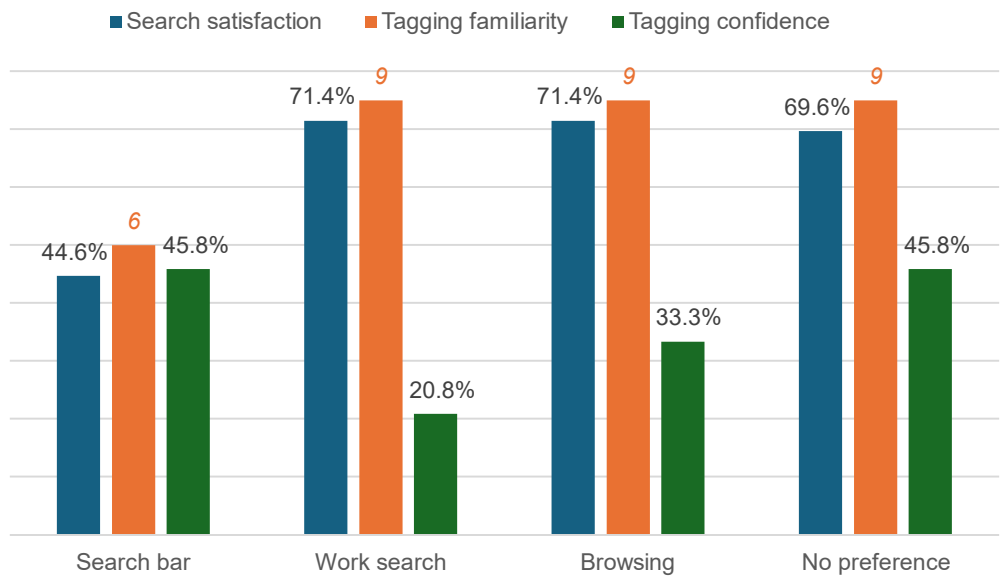


Figure 5.13

Median Score Totals by Search Method Preference



An ocular inspection of these charts indicates that respondents who prefer to use the search bar appear to get lower scores on *Search satisfaction* and *Tagging familiarity* and that respondents preferring work search appear to be less confident at tagging. However, the search bar was only preferred by a total of three respondents, and only five respondents preferred work search and also filled out the *Tagging confidence* matrix, so in actuality, it is hard to draw any conclusions at all about these groups.

5.2 Qualitative content analysis

The open-ended questions 6, "Do you have any further thoughts about searching for fanworks on AO3?" and 10, "Is there anything else you'd like to say about tagging on AO3?" as well as transcripts from the CCI sessions, were analysed qualitatively with a focus on attitudes to and factors affecting search and tagging. Question 6 was answered by 31 people and question 10 by 23 people. In total, 33 people gave a response to at least one of the two questions.

A number of themes emerged from the coding – see Table 9 below. Responses which were given in Swedish have been translated by the author, with original quotes included in footnotes.

Table 9

Categories and Themes Emerging From the Open-Ended Questions and Interview Transcripts

Themes and categories	Example quotes
Power and specificity of search	
Filtering system	"I like the tagging system and ability to sort and filter." (R19)[4]
	"I really like the Exclude option. It allows to filter for the work I'm looking for even better." (R60)
Tagging system makes search great	"AO3 has the absolute best search feature I know thanks to the fantastic tagging system." (R14)
AO3 search gives high level of control	"The more you learn about it, the more advanced it is."[5] (I2)
	"It's like an academic database to search to me, which I'm comfortable and familiar with." (R63)
	"Unlike Google search or YouTube search these days, AO3 lets you search for precisely what you want, and respects the search terms you entered." (R12)

[4] For the purposes of anonymisation, survey respondents are referred to as 'R' plus their number in the survey data set, and interviewees are referred to as I1 and I2, respectively. Although the survey was completed by 62 people, some respondent numbers are over 62.
[5] "Ju mer man lär sig om det, desto mer avancerat är det ju."

Themes and categories	Example quotes
Effects of author-tagging	
Tags as content declaration	"I NEED to know what I'm about to read in order to decide whether the work in question is worth the effort." (R48)
	"Tags are a good way to get an idea of what kind of fiction it is, and also useful to search for specific things you want to read (or avoid)." (R52)
	"There's a lot of, like, reader insert stuff and […] I usually exclude that sort of thing."[6] (I2)
Too many or too few tags is a problem	"I sometimes worry I over tag, but I also don't want to jump scare people." (R32)
	"I personally hate to read through a wall of tags on a story." (R60)
	"If you come across a severely undertagged fic, it could mean that it's either the best fic you've read in your life, or a terrible fic, there's no in between." (R5)
Author-tagging is inconsistent	"If I want to search for a specific idea or tag and the author didn't feel the need to include it […] I wouldn't find it." (R60)
	"The search function working as intended wholly depends on writers tagging their fics correctly." (R64)
	"I often struggle pulling out the most important elements in my fics to tag." (R63)
Commentary-style tags	"I also love a good silly tag or two, but I know some people find that annoying." (R32)
	"While jokey tags can be fun to read, I think they should be used as rarely as possible on actual fic headers." (R18)
Tags are hard to understand	"I don't really get how they work or what a lot of tags mean." (R27)
	"Then someone said there are tags that have sub-tags, and […] if you exclude the main tag, sub-tags disappear too. I had no idea about this."[7] (I2)
Issues with author-tagging are hard to solve	"Since these are all self-applied tags I suppose that would be pretty difficult to coordinate." (R27)
	"Determining synonyms and hierarchies of tags, in the fandom space which is full of super-niche subculture jargon and rapidly-shifting language? That's amazing." (R12)

[6] "Det är mycket såhär reader insert-grejer och […] jag brukar ta bort sånt."
[7] "Och det var någon som sa att det finns tags som har undertags, och […] tar man bort huvudtagen så försvinner undertags också. Det här hade inte jag en aning om."

Themes and categories	Example quotes
Specific lookup vs. browsing	
Specific lookups are sensitive to minor errors	"The title can be typed in as precisely as expected with absolutely zero results." (R9)
	"If I know enough about a particular fiction I'll get there quickly, but if [I] only half-remember something it can take a while." (R24)
Broad searches/browsing works well	"It's most useful for well-populated tags." (R18)
	"I prefer to use it when I'm casting a wide net." (R30)
	"I find the brows [sic] feature the easiest and most reliant way to search." (R51)
Iterative modification of search terms/filters as search method	"I often have to change some things in my search before I finally find what I was looking for." (R22)
	"I really like doing iterative searches, starting really broad and then narrowing it down." (I1)
	"The thing is that when you search for tags you have to use your imagination. It's rarely enough to just chose [sic] one and [be] done." (R48)
Search method depends on goal of search	"The answer to question 1b depends on what the information […] I am searching [for] and what information I already possess." (R53)
	"If I know the name of a fic, I'll go straight to Work Search." (I1)
	"Search bar at [the] top when I know it'll be a small fandom or character list, browse with lots of options selected when I'm in a huge fandom […] and work search when I can half remember a title or author's name." (R63)
Interface and discoverability	
Interface is confusing, frustrating	"It's counterintuitive that, in order to exclude tags from a search, you have to click 'Fandom' and select a fandom tag first, then filter. I have to explain this to new users a lot." (R12)
	"I don't think it's a very intuitive search page" (R65)
	"It's annoying that when you search for a specific relationship it also counts friendship, not just 'love'." (R59)
Finding canonical tags with a specific meaning	"It's hard for me to tag things […] because I can't always find already existing tags that match what I want to convey." (R56)
	"Why is the canonical [tag] for turtleneck Crowley 'Crowley's Polo Neck Jumper'? I clicked on that canonical tag once and not a single person used that canon form." (R32)

Themes and categories	Example quotes
Belief that existing features are missing; discoverability issues	"I've seen comments on the ao3 subreddit where people struggle with it (eg 'I wish we could exclude tags' when the sidebar search has that feature.)" (R32)
	"I know there's tags that become official and tags that are just kind of user created, and I don't know if AO3 searches for […] user created tags." (I1)
Steep learning curves	"It took me some time to get used to and to understand what everything meant." (R22)
	"I still haven't got the hang of bookmarks."[8] (I2)
	"It's time consuming to learn how to be better at it." (I1)
The user's responsibility in search and tagging success	"Though, I do believe that this issue is completely on me. I took a long time to read the FAQ or to look something up. I took some things at face value." (R22)
	"I just kind of do a self-tutorial to get better at it." (I1)
	"The more complex one's search needs are, the less likely it is to find results. However, this may be a matter of a user's over specificity." (R18)
Appreciation and affect	
Appreciation towards AO3	"AO3 is perfect in my opinion!" (R56)
	"It really is a useful thing […] I'm happy it exists."[9] (I2)
AO3 compares favourably to other platforms	"I've used both Wattpad and Fanfiction.net before discovering AO3 and there's simply no comparing the three." (R50)
	"I feel like the Google searches, especially if I'm looking for an unusual fic, don't yield any better results than AO3 would." (I1)
	"It's a lot easier than Livejournal and Dreamwith!" (R63)
Love for fans and fandom	"Love the creativity some authors develop in tagging their works – in some instances, reading the tags [is] just as interesting as reading the stories." (R21)
	"It's a whole new world opening up. It makes me so happy."[10] (I2)
	"It's so fascinating […] how many wonderfully talented people there are creating things for others."[11] (I2)

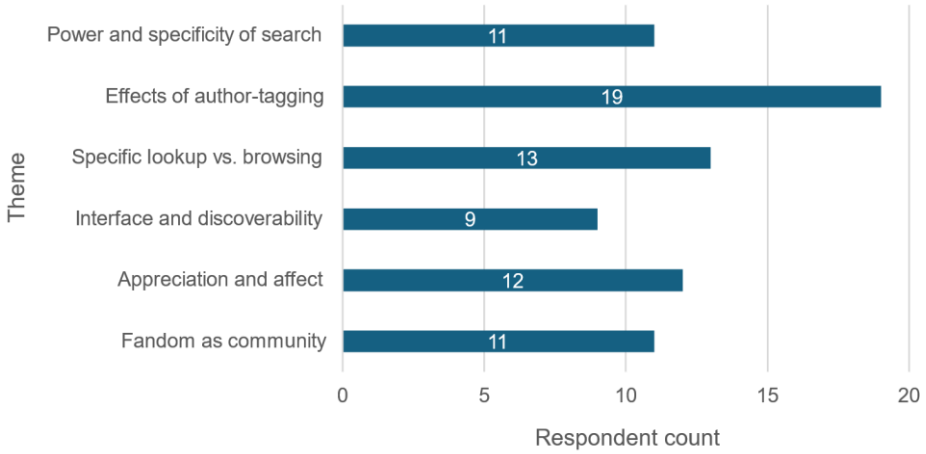
[8] "Bookmarks har jag ju fortfarande inte fått häng på."
[9] "Det är verkligen [en] användbar grej […] man är glad att den finns liksom."
[10] "Det är en helt ny värld som öppnas. Man blir ju bara så glad."
[11] "Det är så fascinerande […] hur underbart många duktiga folk det finns som gör saker för andra."

Themes and categories	Example quotes
Fandom as community	
Social spaces	"And then as I became personally involved in like Discord servers, definitely, that's when my Marked for Later list became three miles long, because I got to know writers." (I1)
	"In this fandom, I'm not a member of a fanfic group or anything like I am in the others."[12] (I2)
	"When I got more into the fandom, I started following recommendations off Tumblr and then off a subreddit." (I1)
Knowledge sharing	"I often struggle pulling out the most important elements in my fics to tag, and often end up getting a friend or people on a server to help." (R63)
	"I've only posted how-to coding works." (R30)
	"I would probably get on Discord and be like, 'What is this fic? Help me!' And see if anyone else can help." (I1)
	"This is another issue that I have to explain to new users a lot." (R12)
Using recommendations from other fans to find works to read	"I like to read what my favorite authors bookmark." (R37)
	"Tags [are] the best way I've found to find new works to read outside of recommendations on Reddit." (R27)
What fan-readers and fan-writers owe each other	"I fear to use [too] many tags. I don't want to overwhelm the reader." (R60)
	"You feel a bit like you just take and take and take, you never give anything back […] I just consume all the time."[13] (I2)
	"I have kind of a complicated relationship to stats on AO3 […] because I know writers and I know they have lots of anxiety about it." (I1)
	"Then the author wrote, 'I'm not getting enough comments, nobody likes this, so this is what I had planned to do.' […] And I'm like, what the hell, if I'd known this – that's not an ending! Then I wouldn't have read it."[14] (I2)

[12] "Den här fandomen är jag inte med i någon fanfic-grupp eller något som jag är [i] de andra."
[13] "Man blir lite såhär att man bara tar och tar och tar, man ger aldrig nånting tillbaka […] jag bara consume:ar liksom hela tiden."
[14] "Då skrev liksom författaren: 'Jag får inte tillräckligt med kommentarer, ingen gillar det här, så det här är vad jag hade tänkt att göra.' […] Och då blir man, men vad fasen, hade jag vetat det här – det är ju inget slut! Då hade jag ju inte läst den."

Figure 5.14

Frequency of Main Themes Among Survey Respondents



Note. A total of 33 survey respondents answered at least one of the open-ended questions. In the interviews, all main themes were expressed by both participants at least once.

5.3 Quantitative CCI results

Two people participated in recorded CCI sessions and interviews over video call, and one additional person participated in an unrecorded CCI session.[15] Recordings amounted to a total of just under 127 minutes, that is, one hour and 3 minutes per video call session on average. Of the interviewees, neither had any specialist experience, both reported using AO3 daily, and both had used AO3 for between two and five years. Both video calls were held at a time of day when participants reported that they would normally be active on AO3.

A total of 13 lookup tasks were performed in the three CCI sessions. The participants needed between one and ten iterations per lookup task, with an average of three iterations per task. Summaries of the quantitative results are given in Table 10 and Table 11.

Most lookup tasks (7) were navigational, that is, the intended goal was a specific work known or assumed to exist. The information available for these tasks varied: in three cases, the participant knew more or less the complete work title; in one case, one word of the title was known; and in three cases, the participant remembered or was provided with a summary description of the plot and setting and the names of some characters appearing in the work, but not the title or author (the 'lost fic' task mentioned in 4.2.2). The remaining six lookup tasks were much less specific, browsing-type searches; for example, in one task, the information need was expressed as, "I want something short and comfortable to read after a long day, something nice and positive." In Broder's terms, these tasks have the characteristics of (mostly) transactional queries, although without any actual monetary transaction taking place.

[15] The third person being the researcher; see sections 4.2.1 and 4.2.3.

Table 10

Lookup Task Counts with Maxima, Minima, Means, and Medians per Participant

Query type	Total	Tasks per participant			
		Max	Min	Mean	Median
All lookup tasks	13	5	3	4.33	5
Navigational queries specific prior information	4	2	1	1.33	1
vague prior information	3	1	1	1	1
Transactional queries	6	3	1	2	2

To account for the "prior information" dimension discussed above, navigational queries were further split into specific versus vague prior information. The transactional queries all had vague prior information, so this category was not split.

No lookup tasks were of the type that Broder calls informational; however, outside of the main observation session, one participant recounted an informational-type search experience in the context of entering an unfamiliar fandom:

> My daughter and I were watching [a show] together and I was like, I wonder what's being written about this […] so I think the first thing I did was, I went to AO3 and just [searched for the show's name] like how many fics are we looking at here?
>
> (I1)

A majority of the lookup tasks – 10 out of 13 – were deemed successful by the participants. Due to the size of the sample, it is difficult to draw any conclusions from the quantitative data; however, it is fairly clear that, for these participants at least, navigational lookup tasks with vague prior information were the most difficult: these tasks required six times as many iterations as navigational queries with specific information, on average, and only one participant succeeded in finding a work fitting the description (see Table 12).

Table 11

Number of Iterations per Lookup Task – Maximum, Minimum, Mean, and Median

	Max	Min	Mean	Median
Iterations per task	10	1	3	2

Table 12

Average Iterations and Perceived Success by Query Type and Information Specificity

Query type	Iterations per task (average)	Succeeded tasks	
		n	%
All lookup tasks	3	10	76.92%
Navigational queries specific prior information	1.25	4	100.0%
vague prior information	7.50	1	33.3%
Transactional queries	1.92	5	83.3%

5.4 Condensed contextual analysis

Sequence models were produced from the 13 individual lookup tasks performed during the three CCI sessions, and these were combined into three consolidated sequence models and one list of breakdowns (that is, points where the participant's work was impeded or stopped in some way). The sequence models cover search preparation, entering and updating search terms or facets, and assessing results, and contain a total of 57 steps, 18 breakdowns, and 12 intents. The full list of breakdowns amounts to 22 items.[16]

The consolidated sequence models are presented in the following sections. The full list of breakdowns is included in Appendix F.

5.4.1 Finding a work by its title

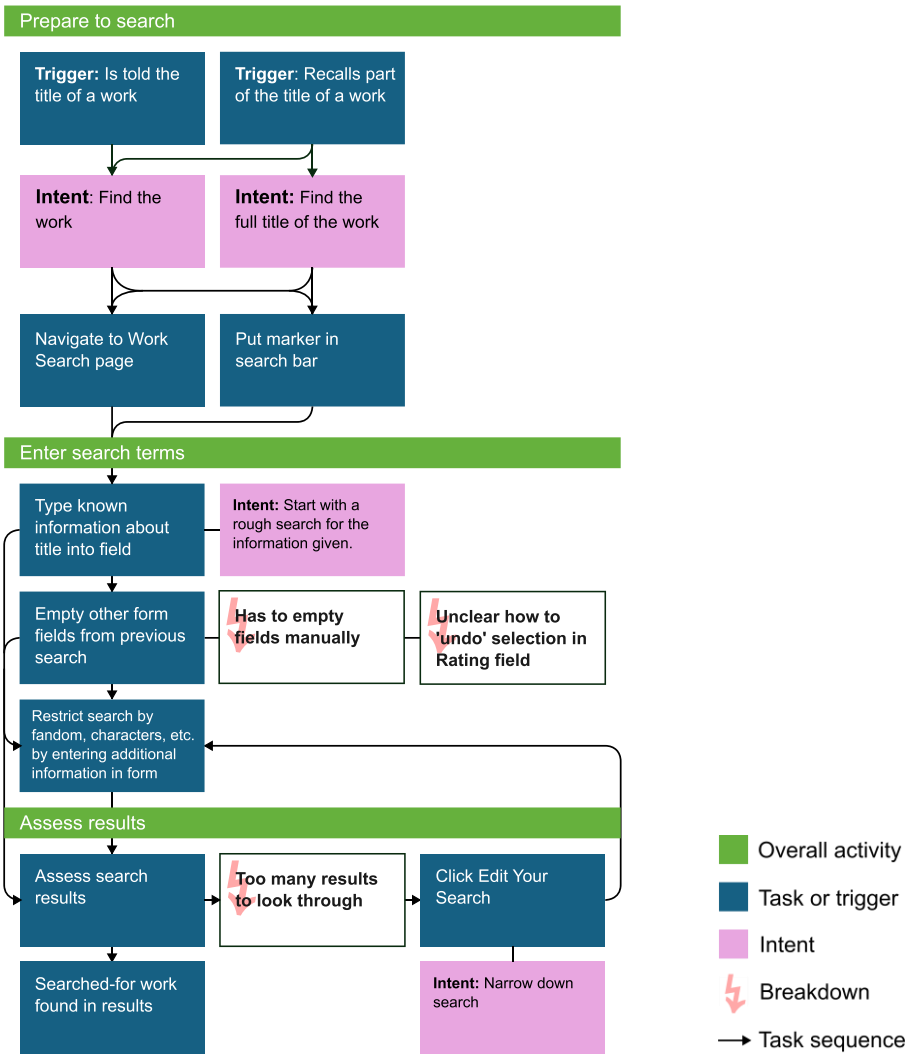
Four separate lookup tasks shared the common goal of finding a fanwork based on a full or partial work title. Participants described this as an easy type of lookup, which is reflected in the simplicity and small number of steps and branches of this sequence model.

All participants used either work search or a combination of work search and the search bar for this task, and all started by simply typing the known information into the search bar or Any Field text box (see Figure A2), reasoning that they would start with a rough, simple search and specify it further only if needed. The majority of these lookup tasks required only a single iteration, as in most cases, a matching work was immediately shown on the search result page. Where a participant needed to narrow down the search, simply adding an appropriate fandom in the Fandoms field was sufficient.

[16] The full list of breakdowns includes every breakdown noted during the lookup tasks, some of which were combined in the consolidated sequence models.

Figure 5.15

Sequence Model: Finding a Work by Title



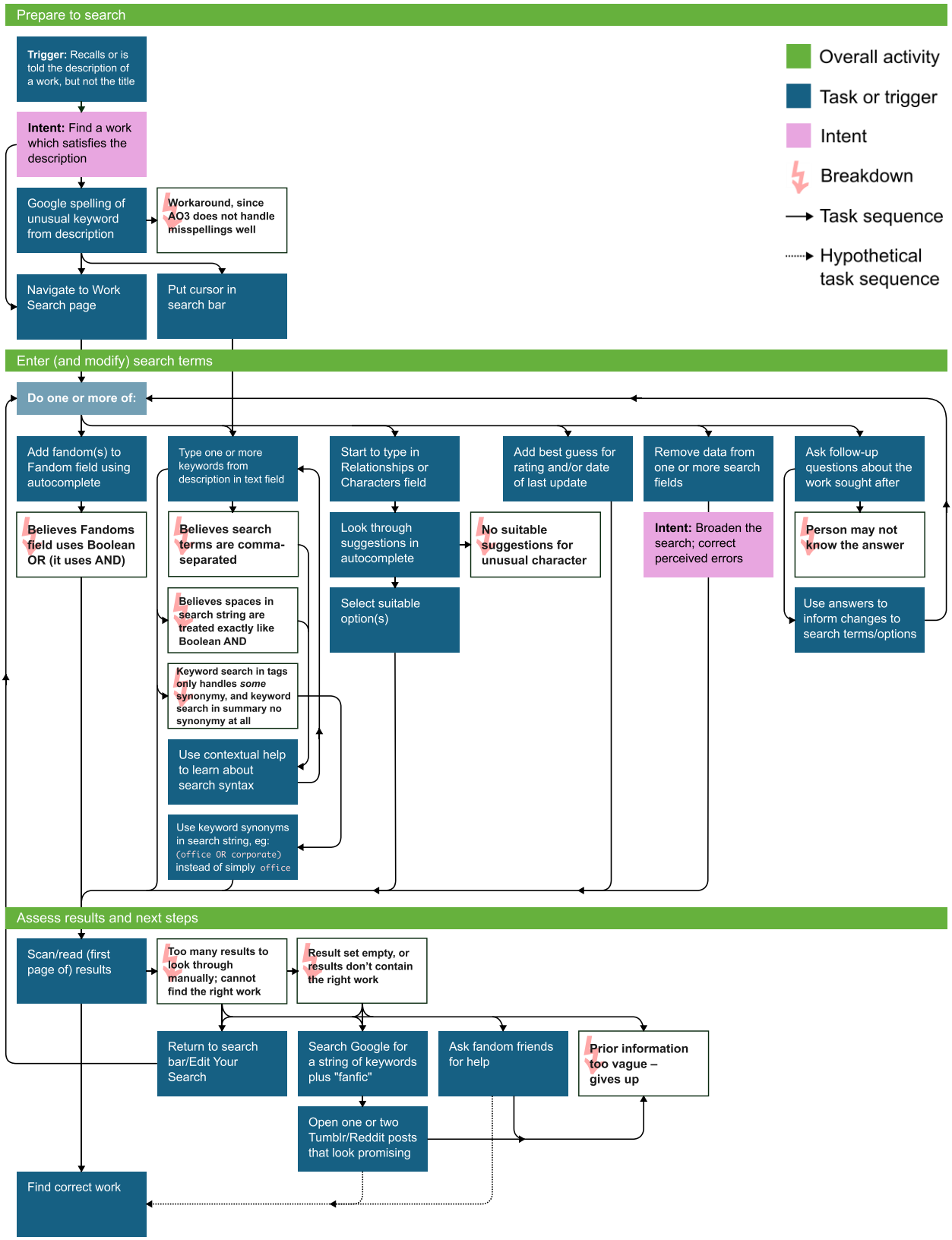
Note. No participants in this study needed to broaden their search during these tasks, and thus, only a "Narrow down search" intent is included under the Assess results activity. In a larger sample, perhaps a "Broaden search" intent would have been needed, as well.

The sequence model in Figure 5.15 shows the two slightly different triggers, all steps taken and intentions expressed, and breakdowns encountered along the way. It divides the task into three overall activities, which summarise the full groups of steps that they contain. Arrows are used to represent motion through the diagram and to show where user decisions, or events otherwise, diverge.

This first model shows that even with this simple type of lookup, participants still encountered problems. Here, for instance, the empty option of the Rating field was not immediately visible to one participant, and another got too many search results to manually go through on the first attempt.

Figure 5.16

Sequence Model: Finding a Work by Description



Note. 'Hypothetical sequences' were only recounted or described hypothetically by participants (e.g. 'Normally, I would…')

5.4.2 Finding a work by description ('lost fic')

Three lookup tasks were consolidated to create this model. In all of them, the participant recalled or had recounted to them a rough description of the plot of a work in the Good Omens fandom, along with its rating and three characters appearing in it, one of whom was an uncommon character in this fandom, who does not appear in the original book or TV show. As can be seen in Figure 5.16, participants all used work search or the search bar, and a variety of different strategies were employed. All participants needed several iterations, and only one succeeded in eventually finding the work – needless to say, this task was perceived as very difficult.

All participants used the auto-complete features in the work search interface, and both interviewees explained that they like to use these whenever possible to avoid any risk of spelling errors and to ensure that they only search for canonical tags. Unfortunately, in some cases, the auto-complete did not provide any suitable suggestions, which led participants to leave out the relevant term completely rather than attempt to type it in manually. Several breakdowns in this model concern misunderstandings about search syntax or the Boolean logic of adding multiple terms to a field – for example, the default search logic is to require *all* search terms to be present, that is, Boolean AND, but two participants expected Boolean OR in the Fandoms field.

5.4.3 Finding something (new) to read or recommend

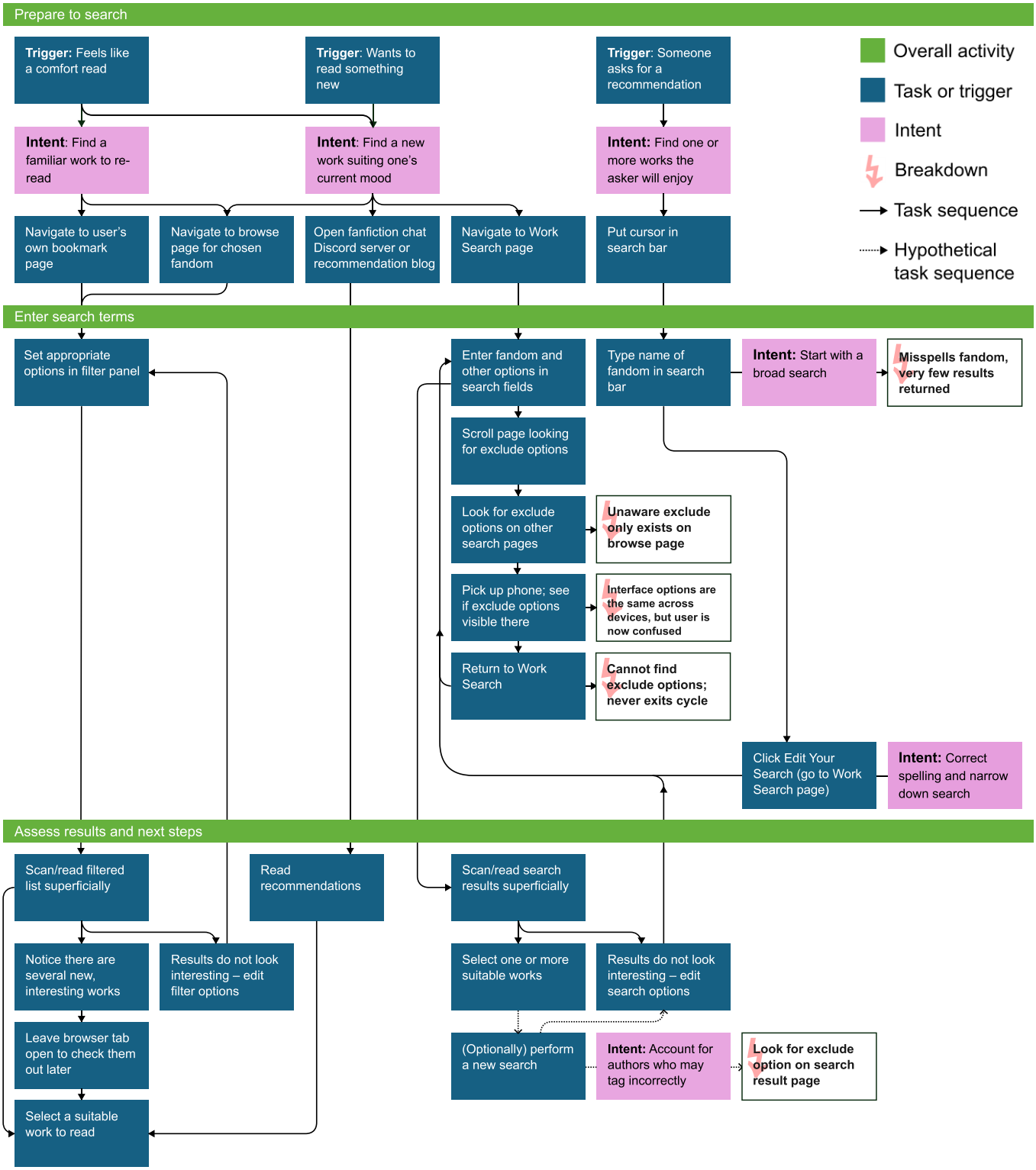
This third sequence model (see Figure 5.17) was consolidated from six lookup tasks, which were all transactional tasks with comparatively vague goals, seeking simply to find a work that the participant or a hypothetical third person might enjoy. Several different search methods were used, with participants choosing between them depending on their familiarity with the fandom (in a familiar fandom, fans might already have their own list of bookmarks or friends to ask for recommendations; in unfamiliar ones, participants instead performed exploratory searches using work search or the browse interface) and the particular information need (when looking for a familiar "comfort read," participants were more likely to look through their own bookmarks or search/browse for a few favourite tags than when looking for something new or searching on behalf of another person).

Interestingly, many breakdowns in these tasks concerned participants wishing to use the filter sidebar to exclude terms, but looking for it in the wrong place. It was not obvious to participants that the filter sidebar only exists on browse result pages and on the user's list of bookmarks, not in result pages for Work Search, for example. As the filter sidebar was brought up several times in the long-form survey responses, this confusion may be shared by many fans beyond the ones observed in this study.

There was also an instance of a participant misspelling the name of a fandom, underlining, according to them, the importance of the auto-complete feature.

Figure 5.17

Sequence Model: Finding Something to Read or Recommend



Note. 'Hypothetical sequences' were only recounted or described hypothetically by participants (e.g. 'Normally, I would…')

6 Discussion

This study set out to investigate the attitudes of fans towards the search and tagging features of AO3, how fans use these features, and whether their attitudes are influenced by any usage or experience factors. This chapter will return to the research questions, summarising and discussing the results of the survey, interviews, and CCI observation sessions as they relate to each one, before turning to a brief reflection on the CCI method and a supplementary point which arose from the qualitative data.

6.1 Fans' attitudes to search and tagging

RQ1, which concerned fans' attitudes to search and tagging on AO3, was investigated in two ways: qualitatively, using the open-ended answers to the survey as well as data from the interviews, and quantitatively, through the constructs of search satisfaction and tagging confidence. The qualitative portion of the study resulted in five themes that are relevant to this research question: *Power and specificity of search*, *Effects of author-tagging, Interface and discoverability, Specific lookup vs. browsing,* and *Appreciation and affect*. It is worth noting that this research question includes aspects of usability, which, although very few respondents comment on it directly, is implied in many responses.

6.1.1 Search satisfaction

The search satisfaction scale was constructed to give a measurement of how satisfied survey respondents are, overall, with AO3's search and browse features. It showed excellent internal reliability (Cronbach's $\alpha = .93$ and item–total correlations between $\rho = .40$ and $\rho = .79$) over the present study's data set, and as it is based on existing generic user satisfaction scales, it is reasonable to conclude (see section 5.1.1) that it has a good construct validity. This means that we can fairly confidently say that the scale measures what we want to measure – and what it measures, in the present study's sample, is that AO3 users are overall quite satisfied with searching the archive, with a median search satisfaction score of 70.5%.

6.1.2 Perceived strengths of and appreciation for AO3 search

The qualitative data nuances this one-dimensional, numerical measurement. On the positive side, the power and level of control granted by AO3's search and browse interfaces were praised by a number of participants, with one likening the power of its search to that of an academic database and several bringing up AO3's very granular filter options as a favourite feature. One participant described the process of figuring out how to best formulate a search using work search as a fun puzzle, and others expressed their appreciation by comparing it favourably to the semantic search of generalist search engines like Google:

> I like the level of control that AO3 search gives to users. Unlike Google search of YouTube search these days, AO3 lets you search for precisely what you want, and respects the search terms you entered.
>
> (R12)

Five separate people specifically mentioned the Exclude filters, which are available on browse result pages, with one participant calling it an absolutely crucial feature which they use every time they search for fanworks. Interestingly, several of Dalton's (2012) participants specifically stated a wish for an exclude feature – which did not exist at the time of administration of Dalton's survey – to be added and for the rest of the filter panel, which had been temporarily removed to conserve system resources, to return (Dalton, 2012, pp. 82–83). Now that both of these features are present, it is encouraging to see that the demand for them that appeared to exist in 2012 has translated to the appreciation seen in the present study.

Many respondents also expressed a general appreciation and love for AO3 and the work of the tag wranglers, including when comparing AO3 to other fanwork platforms:

> I've used both Wattpad and Fanfiction.net before discovering AO3 and there's simply no comparing the three.
>
> (R50)

> I think it's brilliant for finding [works.] Compared to the [Fanfiction.net] days it's a blessing!
>
> (R48)

> Determining synonyms and hierarchies of tags, in the fandom space which is full of super-niche subculture jargon and rapidly-shifting language? That's amazing.
>
> (R12)

> AO3 has the absolute best search feature I know thanks to the fantastic tagging system.
>
> (R14)

Some put this in the context of a love for fandom and the joy it gives them:

> I see two people in a show who have this chemistry and I think, these people should be a couple. Then I look for that, and it exists! Because everyone thinks like this! […] It's a whole new world opening up. It makes me so happy.[1]
>
> (I2)

[1] "Jag ser att två personer i en serie har en sån kemi och man tänker att de här ska bli tillsammans. Då letar jag efter det och så finns det! För alla tycker så här! […] Det är en helt ny värld som öppnas. Man blir ju bara så glad."

This sense of appreciation for AO3 is present across the literature as well, including how it compares positively to other fanwork platforms. Those of Dalton's (2012) survey participants who had used both Fanfiction.net and AO3, for example, overwhelmingly preferred searching on the Archive of Our Own, and Miller (2022) reports that a large majority of her interviewees prefer reading fanfiction on AO3 compared to Fanfiction.net or Tumblr. In Waugh's (2019) study, one participant said that being in the online spaces of her fandom "feels just like home," and over 70% of Dalton's respondents expressed a high overall level of satisfaction with searching on AO3 (Dalton, 2012, p. 45).

6.1.3 Perceived weaknesses of AO3 search

However, participants in the present study were also clear about the frustrations they have or have had with AO3's search system. Thirteen survey respondents, as well as both interviewees, described difficulties when trying to locate specific fanworks, either reasoning that their prior information was too vague to construct a good search or filter query, or stating that minor mistakes in spelling or wording are liable to make searches fail completely. Again, some made the perhaps natural comparison with Google and its better ability to 'guess' what the searcher means:

> The search bar is only useful [for] searches for exact works or authors, but it doesn't always work, in my experience.
>
> (R53)

> I just kind of put some words into the Google bar and trusted Google to sort those words into the different categories for me. Whereas in AO3 […] that work is front loaded onto you, where you have to decide, well, am I going to put [this keyword] as a tag, as "all fields," as title […] knowing that that will yield significantly different results?
>
> (I1)

> It's most useful for well-populated tags and straightforward searches. The more complex one's search needs are, the less likely it is to find results.
>
> (R18)

Meanwhile, in the CCI study, all three participants had significant trouble finding a specific work given only a summary description of the story (the 'lost fic' lookup task), and two did not find a work fitting the description at all. For one participant, this was the only lookup task in which they did not find at least one work satisfying their information need.

Nine survey respondents described having had trouble finding particular features, feeling confusion about the interface, or of having spoken to other fans who have these problems, and all CCI participants[2] got stuck at least once

[2] Including the researcher.

due to a misunderstanding of some aspect of the interface. One interesting portion of the responses on this theme were those that confidently stated that some existing feature was missing, or that some feature worked differently from how it works in reality. One participant wrote,

> I just wish AO3 gave us an "advanced" search option that allowed us to write our own search queries in some type of DQL (Data Query Language) that would give more flexibility (e.g. giving the option to use the "OR" operator) than the [current] search interface allows for.
>
> (R21)

However, AO3's search fields do in fact allow for Boolean search expressions, including the OR operator, as well as advanced syntax for, for example, exclusion of terms or searching in specific metadata fields (see also section 2.3). Other respondents reported noticing *other* fans harbouring this type of misunderstanding:

> I find the search interface fairly easy to use, but I've seen comments on the ao3 subreddit where people struggle with it (eg "I wish we could exclude tags" when the sidebar search has that feature.)
>
> (R32)

Some participants implied that the user themselves has a responsibility for how successful they are at using and searching on AO3, citing the availability of help pages or explaining that they, the respondent, had been too impatient to learn the system 'properly' yet. This is, however, not entirely fair. In the case of the two quotes above, for instance, the information that the search fields allow Boolean operators and other types of advanced syntax is not necessarily easy to find, and while the Exclude feature does indeed exist now, it is only *visibly* present[3] on browse result pages – not on the search result page, which otherwise looks almost exactly the same (see Figure A1 and Figure A3 for comparison). Confusing two nearly identical views for each other or assuming that a feature which is not clearly presented does not exist are certainly not signs of individual failure.

6.1.4 The tagging system

The *Effects of author-tagging* theme shows, on the one hand, an awareness on the part of fans of the strengths and weaknesses of the folksonomic aspect of AO3's tagging system, and on the other, it reveals some more or less controversial topics related to tagging practices, some of which echo the results of previous research.

The fact that works on AO3 are tagged by the authors themselves appears to be common knowledge among the participants of the present study. Many

[3] As previously mentioned, the NOT operator allows for excluding terms when using any of the search fields, but this feature is likely somewhat overly complicated for casual users.

participants brought up risks of author-tagging, some even calling author-tagging the chief weakness of the tagging system. Yet, several respondents also demonstrate an understanding that not all of these risks are possible to mitigate:

> The main problem I have is the tagging is done by the author themselves. If I want to search for a specific idea or tag and the author didn't feel the need to include it (even if it would technically apply) I wouldn't find it. Of course it's not really possible to do any other way.
>
> (R60)

> The search function working as intended wholly depends on writers tagging their fics correctly, so the ability to search certain tags can vary.
>
> (R64)

> The main problem with finding works is authors' tags, not the interface.
>
> (R49)

> I do wish tags were more consistent. It's hard to find things sometimes if you don't already know what tags have been applied to it. However, since these are all self-applied tags I suppose that would be pretty difficult to coordinate.
>
> (R27)

Beyond the fact that authors tag their own works, some respondents expressed confusion about certain aspects of the tagging system. The *Tagging familiarity* questions indicate that many feel they are reasonably or very familiar with most of the topics named in this question, with a median total familiarity score of 9 out of 12. However, the item about the tag hierarchy received markedly fewer "very familiar" responses, with almost half of respondents reporting that they had never heard of it or only just know it exists. This is, to some extent, borne out by the long-form responses, where a few respondents admit outright that they had no idea there was a tag hierarchy until the survey brought it up.

Some participants also state that they "don't really get how [tags] work or what a lot of tags mean" (R27) or that they "struggle pulling out the most important elements in [their] fics to tag" (R63) and often need to ask other fans for help. There was a certain sense of confusion about how, exactly, canonical tags work, with one respondent writing,

> It is hard for me to tag things, but only because I can't always find already existing tags that match what I want to convey.
>
> (R56)

This is an understandable concern – one gets the sense that this creator wants their works to be easily findable, and using canonical tags is a natural step

towards ensuring that. However, it is not in fact necessary to tag a work with canonical tags for that work to be findable with a search which uses canonical tags: as discussed in section 2.3.1, one of the features of the tag wrangling system is that user tags are linked to synonymous canonical tags such that a search for the canonical tag will also bring up results tagged with its synonyms. This subtlety is clearly not known by all users; in fact, even the author was tripped up by this during his CCI session, and one of the interviewees admitted that while they were aware that some tags are canonical and others are not, they were not sure whether or not AO3 allows searching "for just like the random kind of user-created tags" (I1). As for the tag hierarchy, only one out of three CCI participants thought to use the top-level fandom tag "Good Omens - All Media Types" instead of using workarounds to search in the numerous Good Omens fandom tags, and even then, only after several attempts trying other methods.

Previous research, particularly Dalton (2012), reports similar findings; in Dalton's study, even a small number of current or former tag wranglers reported being confused by the tag hierarchy (Dalton, 2012, p. 43). Dalton's respondents also agreed with the present study's respondents on the fact that author-tagging is often inconsistent, which limits the usefulness of the system, with some of Dalton's participants wishing that AO3 would enact stricter tagging guidelines or put more effort into explaining the system to its users.

6.1.5 Tagging confidence

The intent of the *Tagging confidence* Likert set was to measure the level of comfort and confidence with applying tags of respondents who had posted works on AO3 themselves. Of the 25 respondents who answered these questions, the maximum total score was 54.2% and the median and average were both around 33%, which indicates a rather low level of confidence on the part of the participants of this study.

However, the internal reliability of the *Tagging confidence* set is weaker than that of the *Search satisfaction* scale over the present study's data set ($\alpha = .63$ and an average item–total correlation of $\rho = .43$ when the two items with the lowest correlations were removed). The small numbers of items and responses may have made Cronbach's α lower than it might otherwise have been, but one must still be somewhat careful when treating this set as a combined scale and discussing its responses.

Either way, the low average score is supported by the answers to the open survey questions: many of the respondents who discuss their own tagging practices express uncertainty about how to choose tags, with comments such as, "I often struggle pulling out the most important elements in my fics to tag" (R63), or "I can't always find already existing tags that match what I want to

convey" (R56). This gives some indication of the convergent validity of *Tagging confidence*.

An interesting point to note is that the items with the lowest item–total correlation in this set – item 3, "I know where to go if I need more information about tags and tagging," and item 6, "There is a high threshold for new users to learn how to tag" – both concerned *learning* about tagging; that is, both were related to the qualitative category of learning curves (part of the *Interface and discoverability* theme). The ease of learning or finding resources to learn about tagging, then, seems to have had no connection with the present study's participants' confidence in using the tagging system.

6.1.6 Tagging practice, tagging culture

A final aspect of the *Effects of author-tagging* theme concerns tagging practices. One recurring aspect of this is that tags are viewed as a type of content declaration. Fans use tags both as selling points for works they might be interested in and as warnings about works they might want to stay away from, and many appreciate the tagging system for giving them this type of information upfront. Some fans like to use AO3's Exclude filters to avoid tags that they dislike:

> Tags rule! I NEED to know what I'm about to read in order to decide whether the work in queston is worth the effort.
>
> (R48)

> Tags are a good way to get an idea of what kind of fiction it is, and also useful to search for specific things you want to read (or avoid).
>
> (R52)

> Often things I want to remove […] I think [this tag] is so frustrating, I don't understand what people get out of it. Honestly, I've never understood it.[4]
>
> (I2)

Connected to this is the preference, expressed by some participants, that authors should put enough tags on their works to be informative, but not so many as to be overwhelming:

> I fear to use [too] many tags. I don't want to overwhelm the reader. I personally hate to read through a wall of tags on a story. This leads to the problem, that I possible [sic] leave out a tag that don't seem important to me but that would apply to my fic and a different user might want to search for.
>
> (R60)

[4] "Ofta saker som jag vill ta bort […] [Den här taggen] tycker jag är så [jobbig], jag fattar inte vad folk får ut av det. Ärligt talat aldrig begripit mig på det."

> If you come across a severely undertagged fic, it could mean that it's either the best fic you've read in your life, or a terrible fic, there's no in between.
>
> (R5)

> My only grudge is when people do not use tags at all/or are very vague.
>
> (R56)

These attitudes echo the *informed reading* paradigm of Silberstein-Bamford (2024), as well as Miller's (2022) study, where participants said that the wealth of information they can see about a work on AO3 before beginning to read it was an important reason for why they preferred looking for stories to read on AO3 rather than, say, in a conventional library.

Another recurring topic which echoes the results of previous research is the debate around commentary-style or expressive tags (Bourlai, 2018; Gursoy et al., 2018; Kennedy, 2024). Participants in the present study write,

> Love the creativity some authors develop in tagging their works – in some instances, reading the tags [is] just as interesting as reading the stories.
>
> (R21)

> I also love a good silly tag or two, but I know some people find that annoying.
>
> (R32)

> [Tagging] can be veery [sic] complex, but I adore seeing joke tags or personal details that get a laugh out of me being part of the fandom's jokes.
>
> (R82)

> While jokey tags can be fun to read, I think they should be used as rarely as possible on actual fic headers.
>
> (R18)

Just like Kennedy (2024) and Dalton (2012, p. 47) have observed, it is clear that some fans enjoy the practice of commentary tags, while others are not impressed.

6.2 Differences between attitudes of 'specialists' and 'non-specialists'

To answer RQ2, "Are there any significant differences between fans with and without specialist experience in their attitudes to AO3's systems?" survey participants' self-reported specialist experience was correlated against *Search*

satisfaction and *Tagging confidence* scores. The idea that fans with specialist experience of the type measured here – that is, fans who are tag wranglers, fannish librarians, or have professional or academic experience in IT – might score noticeably different on these scales is not borne out by the data. *Specialist experience* shows no correlation with *Search satisfaction* ($\rho = -.01$ when specialist experience is used as a dichotomous variable of whether the respondent has *any* such experience, and $\rho = -.03$ when the number of "yes" answers to *Specialist experience* questions are counted), and only a weak correlation with *Tagging confidence*, which, perhaps interestingly, is negative ($\rho = -.41$ and $\rho = -.40$, respectively). As these numbers are all quite close to zero, any conclusions drawn from them are going to be highly speculative.

6.2.1 The (weak) correlation between tagging confidence and specialist experience

While the correlation between *Search satisfaction* and *Specialist experience* is essentially non-existent in our sample, *Tagging confidence* and *Specialist experience* show a weak negative correlation. Although not enough to claim anything conclusive, one might still speculate about what this could signify if it were to hold true in a larger data set.

The negative sign of this correlation suggests, if it holds, that more specialist experience would be connected to lower *Tagging confidence* and vice versa. On its face, this appears counterintuitive. However, a person with more experience using knowledge organisation systems or who is highly skilled at searching for information might be more aware than someone less experienced of just how much more there is to know, which could cause them to report lower confidence – an effect whose inverse was famously demonstrated by Kruger and Dunning (1999), who showed that to correctly *evaluate* skill in a domain, one must also *possess* a sufficient level of skill in that domain.

6.2.2 Possible causes of the weakness of the correlations

Leaving aside the possibility that the results of this study may be unique to its particular sample, it is worth discussing why it might be that these variables correlate only weakly with each other.

The assumption which motivated this research question was that specialist experience might cause a person to, in some sense, have a better search and tagging experience on AO3. Firstly, "a better search experience" does not necessarily mean a higher search satisfaction, but could also refer to search *success* – whether in terms of precision, recall, or some other measurement related to the search results. A person experienced in searching different databases or expressing information needs with Boolean operators might have a higher rate of search *success* without necessarily being any more *satisfied* with the search tool overall. Bar-Ilan et al. (2012), for example, showed that

user satisfaction, in a general sense, does not necessarily correlate with search recall; they reason that satisfaction might be affected by what sort of interfaces users are used to, so that familiar interfaces might yield higher user satisfaction scores simply by making users feel 'at home.' Furthermore, Ryker et al. (1997) showed that user satisfaction is affected by user *expectations* of a system and by the sources from which they have got these expectations. It is possible that 'expectation effects' might be at play in AO3's case, as well.

A person with more experience using different tools to search for information might also be able to more easily spot problems in a system, which could make them *less* satisfied, despite a high level of search recall or precision. Further affective variables, such as the high level of appreciation for AO3 and love for fandom in general which were observed in the long-form answers and interviews, might have affected fans' reported satisfaction in a positive direction. In some participants' responses, there was a sense that while they may be frustrated by the odd quirk of AO3's systems, they were reluctant to express directly negative opinions, instead making certain to nuance or excuse any negative statement with counter-statements such as,

> I do believe that this issue is completely on me. I took a long time to read the FAQ or to look something up. I took some things at face value.
>
> (R22)

> There's probably a more effective way to do it, but that's what I've always done.
>
> (I1)

On the other hand, a person who does not have any specialist experience cannot necessarily be assumed to be unskilled or unconfident at searching or tagging. Indeed, a hobbyist can be as good as or better in a given field than a skilled professional, as is widely recognised by information science researchers studying serious leisure (e.g. Hartel, 2003). Thus, even if the skills gained from professional or academic experience in LIS or IT are predictive of search success, search satisfaction, tagging skill, or tagging confidence, it may be that those skills can equally well be gained from experiences outwith what is considered here to be 'specialist experience' – including, perhaps, by engaging in fandom itself.

Finally, it is worth noting that in Dalton's (2012) study, respondents who had a background in LIS or had been tag wranglers or otherwise volunteered for AO3 did not report noticeably different levels of satisfaction with AO3's features compared to other respondents (Dalton, 2012, pp. 63–72). Thus, the inconclusive results of the present study appear to agree with the research that exists on this topic.

6.3 Search behaviour and its influence on attitudes

RQ3, "How do fans search AO3 in practice, and does this influence their attitudes to AO3 search?" was investigated using direct survey questions about search methods, qualitative coding of long-form answers and interviews and, most importantly, the three CCI sessions.

6.3.1 Preferred search methods

A majority of survey respondents had tried all three search methods listed (the search bar, the work search page, and browsing by tag or fandom) and almost two thirds preferred browsing, while the second most popular method – the work search page – was the preference of just over 20% of respondents. Many respondents clarified, however, that their choice of search method is dependent on both their information need and the specificity of the information they already have. This was often qualified by a comment on AO3's perceived weakness in locating specific works given vague information:

> The answer to question 1b depends on what the information […] I am searching [for] and what information I already possess. The search bar is only useful to searches for exact works or authors, but it doesn't always work, in my experience.
>
> (R53)

> I tend to use the different options for looking for different things – search bar at [the] top when I know it'll be a small fandom or character list, browse with lots of options selected when I'm in a huge fandom […] and work search when I can half remember a title or author's name.
>
> (R63)

The CCI study echoed this result, with participants choosing different methods for different types of task, sometimes without seeming to consciously think about it. One participant claimed to exclusively browse by fandom when looking for works using AO3 itself, but when asked to find a fanwork with a specific title, they quite naturally navigated to work search. Additionally, both interviewees reported favouring different methods depending on which fandom they wanted to find works within – in some fandoms, they were active in forums or chat groups, where they could browse recommendation threads and ask for tips from other readers or from fan-creators themselves; whereas in fandoms where they were not as socially involved, browsing AO3, reading other users' bookmarks, or following 'tag trails' (that is, rhizomatic reading) were the favoured options.

Interestingly, when Dalton (2012) asked participants what they thought of AO3's various search methods, browsing by fandom was the method with

which the highest proportion of respondents reported being satisfied (Dalton, 2012, p. 37). As has already been mentioned (section 6.1.2), Dalton's respondents also expressed wishes for the then-absent filter panel to return, which also indicates a tendency towards favouring the browse feature, as the filter panel is not present on result pages for the search bar and work search.

In addition to the three search methods listed in the survey, many respondents wrote that they like to look through other users' bookmarks or browse favourite authors' lists of works, or, like the interviewees, described the rhizomatic reading method:

> I almost always click on a tag from something I've already read and then filter out the resulting list to find what I'm looking for […] I can usually find stuff similar to something I've read and enjoyed by clicking on the tags. (Which is the whole point of tags, I suppose)
>
> (R27)

Seven people also wrote about using external, general-purpose search engines, like Google, Ecosia, or DuckDuckGo. Other respondents, however, explained why they thought general-purpose search engines were worse than AO3 (see section 6.1.2), and one interviewee pointed out another potential problem:

> Sometimes I'm brave and I just do a Google search, but [that's] not the most effective way […] It [also] goes into your search history, and there's a stigma against fanfiction.
>
> (I1)

This refers to the fact, mentioned by Dym and Fiesler (2020) and others and well known by many fans, that outsiders to fandom often have negative perceptions of it.

Given the doubts expressed by some participants about using general-purpose external search engines to search for fanworks, it is somewhat interesting that a few respondents still claimed to use this method. Of course, no subculture is a monolith. However, there are problems associated with using external search engines which are unrelated to Google's data protection policies or the efficacy of the search engine itself. When creators post a work to AO3, they are given the option to 'lock' the new work so that it will only be visible to logged-in users. Additionally, AO3 gives users the option to hide all their works from search engines, which leaves works accessible to unregistered users, but programmatically requests that search engines refrain from adding them to their indexes. A search for fanworks using an external search engine will simply not show any locked or hidden works at all, which, of course, reduces the effectiveness of this method.

It is likely that users who still choose to use external search engines either are not aware that this will render many works on AO3 invisible to them, or find

that there are sufficient advantages to these search engines to make up for the problem of hidden works. AO3's perceived weakness on specific lookups with vague information (see section 6.1.3) may be a possible reason for this.

6.3.2 Search strategies

Several participants described modifying search or filter terms iteratively in response to unsatisfying search results:

> Sometimes I run into problems when I look for an author or fic [title]. Here I often have to change some things in my search before I finally find what I was looking for.
>
> (R22)

> It seems easiest to just open the entire fandom I'm interested [in] and then use the filter toolbar on the right to adjust the results until I find what I'm looking for.
>
> (R70)

> I really like doing iterative searches, starting really broad and then narrowing it down. Because if I start too narrow, then where do you go from there?
>
> (I1)

Some argued that searching is not a mechanical process, but in itself something of a creative endeavour:

> The thing is that when you search for tags you have to use your imagination. It's rarely enough to just chose [sic] one and [be] done. You get so many more satisfying results if you're a bit creative.
>
> (R48)

This iterative and creative approach was also demonstrated by all three CCI participants, who needed an average of three iterations per lookup task, and who often tried different approaches when searches did not yield the results they hoped for.

When searching for a work based on a description – the 'lost fic' lookup task, which was described as the most difficult by the interviewees – participants needed a full 7.5 iterations on average and tried a wide range of different approaches (see section 5.4.2 and Figure 5.16). Participants began with either a straight-forward keyword search using the search bar, or by going directly to the work search page and adding text to the Any Fields text box. They then went back and forth between the search result page and work search, adding and removing search options and trying out different syntax and keywords. During the process, participants sometimes 'took a step back' to find more information to help their search process – for example, by asking the researcher for more information, looking up keywords on Google, or reading AO3's context-based help. Upon eventually giving up on the task, one participant said

that had they not been in an interview setting with a time limit, they would likely have reached out to friends or started browsing recommendation blogs to see if they could find someone there who knew the work from the description.

During the 'lost fic' task, the difference between participants did not lie in which overall search *method* they chose: they all ended up, in one way or another, mostly using work search. The specific search *strategies* chosen, however – that is, the search fields used, search terms and syntax, and so on – varied. On the other hand, when performing exploratory-type transactional lookup tasks with the aim of finding one or more (non-specific) works for themselves or a friend fitting only a loose specification (see section 5.4.3 and Figure 5.17), the search *method* chosen also varied widely and seemed to depend partly on the specific information need, partly on in which fandom the participant chose to look for works, and partly on habit and preference. Here, browsing was much more popular, although all three search methods mentioned in the survey were used at some point, as well as recommendations and own bookmarks.

Finally, regardless of the information need or the search method or strategies chosen, all three participants used (and both interviewees expressed appreciation for) the filter panel and tag auto-complete features. Interviewees explained that they use auto-complete to ensure they spell tags and names correctly and to find appropriate canonical terms to search for. Both of the interviewees were aware of the concept of canonical tags and appeared to be implicitly aware that tags can have synonyms, and they seemed to reason that using canonical tags in searches was the 'safest' way to get as many results as possible (that is, to increase recall). However, there were moments when the auto-complete feature failed participants – such as when typing in the name of a very uncommon character. When this happened, participants either tried to work around the problem by entering the name of the uncommon character in Any Fields, or left it out entirely.

6.3.3 Discoverability issues and breakdowns

The CCI participants ran into various problems, of course, and situations similar to some of those problems were also recounted by some survey respondents. In the CCI study, these problems are reported as breakdowns (see Appendix F).

The breakdowns encountered in the CCI study fall into six categories: unsatisfying search results; problems caused by functionality which was understood by participants and was working as intended; problems caused by participants misunderstanding the system; discoverability issues or unclear interfaces; participants needing or believing that they need to use workarounds; and a miscellaneous 'other' category, which includes simple mistakes.

For the purposes of this study, these breakdown categories can be used to discuss fans' expectations of AO3 and to what extent they align with reality.[5] A common subtype of the *misunderstanding* category, which has already been mentioned in this chapter, was the expectation that Exclude filters would be present on all types of search result pages or that they would exist in some form in the work search interface. Work search, however, does not have any obvious or visible options to exclude particular tags or other metadata (although the functionality exists in the form of advanced search syntax – see section 6.1.3), and the Exclude filters, as has been mentioned, only exist on browse pages. It is reasonable to assume that the similarity in appearance between search result pages and browse pages contributes to this expectation; indeed, one participant had significant trouble finding their way to a browse page at one point, and insisted that they had never browsed by tag or fandom despite also reporting being used to the filter sidebar. This confusion is recognised by some survey respondents as well:

> It is counterintuitive that, in order to exclude tags from a search, you have to click "Fandom" and select a fandom tag first, then filter. […] Technically it is possible in the Search Works menu, but it is not discoverable.
>
> (R12)

> I've seen comments on the ao3 subreddit where people struggle with it (eg "I wish we could exclude tags" when the sidebar search has that feature.)
>
> (R32)

The two respondents above are clearly aware of the difference between the different search methods and of which interface is which, but other participants appeared to be unsure about the distinction between work search and browsing, with one describing the browse-by-fandom feature using different words upon being asked if they had used any search methods other than those listed[6] in the survey:

> There is the 'Fandom' possibility in the bar at the left. And then all sorts of fandoms […] by number or alfabetically [sic].
>
> (R79)

Other common misunderstandings leading to CCI breakdowns included misconceptions about the details of search syntax and the Boolean logic of some work search fields. One gets a sense that AO3 has tried to mitigate this by including context-based help and FAQs but, of course, it is a fairly well-

[5] A study with a different focus could have used these results to directly inform decisions about functionality and usability improvements on AO3. A common misunderstanding, for instance, might suggest that parts of the interface need to be made clearer, and a common workaround might be turned into a new feature negating the need for that workaround.
[6] Which, to reiterate, *included* browsing by tag or fandom.

known problem that many web site users do not read manuals (Rogers et al., 2019, Section 7.2.5) – which some of this study's participants recognise (see e.g. R22's remark in section 6.2.2). CCI participants' expectations on the Boolean logic of search fields also appear to vary by field – for example, participants understood that the Characters and Additional Tags fields used AND logic, but mistakenly believed that the Fandoms field used OR logic.

Finally, survey responses as well as CCI sessions showed that while fans are fairly well aware of the basics of how AO3's tagging system works, there are subtleties which are not clear to all AO3 users. Many participants were not very familiar with the tag hierarchy, for instance, and the fact that searches for canonical tags also bring up results tagged with user-created synonyms of those tags seemed unclear to the CCI participants. In participants' defence, there is no way to tell from the work search or browse pages which tags are parent-tags or child-tags of which others. To find this information, one must use the tag search and navigate up and down the hierarchy using a possibly overwhelming interface (see Figure A4 and Figure A5) – and in fact, not every participant even knew that the tag search feature exists or that tags are hierarchical:

> And this thing you showed me,[7] these tags […] I've never even bothered to go in and click those things. Like, I've never thought about the fact that they exist.[8]
>
> (I2)

> I had no idea there was a hierarchy of tags.
>
> (R6)

The auto-completes, which both interviewees as well as several survey respondents reported an appreciation for, do not show any indications of this hierarchy either, but simply display suggestions as a one-dimensional, flat list, ordered by similarity to the characters the user has already typed. It is possible that this contributes to the wider lack of knowledge about the tag hierarchy.

6.3.4 Effects on attitudes

In the present study, no correlation could be found between search method preference or search methods used, on the one hand, and search satisfaction or tagging confidence, on the other. This was, in a large part, due to the small sample size, which made it difficult to apply statistical methods to investigate these correlations (see section 5.1.3).

The CCI study, meanwhile, was far too small to draw any conclusions regarding the effects of actual usage on fans' attitudes. Additionally, both

[7] The participant found this on their own during the session; they are misremembering here.
[8] "Och det här som du visade, de här tagsen […] Jag har aldrig brytt mig om att gå in och klicka på de där sakerna ens en gång. Aldrig liksom tänkt på att de existerar."

interview participants expressed similar attitudes towards AO3, meaning there were no attitude differences in this sample to investigate.

6.4 The influence of other experience factors

RQ4 asks whether experiences related to fandom and AO3 itself have any connection to fans' attitudes towards AO3. Survey participants rated their familiarity with four aspects of AO3's tagging system (the tag hierarchy, relationship tags, canonical tags, and popular genres and tropes) and answered questions on how long they had used AO3 (whether as a registered user or not), how often they use AO3, and whether they had used any other platforms to search for fanworks. Most respondents had used AO3 for 1–5 years, closely followed by more than five years, and a large majority used AO3 several times per week. About half of respondents had used one or two other fanwork platforms, but about 40% had only ever searched for fanworks on AO3. Self-reported *Tagging familiarity* was, as mentioned in section 6.1.4, generally high, except for the tag hierarchy item, where a larger proportion of respondents reported being unfamiliar with the concept, a finding which is reinforced by the results of the qualitative content analysis.

Statistical correlations between these variables and search satisfaction or tagging confidence were weak. The strongest correlation found was between *Tagging familiarity* totals and *Search satisfaction* scores which, at $\rho = .48$, was just low enough to be considered weak in this study (see section 5.1.3). The correlation between *Search satisfaction* scores and the number of other platforms used was $\rho = .41$ and the correlation between *Search satisfaction* scores and the length of time respondents had used AO3 was $\rho = .39$. All other correlations were very weak to negligible.

6.4.1 Caveats to the correlations

It is worth noting that both time using AO3 and number of other platforms used were questions with three options, two of which overwhelmingly dominated the responses: only three respondents had used AO3 for less than a year, and only eight out of the total 62 had used more than two other fanwork platforms. Thus, both variables are, in practice, almost dichotomous in the present study's sample. It is possible that more 'granular' sets of response options (such as allowing participants to give an approximate length of time in years, or the actual number of other platforms they had used) would have yielded different statistical results. However, having to recall these numbers from memory might have led survey respondents to simply choose numbers that 'sound likely,' leading to a higher risk of invalid results. When the question about time using AO3 was posed during the interviews, however, both participants were happy to reminisce about when they had first found their way to the site, relating this to various events in their lives and in the fandoms in which they were active. It is possible that this type of question is simply more suited to a semi-structured

interview context, where the answer can be teased out from a longer thread of reasoning, than a self-administered questionnaire. Similarly, the *Tagging familiarity* total score was a 12-step scale measuring only self-reported knowledge of four different aspects of the tagging system. Asking about further aspects, or measuring this variable in a different way, might have led to different results.

Finally, whenever comparisons are made between many variables, there is a heightened risk that false correlations are found; that is, the risk of a Type I error increases (Crystal & Wildemuth, 2017a, p. 399). As this study compared several different variables to the total scores on *Search satisfaction* and *Tagging confidence*, this risk must be kept in mind when speculating about the meanings of these results.

6.4.2 Possible meanings of the correlations

Either way, although the correlation statistics are too weak to be conclusive, it is, as argued in section 6.2, still worth discussing what they might mean if they do hold.

Seeing a positive correlation between the amount of time using a system and one's satisfaction with that system is, perhaps, not surprising. Ryker et al. showed (1997, p. 535) that expectations formed by users' experiences of using information systems had an effect on user satisfaction, and experience of the very information system under investigation is most likely no exception. The power of the mere-exposure effect (Zajonc, 2001) also cannot be discounted, and, of course, practice using any tool tends to lead to increased skill using that tool. One can expect the same to hold true for AO3's search features.

The positive correlation between the number of other fanwork platforms used and satisfaction with AO3's search is somewhat more interesting. As has previously been noted, participants who mentioned other fanwork platforms in their long-form answers expressed overall positive attitudes to AO3 in comparison to those platforms, saying that there is "simply no comparing" the search features of AO3, Wattpad, and Fanfiction.net (R50), and that compared to using Fanfiction.net, AO3's system is "a blessing" (R63). This echoes results of previous studies; for example those by Dalton (2012) and Miller (2022). If this correlation holds true in general, it could imply that AO3 is just so much better than these other platforms that trying one's luck elsewhere simply underscores the weaknesses of those platforms. However, this is, of course, highly speculative.

Finally, *Tagging familiarity* and *Search satisfaction* showed the strongest correlation of any two variables in this study's data set. It is not difficult to imagine possible reasons for this correlation; indeed, as the tagging system is so central to the functioning of AO3's search system, a general understanding of how tagging works is likely to make searching less confusing. In addition,

when respondents expressed frustration with searching on AO3, several did so citing confusion with the tagging system, as well:

> The main problem I have is the tagging is done by the author themselves. If I want to search for a specific idea or tag and the author didn't feel the need to include it (even if it would technically apply) I wouldn't find it.
>
> (R60)

> I do wish tags were more consistent. It's hard to find things sometimes if you don't already know what tags have been applied to it.
>
> (R27)

It would be interesting to see a deeper investigation into this connection, whether it holds true over a larger sample and across other aspects of the tagging system, and what it might signify.

6.5 The applicability of the CCI method

In addition to the research questions, this study was also interested in trying out methods from HCI in a fan studies/fanLIS research context. Thus, in this section, the data and observations directly related to condensed contextual inquiry as a method will be discussed, as well as some researcher reflections.

As a method of usability evaluation, CCI over video call is a compromise between convenience and the richness of naturalistic observation. Compared to traditional contextual inquiry, the condensed format with more or less well-defined tasks offers some structure and a measure of repeatability, while the still mostly participant-driven master–apprentice model appears to empower participants to share in-the-moment reflections and lines of reasoning which might not have been brought up using a more structured protocol.

The types of insight which might be gained from CI and CCI studies depend on the choice of contextual analysis methods. The sequence model used in the present study offers insight into the 'how' of carrying out different tasks – the sequences of steps and decisions – as well as some insight into intentions and problems that arise along the way. Other models can offer insight into social structures, user identity, and more.[9] All of these can potentially be useful to an information practices researcher.

Both interviewees chose to use the card sorting exercise to share their feedback on the CCI experience. Overall, participants reported enjoying the sessions, choosing word cards such as "organised," "motivating," "comfortable," and

[9] There are some contextual models for which the video call version of CCI is not suitable, such as the physical and artifact models, which do still need the researcher to physically be in the room with the participant.

"friendly." The master–apprentice format was described in positive terms, with participants appearing to enjoy taking on a leading role in the sessions, and explaining their process and reasoning seemed to work well, with neither participant needing much prompting from the researcher to keep talking. This echoes the reflections by Kronberg (2020, p. 39), who notes that his CI participants enjoyed their observation sessions more than those participants who participated in traditional field usability tests, and who theorises that the cause might be CI's stronger focus on the system rather than the participant's performance.

An interesting side-effect of the CCI format was that the experience seemed to lead participants to reflect on their knowledge and usage of the system under investigation, as well as notice features of AO3 that they had not used before:

> The more you learn about it, the more advanced it is. I still have some things that […] I had never looked at even. I've never even bothered to go in and click those things.[10]
>
> (I2)

> I found that as I was describing how to search things, I was like, I wish I knew a better way to do this […] Doing this illuminated, for me, areas in which I can grow or […] learn more on AO3.
>
> (I1)

As we have seen and as other research (e.g. Koven Matasy, 2013; Waugh, 2019) has shown, online fandom is a 'learning-intensive' hobby – it is a serious leisure activity (H. Hill & Pecoskie, 2017; Price & Robinson, 2017, 2021) whose practitioners not only frequently teach and learn from each other, but also enjoy learning and seek it out. The CCI participants in the present study spoke of the realisations and learning opportunities alluded to above in tones of excitement; they were clearly happy to get the chance to learn more about this tool that they use every day.

Participant enjoyment is perhaps not the first consideration brought up in most discussions of scientific method, yet it is implicit in many arguments both ethical and practical. One can very reasonably argue that it ought to be a consideration, at least implicitly, in many types of research, especially when studying serious leisure.

Lastly, from a researcher perspective, it was fascinating both to see the participants' search process 'in action,' as opposed to only having it recounted approximately in an interview setting, and to gain some insight into their thought processes and mental models while they were being observed. The

[10] "Ju mer man lär sig om det, desto mer avancerat är det ju liksom. Vissa saker har jag ju fortfarande som […] jag hade aldrig tittat på ens en gång. Jag har aldrig brytt mig om att gå in och klicka på de där sakerna ens en gång."

latter would, of course, have been difficult had this been a more 'hands-off' observation or a more structured type of user study.

6.6 Fandom as a community

During the initial qualitative content analysis of the open-ended survey responses, a theme of knowledge sharing and norms and expectations in fandom was repeatedly observed. Except for comments about finding fanworks through recommendations, this did not fit neatly into any themes related to the research questions. It was clear, however, that this topic was important enough for participants to bring it up repeatedly, despite the fact that no questions were asked specifically about the fandom experience beyond attitudes to search and tagging. The interviews cemented the necessity of this as a separate theme, as participants spoke extensively about the social side of fandom without being prompted about it:

> I like the community aspect of it. I know it's an archive, which is […] not a social media site, but I do like commenting and leaving kudos on fics, leaving bookmarks […] Sometimes I search just for works in progress, because I actually really enjoy finding works in progress to leave comments and encourage writers.
>
> (I1)

> It's so motivational, I think […] it's so fascinating […] how many wonderfully talented people there are creating things for others. And then […] you feel a bit like you just take and take and take, you never give anything back.[11]
>
> (I2)

> Then as I became personally involved in like Discord servers, definitely, that's when my Marked for Later list became three miles long […] At this point with Good Omens, I mostly just read fics that have been recommended to me via people that I know in my writing groups.
>
> (I1)

That fandom is a social activity is in no way a ground-breaking discovery – indeed, this fact has been central to fan studies since the inception of the field, and much has been written about it. Waugh (2019), for instance, found prevalent themes of relationships and social activities in her data; her observations of a fandom forum showing members often coming online simply to "hang out." Koven Matasy (2013) and Price and Robinson (2017) both describe how being in fandom allows fans to form meaningful relationships,

[11] "Det är så motiverande, tycker jag […] det är så fascinerande […] hur många underbart duktiga folk det finns som gör saker för andra. Och då […] man blir lite såhär att man bara tar och tar och tar, man ger aldrig nånting tillbaka."

even meeting romantic partners, while other authors (e.g. Joyce, 2025; Rebaza, 2009; Russ, 1985/2014; Wagenaar, 2024) show fans using fandom as a safe space to explore stigmatised identities or interests, while actively protecting their own and others' privacy from bullies and outside observers.

That there is stigma associated with fandom was, as has already been discussed, recognised by some of the present study's participants:

> Sometimes I'm brave and I just do a Google search, but […] then it goes into your search history, and there's a stigma against fanfiction. [But] I don't know why I [feel like] that, since I use AO3 on my browser anyway, right?
>
> (I1)

It is also clear that fans like to share knowledge and information with and to learn from each other. This is most obviously expressed by those participants who state that they sometimes, often, or even exclusively find new fanworks using recommendations or bookmarks from other fans:

> I very rarely search for a Good Omens fic using the AO3 browse function unless I'm specifically searching for one that's been recommended […] and I've lost the link.
>
> (I1)

> [I use the] rabbit-hole method of clicking on an author name and scrolling through their works, works in a series, bookmarks, etc.
>
> (R27)

> I usually go to where people share things and I check Discord […] where you can get tips about different [fanfics].[12]
>
> (I2)

Wider themes of fans offering, asking, and receiving help from other fans were also expressed:

> [I had to] figure out which ship specifically I'm interested in […] so then I went to Tumblr to figure out ship names and then came back and searched for that.
>
> (I1)

> In order to exclude tags from a search, you have to click "Fandom" and select a fandom tag first, then filter. I have to explain this to new users a lot.
>
> (R12)

[12] "Jag [går] oftast in på när folk delar olika saker och jag kollar på Discord […] där man får tips om olika [fanfics]."

> I complained about this, how to, like, exclude this kind of thing. And then someone said there are tags that have sub-tags, and […] if you exclude the main tag, sub-tags disappear too. I had no idea about this; it was […] brand new information. That was exciting.[13]
>
> (I2)

> Since it's [the] Good Omens fandom, I would probably get on Discord and be like, "What is this fic? Help me!" And see if anyone else can help.
>
> (I1)

There are of course many dimensions to this theme. In the literature, Hellekson (2009) and Rebaza (2009, Chapter 5) describe fandom itself as a gift culture, with fans often creating fanworks as gifts for other fans and then offering them up for all of fandom to enjoy. Koven Matasy (2013, Section 2.4) and Waugh (2019) describe ways in which fans teach and learn skills within fandom, and themes of information sharing, curation, and organisation recur in Price and Robinson (2017), Fiesler et al. (2016), Kem (2005), and many others. Many of these dimensions could not be investigated using the data from the present study, chiefly because no direct questions were asked about them. However, it is nonetheless interesting that a study which only asked about attitudes to and usage of the search and tagging features of AO3, an online archive – "not a social media site," as emphasised by interviewee 1 – still inspired participants to speak so extensively about their experiences of fandom as a social activity.

[13] "Jag klagade på det här, hur man liksom får bort såna här saker då. Och det var någon som sa att det finns tags som har undertags och […] tar man bort huvudtagen så försvinner undertags också. Det här hade inte jag en aning om; det var […] brand new information. Det var spännande."

7 Conclusions

> You don't know what you're going to find. In fact, you don't know what you're looking for until you find it.
>
> —Antony Beevor[1]

The Archive of Our Own is the largest fanfiction archive on the internet (Rouse & Stanfill, 2025; Stanfill, 2024) and a significant daily part of many of its users' lives. The study of attitudes towards and usage of it, therefore, is directly relevant to well over ten million people (AO3, 2026a, 2026b), and viewed as a search engine with an interesting and somewhat unique knowledge organisation system, studies into AO3's successes and challenges can translate into insight about other information retrieval and knowledge organisation systems and about search strategies more broadly.

Mr Beevor's remark above – although it was uttered in reference to archival research – applies equally well to any research activity, and indeed any type of *search* activity. The present study aimed to investigate fans' attitudes towards the search and tagging systems of AO3, as well as experience and usage factors affecting those attitudes. While no strong correlations between the quantitative variables were found, the results nonetheless corroborate previous research in showing that fans feel a high level of satisfaction with AO3's search features, and that they both appreciate and are occasionally baffled by its tagging system. Some elements of the interface, such as the Exclude filters and auto-complete feature, are highly valued by fans, while other aspects cause confusion, such as the differences between the search and browse result pages and the logic and syntax of the search fields. Overall, however, a strong appreciation for AO3 is abundantly clear across the responses, especially when respondents compare it to other fanwork platforms.

The participants in this study demonstrate a high awareness of the benefits and trade-offs of AO3's curated folksonomy. They realise that authors tagging their own works – that is, the folksonomic aspect of AO3's tagging system – leads to inconsistencies in tagging practice, and some are frustrated by the effect this has on their ability to find fanworks. At the same time, participants recognise that it would be difficult to construct a more consistent tagging system which is as flexible as the current one, and fans who have used other fanwork platforms with freer or more restrictive categorisation systems agree that AO3's hybrid approach is more successful than both extremes; some arguing that the tagging system is what makes its search and browse features so precise and powerful.

Search strategies and preferences vary between participants, but in general, fans seem to prefer browsing over searching, and many respondents like to find

[1] The quote appears in an interview by Deborah Orr, published in The Independent on 6th June, 2009.

fanworks through recommendations from other fans or by reading other users' or authors' bookmarks. Participants tend to choose search methods depending on their information needs and the specificity of the information available and use iterative methods to narrow or broaden searches as needed. Tags and other metadata on fanworks are used as both selling-points and warnings, as well as ways to navigate to other works of interest – that is, participants in this study demonstrate Silberstein-Bamford's (2024) *informed* and *rhizomatic* reading paradigms – and the debate about commentary-style tags which is present elsewhere in the literature (e.g. Kennedy, 2024; Price & Robinson, 2021) is echoed in this study as well.

Finally, even though no questions were asked about the wider social context of or general attitudes towards fandom, many of the fans in this study still took the chance to express their love for the community they are part of and to talk about the social spaces they inhabit, the help they offer and receive from other fans, and the social norms of their fandom(s). Even when it is not the topic of conversation, then, it is obvious that this social aspect is absolutely central to fans' experience of being in online fandom.

7.1 Further research

The attitudes and usage patterns found in the present study point to several possible directions for future research. Firstly, while the information needs of fans searching AO3 largely echo those of patrons searching for leisure reading in library catalogues, library catalogues do not generally have a system of index terms as extensive and flexible as AO3's tagging system. Exploring how folksonomic (or even hybrid folksonomic) tagging could be used to improve library catalogues, perhaps building on the research by Peterson (2009), Holley (2010), and Noorhidawati et al. (2013), could contribute to more precise catalogue search tools, simplified workflows for library staff, and better patron experiences. Additionally, the specific findings regarding discoverability, breakdowns, and fans' favourite features on AO3 could be used to improve both AO3 itself and other types of information retrieval systems.

Secondly, the search satisfaction scale developed here could be reused in other studies where a simple, numerical measure of user 'happiness' is required. Investigating, analogously to Ryker et al. (1997), how fans' expectations affect search satisfaction on AO3 would be an interesting avenue of research; the scale could also be expanded to cover additional facets of satisfaction or search success. Furthermore, improvements to the tagging confidence scale could allow for a deeper study of factors affecting attitudes to tagging on AO3.

Lastly, fan studies, and even more so fanLIS, is a young subject, still finding its methodological footing (Benecchi et al., 2022; Jacobs, 2020). Although information practices are central to the field, it is as yet rare to see methods from human–computer interaction – which are not uncommon in the wider LIS

literature – used in fan studies research. It is the hope of the present author that this study will not be the last to apply contextual inquiry, the user satisfaction concept, or HCI methods in general to research problems in fanLIS.

7.2 Limitations

Sample sizes: The number of participants was limited, particularly in the CCI phase of the study, mainly due to the limited time and resources available. A study with a larger sample might have found statistical correlations that this study did not, and could have made additional qualitative findings.

Population size: All populations that could have been defined for this study (users of AO3; AO3 users in the Good Omens fandom; members of the relevant forums) are difficult to estimate and, in some cases, considerable in size. Thus, even had the sample been larger, it would have been difficult to make any estimations regarding statistical significance.

Team size and researcher experience: Contextual inquiry manuals recommend working in teams (see section 4.3.3), and having multiple coders during qualitative content analysis is important for ensuring inter-observer reliability. The author also lacks experience in these methods, which will naturally impact what is noticed during analysis.

Limited pilot testing: Related to the time and resource limitations, the survey would have benefitted from a second round of pilot testing, and a pilot test of the CCI session protocol would also have been prudent. It is possible, thus, that issues with the questions which could otherwise have been noticed and remedied, were not.

References

Acafan. (2025, December 11). Fanlore. https://www.fanlore.org/w/index.php?title=Acafan&oldid=2908291

Archive of Our Own. (n.d.). *Tags FAQ*. Retrieved 4 June 2026, from https://archiveofourown.org/faq/tags?language_id=en

Archive of Our Own. (2026a, January 6). *AO3 statistics – 2025 update*. AO3 News. https://archiveofourown.org/admin_posts/33966

Archive of Our Own. (2026b, March 6). *AO3 celebrates 17 million fanworks*. AO3 News. https://archiveofourown.org/admin_posts/34516

Bacon-Smith, C. (1992). *Enterprising women: Television fandom and the creation of popular myth*. University of Pennsylvania Press.

Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, *29*(5), 530–545. https://doi.org/10.1287/mnsc.29.5.530

Bar-Ilan, J., Zhitomirsky-Geffet, M., Miller, Y., & Shoham, S. (2012). Tag-based retrieval of images through different interfaces: A user study. *Online Information Review*, *36*(5), 739–757. https://doi.org/10.1108/14684521211276019

Bawden, D., & Robinson, L. (2022). *Introduction to information science* (2nd edn). Facet Publishing.

Benecchi, E., Porlezza, C., & Pranteddu, L. (2022). Filling the gap: An exploration into the theories and methods used in fan studies. *Transformative Works and Cultures*, *37*. https://doi.org/10.3983/twc.2022.2243

Benedek, J., & Miner, T. (2002). Measuring desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of UPA 2002 Conference*. Usability Professional Association Conference, Orlando, Florida.

Bennett, L. (2017). Surveying fandom: The ethics, design, and use of surveys in fan studies. In M. A. Click & S. Scott (Eds), *The Routledge companion to media fandom* (pp. 36–44). Routledge. https://doi.org/10.4324/9781315637518

Beyer, H., & Holtzblatt, K. (1998). *Contextual design: Defining customer-centered systems*. Morgan Kaufmann.

Bikos, L. H. (2024). *ReCentering psych stats: Psychometrics*. Seattle Pacific University Library. https://digitalcommons.spu.edu/open_books/6

Bourlai, E. E. (2018). 'Comments in tags, please!': Tagging practices on Tumblr. *Discourse, Context & Media, Discourse of Social Tagging*, *22*, 46–56. https://doi.org/10.1016/j.dcm.2017.08.003

Brooke, J. (1996). SUS - A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis. https://doi.org/10.1201/9781498710411

Bryman, A., Clark, T., Foster, L., & Sloan, L. (2021). *Bryman's social research methods* (6th edn). Oxford University Press.

Buckland, M. (2017). *Information and society*. MIT Press. https://doi.org/10.7551/mitpress/10922.001.0001

Bullard, J. (2018). Curated folksonomies: Three implementations of structure through human judgment. *Knowledge Organization*, *45*(8), 643–652. https://doi.org/10.5771/0943-7444-2018-8-643

Busse, K. (2017). The ethics of studying online fandom. In M. A. Click & S. Scott (Eds), *The Routledge companion to media fandom* (pp. 9–17). Routledge. https://doi.org/10.4324/9781315637518

Busse, K., & Hellekson, K. (2006). Introduction: Work in progress. In K. Hellekson & K. Busse (Eds), *Fan fiction and fan communities in the age of the internet: New essays* (pp. 5–32). McFarland.

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Ceri, S., Bozzon, A., Brambilla, M., Della Valle, E., Fraternali, P., & Quarteroni, S. (2013). *Web information retrieval*. Springer.

Chen, R., Mias, G. I., Li-Pook-Than, J., Jiang, L., Lam, H. Y. K., Chen, R., Miriami, E., Karczewski, K. J., Hariharan, M., Dewey, F. E., Cheng, Y., Clark, M. J., Im, H., Habegger, L., Balasubramanian, S., O'Huallachain, M., Dudley, J. T., Hillenmeyer, S., Haraksingh, R., … Snyder, M. (2012). Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes. *Cell*, *148*(6), 1293–1307. https://doi.org/10.1016/j.cell.2012.02.009

Chin, M. G., Fisak Jr., B., & Sims, V. K. (2002). Development of the Attitudes Toward Vegetarians Scale. *Anthrozoös*, *15*(4), 332–342. https://doi.org/10.2752/089279302786992441

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Choemprayong, S., & Wildemuth, B. M. (Eds). (2017). Developing new measures. In *Applications of social research methods to questions in information and library science* (2nd edn, pp. 291–304). Libraries Unlimited.

Click, M. A., & Scott, S. (2017). Introduction. In M. A. Click & S. Scott (Eds), *The Routledge companion to media fandom* (pp. 1–5). Routledge. https://doi.org/10.4324/9781315637518

Coppa, F. (2006). A brief history of media fandom. In K. Hellekson & K. Busse (Eds), *Fan fiction and fan communities in the age of the internet: New essays* (pp. 41–59). McFarland.

Crystal, A. J., & Wildemuth, B. M. (2017a). Comparing means: T tests and analysis of variance. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 395–404). Libraries Unlimited.

Crystal, A. J., & Wildemuth, B. M. (2017b). Naturalistic research. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 60–70). Libraries Unlimited.

Dalton, K. L. (2012). *Searching the Archive of Our Own: The usefulness of the tagging structure* [Master's thesis, University of Wisconsin-Milwaukee]. UW Milwaukee Electronic Theses and Dissertations. http://digital.library.wisc.edu/1793/92558

David, L. A., Materna, A. C., Friedman, J., Campos-Baptista, M. I., Blackburn, M. C., Perrotta, A., Erdman, S. E., & Alm, E. J. (2014). Host lifestyle affects human microbiota on daily timescales. *Genome Biology*, *15*(7), R89. https://doi.org/10.1186/gb-2014-15-7-r89

Derecho, A. (2006). Archontic literature: A definition, a history, and several theories of fan fiction. In K. Hellekson & K. Busse (Eds), *Fan fiction and fan communities in the age of the internet: New essays* (pp. 61–78). McFarland.

Dinneen, J. D., Asadi, B., Frissen, I., Shu, F., & Julien, C.-A. (2018). Improving Exploration of Topic Hierarchies: Comparative Testing of Simplified Library of Congress Subject Heading Structures. *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval, CHIIR '18*, 102–109. https://doi.org/10.1145/3176349.3176385

Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. *MIS Quarterly*, *12*(2), 259–274. https://doi.org/10.2307/248851

Donnelly, T., Rahman-Vyas, H., Mandelbaum, J., Hatch, M., & Bhattacharya, B. (2021). Virtual observation: Conducting contextual inquiry remotely. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *10*(1), 233–237. https://doi.org/10.1177/2327857921101213

Duggan, J. (2023). Trans fans and fan fiction: A literature review. *Transformative Works and Cultures*, *39*. https://doi.org/10.3983/twc.2023.2309

Dym, B., & Fiesler, C. (2020). Ethical and privacy considerations for research using online fandom data. *Transformative Works and Cultures*, *33*. https://doi.org/10.3983/twc.2020.1733

Fiesler, C., Morrison, S., & Bruckman, A. S. (2016). An archive of their own: A case study of feminist HCI and values in design. *34th Annual CHI Conference on Human Factors in Computing Systems*, 2574–2585. https://doi.org/10.1145/2858036.2858409

Flegel, M., & Roth, J. (2014). Legitimacy, validity, and writing for free: Fan fiction, gender, and the limits of (unpaid) creative labor. *Journal of Popular Culture*, *47*(6), 1092–1108. https://doi.org/10.1111/jpcu.12198

Garcia, P., & Cifor, M. (2019). Expanding Our Reflexive Toolbox: Collaborative Possibilities for Examining Socio-Technical Systems Using Duoethnography. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–23. https://doi.org/10.1145/3359292

Given, L. M., Case, D. O., & Willson, R. (2023). *Looking for information: Examining research on how people engage with information* (5th edn). Emerald Publishing Limited. https://doi.org/10.1108/S2055-5377202315

Gursoy, A., Wickett, K., & Feinberg, M. (2018). Understanding tag functions in a moderated, user-generated metadata ecosystem. *Journal of*

Documentation, *74*(3), 490–508. https://doi.org/10.1108/JD-09-2017-0134

Hale, L. M. (2005, November 22). *History of fan fiction*. The Fanfic Symposium. https://web.archive.org/web/20051122175602/http://www.trickster.org/symposium/symp173.htm

Hank, C., Wilkins Jordan, M., & Wildemuth, B. M. (2017). Survey research. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 272–283). Libraries Unlimited.

Hart, C., Schoolbred, M., Butcher, D., & Kane, D. (1999). The bibliographical structure of fan information. *Collection Building*, *18*(2), 81–89. https://doi.org/10.1108/01604959910265869

Hartel, J. (2003). The Serious Leisure Frontier in Library and Information Science: Hobby Domains. *Knowledge Organization*, *30*(3–4), 228–238. https://doi.org/10.5771/0943-7444-2003-3-4-228

Hartson, R., & Pyla, P. S. (2012). *The UX book: Process and guidelines for ensuring a quality user experience*. Morgan Kaufmann.

Hellekson, K. (2009). A fannish field of value: Online fan gift culture. *Cinema Journal*, *48*(4), 113–118. https://doi.org/10.1353/cj.0.0140

Hellekson, K., & Busse, K. (2014). Introduction: Why a fan fiction studies reader now? In K. Hellekson & K. Busse (Eds), *The fan fiction studies reader* (pp. 1–18). University of Iowa Press. https://doi.org/10.2307/j.ctt20p58d6

Hill, C. D. (2024). Developing a working taxonomy for pan-fandom information behavior comparison. *Transformative Works and Cultures*, *43*. https://doi.org/10.3983/twc.2024.2443

Hill, H., & Pecoskie, J. J. (2017). Information activities as serious leisure within the fanfiction community. *Journal of Documentation*, *73*(5), 843–857. https://doi.org/10.1108/JD-04-2016-0045

Holley, R. (2010). Tagging full text searchable articles: An overview of social tagging activity in historic Australian newspapers August 2008 – August 2009. *D-Lib Magazine*, *16*(1/2). https://doi.org/10.1045/january2010-holley

Holtzblatt, K., & Beyer, H. (2016). *Contextual design: Design for life* (2nd edn). Elsevier Science & Technology.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Ives, B., Olson, M. H., & Baroudi, J. J. (1983). The measurement of user information satisfaction. *Communications of the ACM*, *26*(10), 785–793. https://doi.org/10.1145/358413.358430

Jacobs, N. (2020). Interdisciplinary methodologies for the fan studies bricoleur. *Transformative Works and Cultures*, *33*. https://doi.org/10.3983/twc.2020.1665

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Jenkins, H. (1992). *Textual poachers: Television fans and participatory culture*. Routledge.

Johnson, S. F. (2014). Fan fiction metadata creation and utilization within fan fiction archives: Three primary models. *Transformative Works and Cultures*, *17*. https://doi.org/10.3983/twc.2014.0578

Joyce, C. (2025). 'I wish my life was like this': Queer identity work in BL fandom. *Transformative Works and Cultures*, *46*. https://doi.org/10.3983/twc.2025.2743

Kantner, L., Sova, D. H., & Rosenbaum, S. (2003). Alternative methods for field usability research. *Proceedings of the 21st Annual International Conference on Documentation, SIGDOC '03*, 68–72. https://doi.org/10.1145/944868.944883

Kaplan, D. (2006). Construction of fan fiction character through narrative. In K. Hellekson & K. Busse (Eds), *Fan fiction and fan communities in the age of the internet: New essays* (pp. 134–152). McFarland.

Kasapović, I. (2025). Choice or inevitability: Social context and the role of cultural capital in information behaviour of young people in career decision-making. *Journal of Documentation*, *81*(4), 1059–1077. https://doi.org/10.1108/JD-01-2025-0027

Kem, J. F. (2005). *Cataloging the Whedonverse: Potential roles for librarians in online fan fiction* [Master's thesis, University of North Carolina at Chapel Hill]. https://doi.org/10.17615/k63q-9m78

Kennedy, K. (2024). 'It's not your Tumblr': Commentary-style tagging practices in fandom communities. *Transformative Works and Cultures*, *42*. https://doi.org/10.3983/twc.2024.2475

Koven Matasy, V. R. (2013). *Fannish librarians: The intersection of fandom and library and information science* [Master's thesis, University of North Carolina at Chapel Hill]. https://doi.org/10.17615/1827-tz31

Kronberg, M. (2020). *A comparison between field user research methods in a redesign context* [Master's thesis, Uppsala University]. DiVA. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-404655

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Kuniavsky, M. (2003). *Observing the user experience: A practitioner's guide to user research*. Elsevier Science & Technology.

Lamb, P. F., & Veith, D. L. (2014). Romantic myth, transcendence, and Star Trek zines. In K. Hellekson & K. Busse (Eds), *The fan fiction studies reader* (pp. 97–115). University of Iowa Press. https://doi.org/10.2307/j.ctt20p58d6 (Reprinted from *Erotic universe: Sexuality and fantastic literature*, pp 236-255, by D. Palumbo, Ed., 1986, Greenwood Press)

Large, A., Beheshti, J., Nesset, V., & Bowler, L. (2004). Designing web portals in intergenerational teams: Two prototype portals for elementary school students. *Journal of the American Society for Information Science and Technology*, *55*(13), 1140–1154. https://doi.org/10.1002/asi.20069

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), 57–78. https://doi.org/10.1080/10447319509526110

Lewis, L. A. (Ed.). (1992). *The adoring audience: Fan culture and popular media*. Routledge.

Lichtenberg, J., Marshak, S., & Winston, J. (1975). *Star Trek lives!* Bantam Books.

Luo, L., & Wildemuth, B. M. (2017). Delphi studies. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 81–90). Libraries Unlimited.

Maish, A. M. (1979). A user's behavior toward his MIS. *Management Information Systems Quarterly*, *3*(1), 39–52. https://doi.org/10.2307/249147

Marcengo, A., Rapp, A., Cena, F., & Geymonat, M. (2016). The Falsified Self: Complexities in Personal Data Collection. In M. Antona & C. Stephanidis (Eds), *Universal Access in Human-Computer Interaction. Methods, Techniques, and Best Practices* (pp. 351–358). Springer Publishing. https://doi.org/10.1007/978-3-319-40250-5_34

Marsh, A., & Lerner, A. (2023). Privacy norms of transformative fandom: A case study of an activity-defined community. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW1), 1–29. https://doi.org/10.1145/3637388

Miller, J. N. (2022). Information-seeking behaviors of young adult readers of fiction and fan fiction. *Transformative Works and Cultures*, *37*. https://doi.org/10.3983/twc.2022.2245

Neustaedter, C., & Sengers, P. (2012). Autobiographical design: What you can learn from designing for yourself. *Interactions*, *19*(6), 28–33. https://doi.org/10.1145/2377783.2377791

Neville, S. (2004). Seeing the forest and the trees when developing a new acquisitions system. *The Serials Librarian*, *46*(1–2), 155–159. https://doi.org/10.1300/J123v46n01_16

Nielsen, J. (1993). *Usability engineering*. Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/C2009-0-21512-1

Noorhidawati, A., Fariza Hanum, N., & Zohoorian-Fooladi, N. (2013). Social tagging in a scholarly digital library environment: Users' perspectives. *Information Research*, *18*(3), C43. 8th International Conference on Conceptions of Library and Information Science.

Norman, D. (1988). *The psychology of everyday things*. Basic Books.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

O'Kane, A. A., Rogers, Y., & Blandford, A. E. (2014). Gaining empathy for non-routine mobile device use through autoethnography. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, 987–990. https://doi.org/10.1145/2556288.2557179

Organization for Transformative Works. (n.d.). *OTW history*. Retrieved 3 April 2026, from https://www.transformativeworks.org/about_otw/our-history/

Peterson, E. (2009). Patron preferences for folksonomy tags: Research findings when both hierarchial subject headings and folksonomy tags are used. *Evidence Based Library and Information Practice*, *4*(1), 53–56.

Pilerot, O., & Lindberg, J. (2024). Ubiquitous but invisible – public librarians' self-imposed professional information practices as articulation work. *Information Research an International Electronic Journal*, *29*(2), 483–494. https://doi.org/10.47989/ir292842

Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., Gorgolewski, K. J., Luci, J., Joo, S. J., Boyd, R. L., Hunicke-Smith, S., Simpson, Z. B., Caven, T., Sochat, V., Shine, J. M., Gordon, E., Snyder, A. Z., Adeyemo, B., Petersen, S. E., … Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, *6*(1), 8885. https://doi.org/10.1038/ncomms9885

Popova, M. (2020). Follow the trope: A digital (auto)ethnography for fan studies. *Transformative Works and Cultures*, *33*. https://doi.org/10.3983/twc.2020.1697

Price, L., & Robinson, L. (2017). 'Being in a knowledge space': Information behaviour of cult media fan communities. *Journal of Information Science*, *43*(5), 649–664. https://doi.org/10.1177/0165551516658821

Price, L., & Robinson, L. (2021). Tag analysis as a tool for investigating information behaviour: Comparing fan-tagging on Tumblr, Archive of Our Own and Etsy. *Journal of Documentation*, *77*(2), 320–358. https://doi.org/10.1108/JD-05-2020-0089

Pugh, S. (2004). The democratic genre: Fan fiction in a literary context. *Refractory: A Journal of Entertainment Media*, *5*. https://refractory-journal.com/the-democratic-genre-fan-fiction-in-a-literary-context-sheenagh-pugh/

Pugh, S. (2005). *The democratic genre: Fan fiction in a literary context*. Seren Books.

Rebaza, C. M. (2009). *The modern coterie: Fan fiction writers as a community of practice* [Doctoral thesis, University of Illinois at Urbana-Champaign]. https://hdl.handle.net/2142/18688

Rickards, G., Magee, C., & Artino, A. R. (2012). You can't fix by analysis what you've spoiled by design: Developing survey instruments and collecting validity evidence [Editorial]. *Journal of Graduate Medical Education*, *4*(4), 407–410. https://doi.org/10.4300/JGME-D-12-00239.1

Robinson, L., & Price, L. (2025). Five years of FanLIS: Papers from the FanLIS 2025 symposium. *Proceedings from the Document Academy*, *12*(1), Article 1. https://doi.org/10.35492/docam/12/1/1

Rogers, Y., Sharp, H., & Preece, J. (2019). *Interaction Design: Beyond Human-Computer Interaction* (5th edn). Wiley.

Rosenbaum, S., & Kantner, L. (2007). Field usability testing: Method, not compromise. *2007 IEEE International Professional Communication Conference*, 1–7. https://doi.org/10.1109/IPCC.2007.4464060

Rouse, L., & Stanfill, M. (2025). Quantifying fandom: Identity, romance, and platform preferences on Archive of Our Own. *New Media & Society*, 1–22. https://doi.org/10.1177/14614448251367033

Ruel, E., Wagner, W. E., & Gillespie, B. J. (2016). Pretesting and pilot testing. In E. Ruel, W. E. Wagner, & B. J. Gillespie (Eds), *The practice of survey research: Theory and applications* (pp. 101–119). SAGE Publications, Inc. https://doi.org/10.4135/9781483391700

Russ, J. (2014). Pornography by women for women, with love. In K. Hellekson & K. Busse (Eds), *The fan fiction studies reader* (pp. 82–96). University of Iowa Press. https://doi.org/10.2307/j.ctt20p58d6 (Reprinted from *Magic mommas, trembling sisters, puritans & perverts: Feminist essays*, pp 79-99, by J. Russ, 1985, Crossing Press)

Rutkowski, C. (1982, October). An introduction to the human applications standard computer interface part 1: Theory and principles. *BYTE Magazine*, *7*(10), 291–310.

Ryker, R., Nath, R., & Henson, J. (1997). Determinants of computer user expectations and their relationships with user satisfaction: An empirical study. *Information Processing & Management*, *33*(4), 529–537. https://doi.org/10.1016/S0306-4573(97)00016-2

Sandvoss, C. (2014). The death of the reader? Literary theory and the study of texts in popular culture. In K. Hellekson & K. Busse (Eds), *The fan fiction studies reader* (pp. 61–74). University of Iowa Press. https://doi.org/10.2307/j.ctt20p58d6 (Reprinted from *Fandom: Identities and communities in a mediated world*, pp 27-40, by J. Gray, C. L. Harrington, & C. Sandvoss, Eds., 2007, New York University Press)

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350–353. https://doi.org/10.1037/1040-3590.8.4.350

Schwartz, N. (2019). Surveys, experiments, and the psychology of self-report. In F. R. Kardes, P. M. Herr, & N. Schwarz (Eds), *Handbook of research methods in consumer psychology* (pp. 34–68). Routledge.

Silberstein-Bamford, F. (2024). "Thank god for tags"—Fanfiction as a reading paradigm. *New Review of Hypermedia & Multimedia*, *30*(1–2), 129–147. https://doi.org/10.1080/13614568.2024.2369508

Smith, L. J. (2016). *Information behaviors of ethical hackers* [Master's thesis, University of North Carolina at Chapel Hill]. https://doi.org/10.17615/rpa6-tj65

Stanfill, M. (2024). The platformization of fandom and its discontents: Understanding platform harms through the Archive of Our Own. *International Journal of Communication*, *18*, 4209–4226.

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Swanson, E. B. (1974). Management information systems: Appreciation and involvement. *Management Science*, *21*(2), 178–188. https://doi.org/10.1287/mnsc.21.2.178

Swedish Research Council. (2025). *Good research practice 2024*. https://www.vr.se/english/analysis/reports/our-reports/2025-07-03-good-research-practice-2024.html

Tullis, T. S., & Stetson, J. N. (2004). *A Comparison of questionnaires for assessing website usability*. Usability Professional Association Conference.

Tushnet, R. (1997). Legal fictions: Copyright, fan fiction, and a new common law. *Loyola of Los Angeles Entertainment Law Review*, *17*(3), 651–686.

Tushnet, R. (2007). Copyright Law, Fan Practices, and the Rights of the Author. In J. Gray, C. Sandvoss, & C. L. Harrington (Eds), *Fandom: Identities and communities in a mediated world* (pp. 68–80). New York University Press.

Wagenaar, W. F. (2024). Discord as a fandom platform: Locating a new playground. *Transformative Works and Cultures*, *42*. https://doi.org/10.3983/twc.2024.2473

Waugh, A. (2019). Feels like home: The digital information practices of teen fans. *The Journal of Research on Libraries and Young Adults*, *10*(2).

https://www.yalsa.ala.org/jrlya/2019/07/feels-like-home-the-digital-information-practices-of-teen-fans/

Wildemuth, B. M. (2017a). Correlation. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 387–394). Libraries Unlimited.

Wildemuth, B. M. (2017b). Descriptive Statistics. In B. M. Wildemuth (Ed.), *Applications of Social Research Methods to Questions in Information and Library Science* (2nd edn, pp. 351–360). Libraries Unlimited.

Wildemuth, B. M. (2017c). Frequencies, Cross-tabulation, and the Chi-square statistic. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 361–372). Libraries Unlimited.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual design: An emergent view of system design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '90*, 329–336. https://doi.org/10.1145/97243.97304

Yin, K., Aragon, C., Evans, S., & Davis, K. (2017). Where no one has gone before: A meta-dataset of the world's largest fanfiction repository. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17*, 6106–6110. https://doi.org/10.1145/3025453.3025720

Zajonc, R. B. (2001). Mere Exposure: A Gateway to the Subliminal. *Current Directions in Psychological Science*, *10*(6), 224–228. https://doi.org/10.1111/1467-8721.00154

Zhang, Y., & Wildemuth, B. M. (2017). Qualitative analysis of content. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2nd edn, pp. 318–329). Libraries Unlimited.

Appendix A
Additional screenshots

Figure A1

Search Result Page on AO3

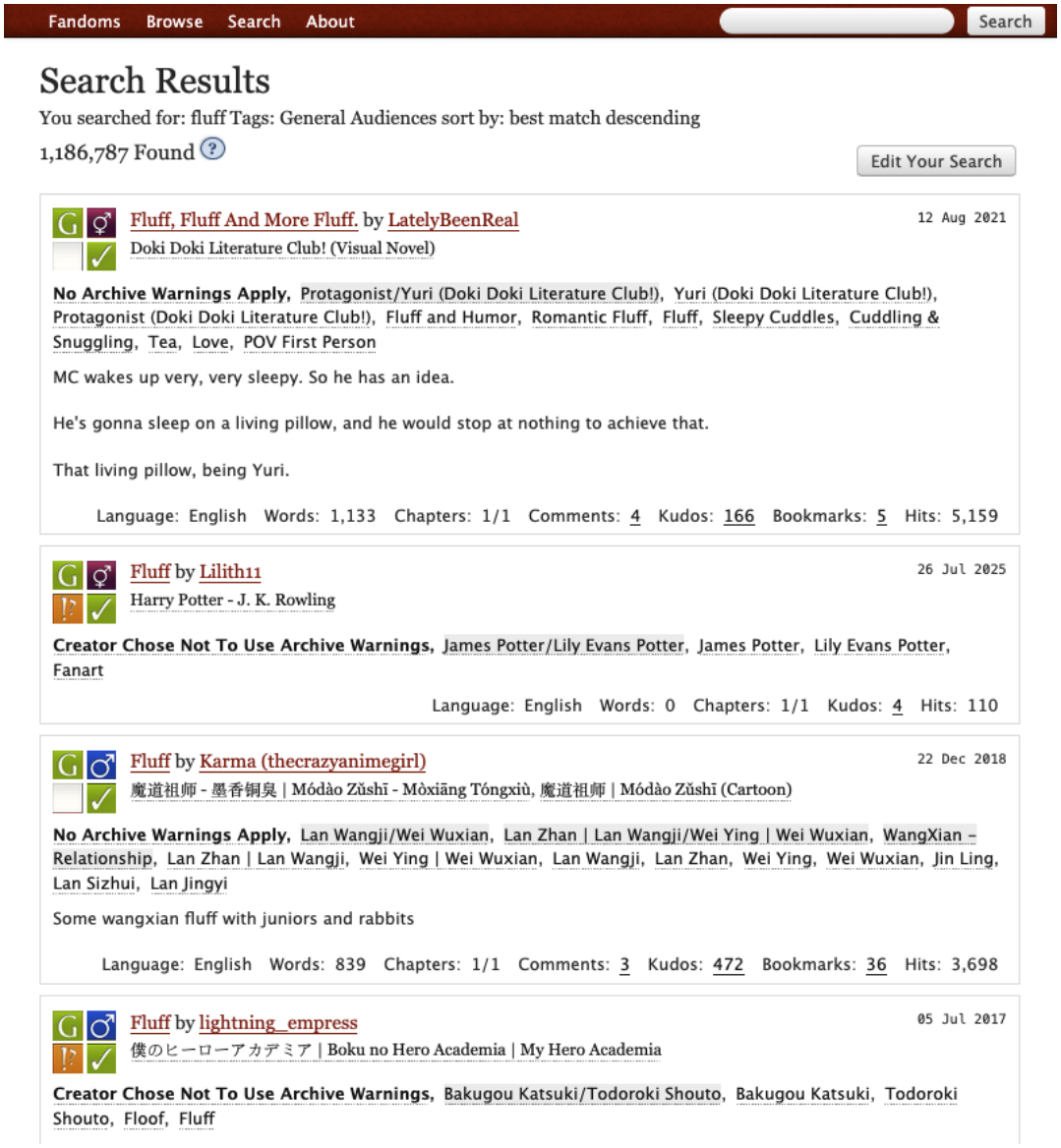
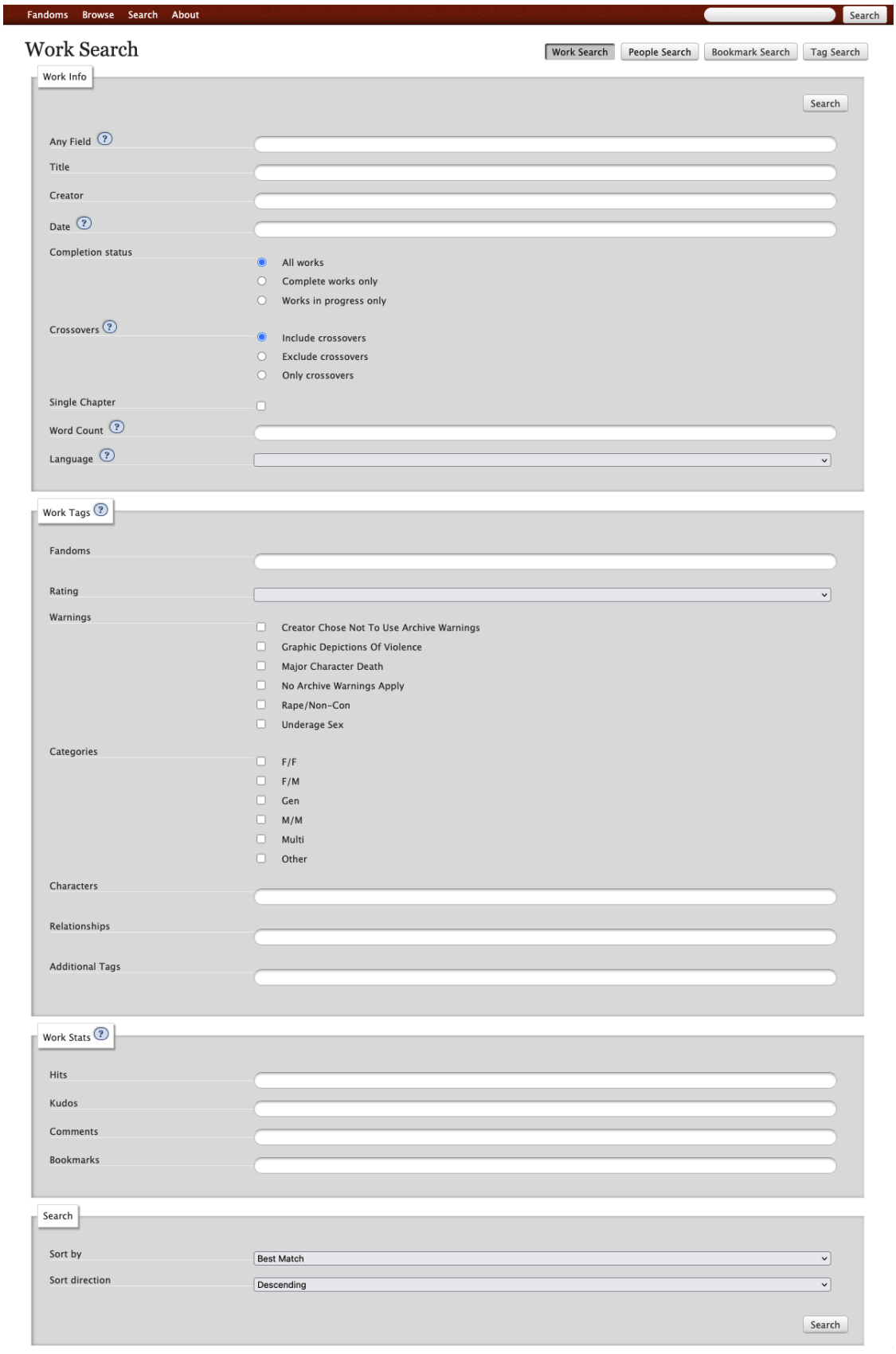


Figure A2

The Work Search Page on AO3



Note. The fields 'Fandoms,' 'Characters,' 'Relationships,' and 'Additional tags' provide auto-complete suggestions of canonical tags (see section 2.2) as users type, but users can also type in tags freely, separated by commas.

Figure A3

Browse Interface on AO3, Showing the Filter Sidebar

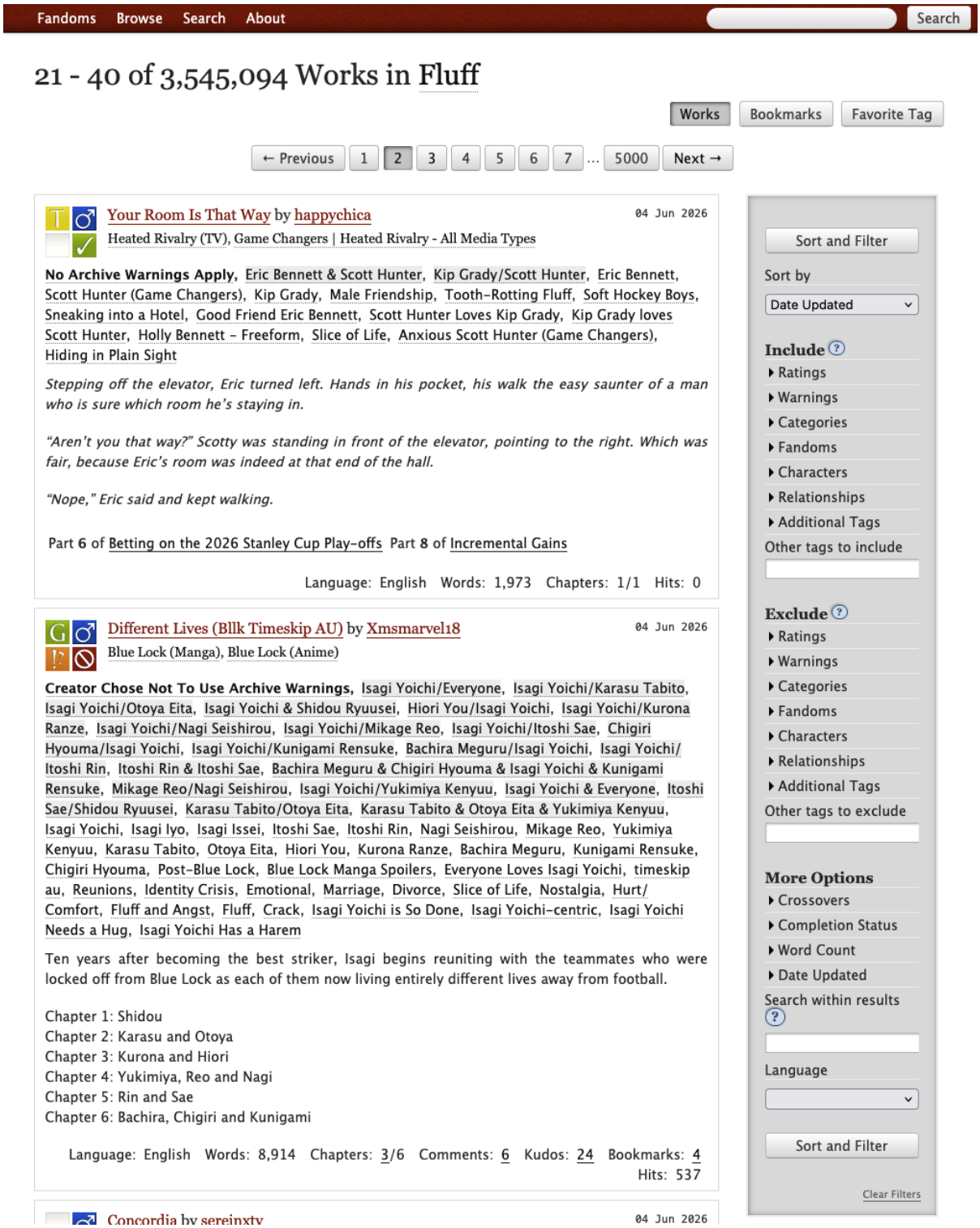
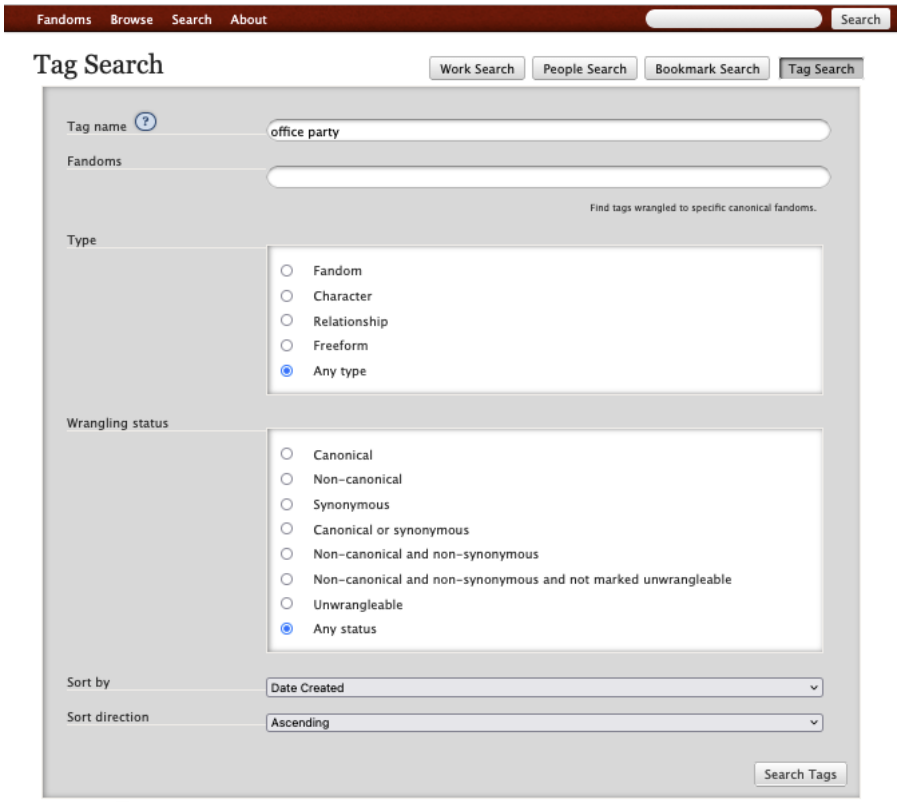


Figure A4

Tag Search Result Page on AO3

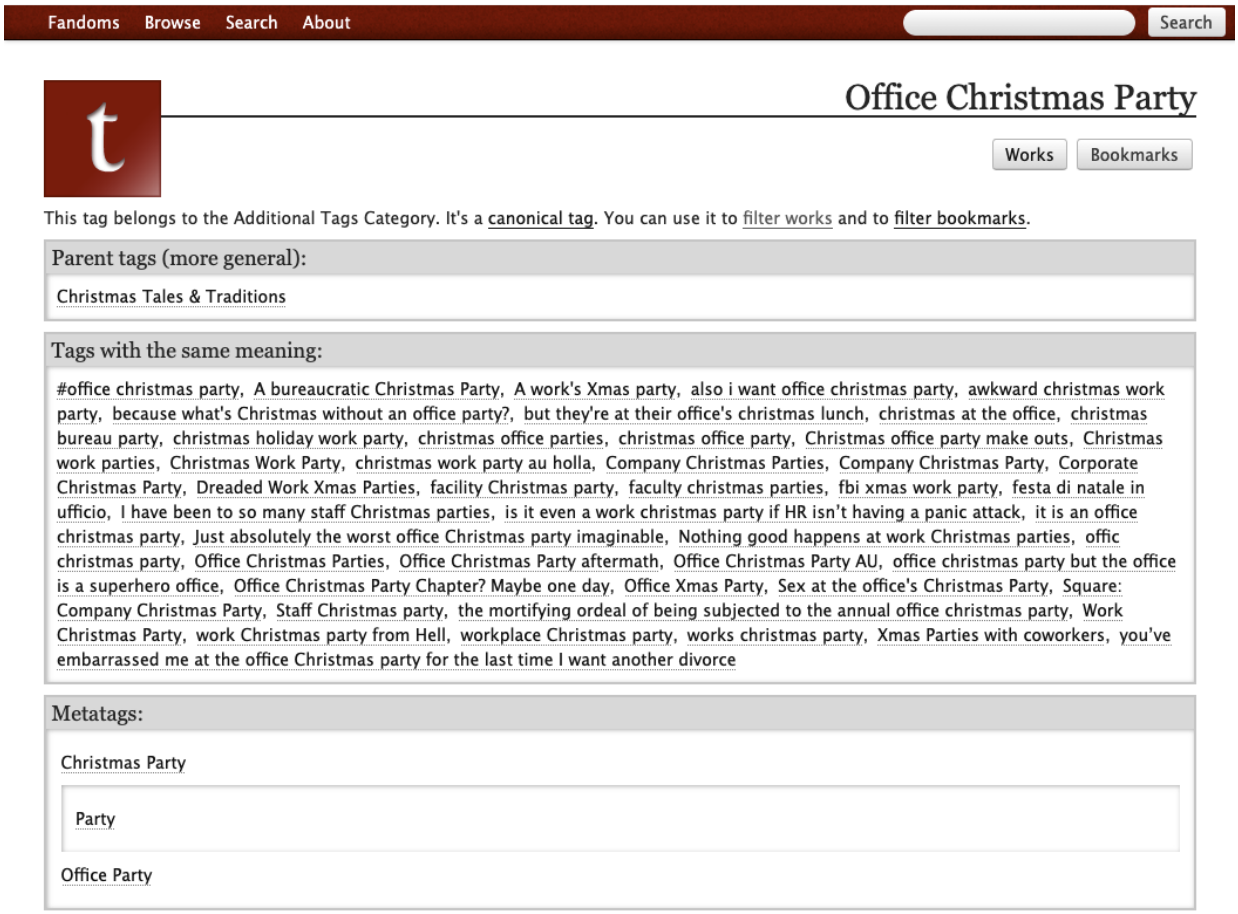


122 Found

Freeform: Underworld Office Party (1)
Freeform: Office Party (1614)
Freeform: christmas office party MI6-style (1)
Freeform: Office Christmas Party (232)
Freeform: End of year office party (1)
Freeform: being interrupted during office party (1)
Freeform: office party au (1)
Freeform: also combined with a holiday office party au my friend found on tumblr (1)
Freeform: #office christmas party (1)
Freeform: Office party with no alcohol (1)
Freeform: Office Christmas Party Movie (1)
Fandom: Office Christmas Party (2016) (2)
Freeform: christmas office party (11)
Freeform: DEO Office Party (1)
Freeform: Office Christmas Party AU (1)
Freeform: there's always that one creepy Twi'lek at the office holiday party (1)
Freeform: Office Holiday Party (16)
Freeform: office party. (1)
Freeform: office christmas party but the office is a superhero office (1)
Freeform: they're at an office christmas party Stan hates (1)
Freeform: Well were at a party before sneaking into an office (1)
Freeform: also imagine the merry & happy mv but as an office holiday party (1)
Freeform: there's an office party (1)
Freeform: Party in Danzo's office (1)
Freeform: Institute Office Party (1)
UnsortedTag: Reylo office party (1)
UnsortedTag: Reylo pining office Christmas party (1)
Fandom: Office Christmas Party (movie) (1)
UnsortedTag: The Wager at the Millennium Office Party (1)
Character: Meghan (Office Christmas Party) (1)
Character: Jeremy (Office Christmas Party) (1)
Character: Clay Vanstone (Office Christmas Party) (1)
Freeform: halloween office party (4)
Freeform: Office Party! AU (1)
Freeform: awkward office holiday party (1)
Freeform: the mortifying ordeal of being subjected to the annual office christmas party (1)
Freeform: Just absolutely the worst office Christmas party imaginable (1)

Figure A5

Tag Page for a Canonical Tag on AO3



Appendix B
Survey

Page 1

Welcome to this survey about searching on AO3!

This survey is being conducted as part of a master's thesis project in library and information science at the University of Borås in Sweden. The aim of the project is to look into people's experiences when searching for fanworks on AO3 (Archive of Our Own) and what things might influence those experiences. If you've ever used AO3 to find fanworks, you're welcome to take part.

The survey contains about 20 questions split into 3 sections and should take about 10-15 minutes to complete. It will be available until the 19th of March or until enough people have responded.

You will **not** be asked for any personal or identifying information during this survey. I would recommend that you avoid sharing anything too personal in the free text fields, as well.

My thesis supervisor is Wout Dillen. You can read about him on his university staff page.[1]

If you have any questions about the survey or the project in general, you're welcome to contact me – Ricky Nilsfred – at the e-mail address below.

Many thanks for your time!
Rickard Nilsfred
S2412240@student.hb.se

I am not affiliated with AO3 or OTW (the Organization for Transformative Works, which owns AO3). While AO3 is the subject of this study, it is not involved in it in any way.

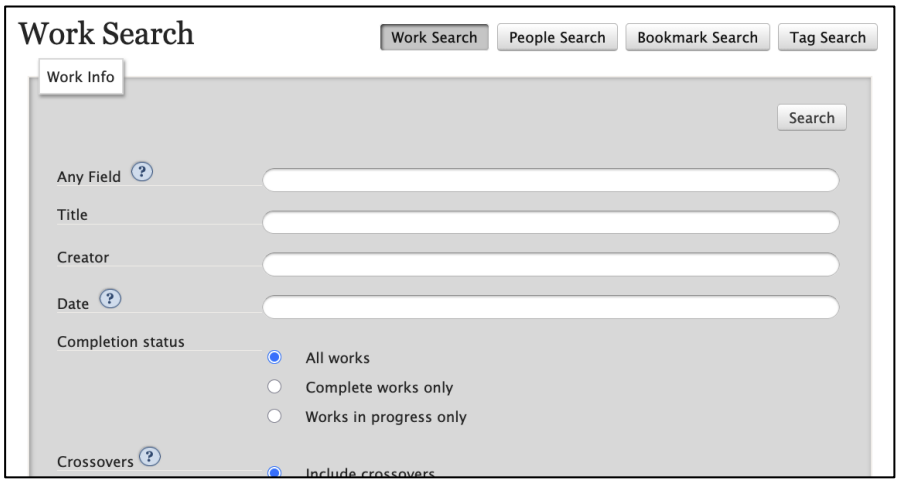
Page 2

Searching AO3

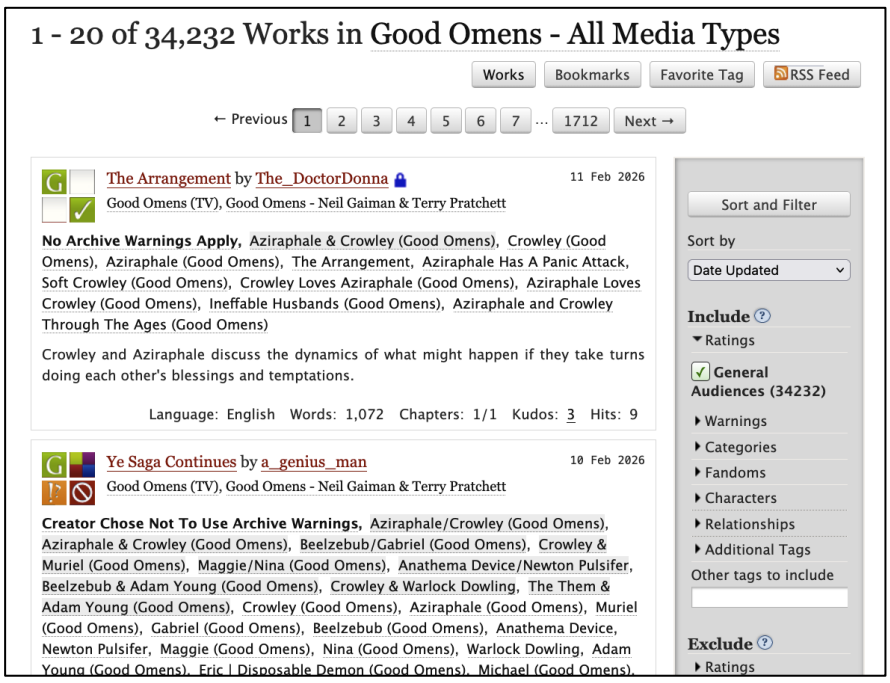
There are **three main ways to search for fanworks on AO3**. One is by using the search bar at the top of every page:

[1] The text "on his university staff page" was a link to the staff web page in question.

Another method is to use the **Work Search page** (archiveofourown.org/works/search), which can be reached by choosing "Works" in the Search menu.



AO3 also has a browse feature, which lets you see all fanworks that use a particular tag or are in a certain fandom. You can reach this by choosing a fandom using the Fandom menu or by clicking on a tag, either on a work or on the Tags page (archiveofourown.org/tags).



1. **Have you used any of these methods to search for fanworks on AO3?**
   a. I have used the search bar (yes/no)
   b. I have used the Work Search page (yes/no)
   c. I have browsed by tag and/or fandom (yes/no)
   d. I have used some other method (yes/no)

2. *(only visible if "yes" to question 1d)*
   What other method have you used? (free text field)

3. *(only visible if "yes" to at least one option in question 1)*
Which of these methods do you prefer? (single-choice)

- The search bar
- The Work Search page
- Browsing by tag and/or fandom
- The other method you described above
- I don't have a preference

4. *(only visible if "yes" to at least one method in question 1)*
Think of a typical time when you've searched or browsed for fanworks on AO3 using any of the methods above. (For simplicity, we'll call them all "the search interface".)
How much do you agree with the following statements?

There are no right or wrong answers! This is all about *your experience* of using AO3.

(strongly disagree/disagree/neither/agree/strongly agree)

4.1. I enjoy using the search interface.
4.2. The search interface is unnecessarily complex.
4.3. The search interface is easy to use.
4.4. I needed (or would need) to read the FAQ or ask a friend before I could understand the search interface.
4.5. The various functions in the search interface suit my needs well.
4.6. It's unclear what some of the functions in the search interface are for.
4.7. I think that most people would learn to use the search interface quickly.
4.8. The search interface is very cumbersome to use.
4.9. I feel confident using the search interface.
4.10. I'm not sure the search interface was made with people like me in mind.
4.11. The search results are usually precise and relevant.
4.12. I think there is too much inconsistency in the search interface.
4.13. It's hard to get the search results I need.
4.14. Overall, I am satisfied with the search interface.

5. Have you used any platforms or web sites other than AO3 to search for fanworks?
This could be, for example, Fanfiction.net, LiveJournal, Wattpad, or web sites/archives for specific fandoms. (single-choice)

- No, only AO3
- Yes, I've used 1-2 others
- Yes, I've used more than 2 others

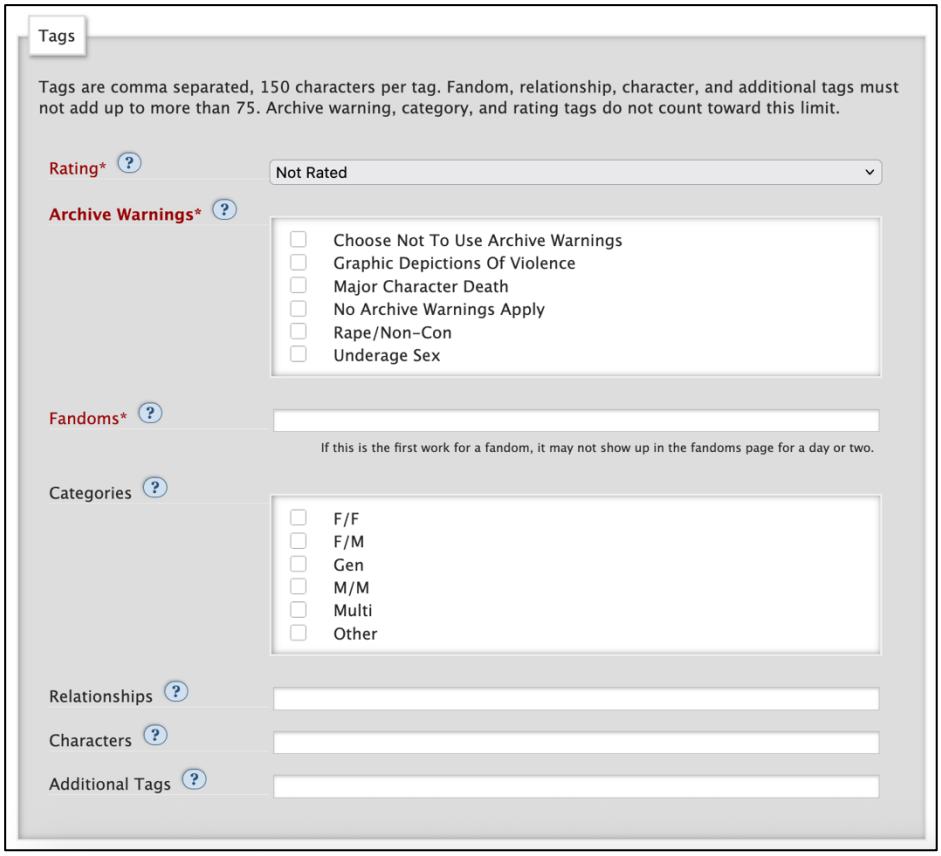
6. **Do you have any further thoughts about searching for fanworks on AO3?**
   This could be comments about the site, how you use it, or how it compares to other ways to search for fanworks. It could also be comments about the questions you've just answered. You can write as much or as little as you like. (free text field)

Page 3

Posting on AO3

7. **Have you posted any works to AO3 yourself?** (yes/no)

(only visible if "yes" to question 7)
When posting works to AO3, **the web site asks you to choose tags** of seven different types: rating; archive warnings; fandoms; category (F/F, M/M, and so on); relationships; characters; and additional tags.



8. *(only visible if "yes" to question 7)*
   Think about when you choose (or have chosen) tags for your works. How do you feel about the following statements?

   There are no right or wrong answers. Once again, **this is all about your experience.**
   (strongly disagree/disagree/not sure/agree/strongly agree)

8.1. I feel like I understand the purposes of the different tag types (fandoms, rating, relationships, and so on).
8.2. I am not sure how important tags are to get my works seen.
8.3. I know where to go if I need more information about tags and tagging.
8.4. Deciding what tags to use is difficult.
8.5. I feel confident at choosing tags for my works.
8.6. There is a high threshold for new users to learn how to tag.
8.7. Overall, I like the tagging system.
8.8. I still think I need more practice or instruction in order to tag works well.

9. How familiar are you with the following?
Please answer as honestly as you can. The purpose of this question is to find out what different users know and don't know – not to test your ability.

(I don't know this at all/I've heard of it/
I think I undestand this reasonably well/I am very familiar)

a. The hierarchy of tags on AO3 (such as the Fluff tag being the "parent tag" of Fluff and Humor, or Good Omens (TV) being a "subtag" of Good Omens (All Media Types))
b. The meaning of "&" and "/" in relationship tags
c. What canonical and non-canonical tags are
d. The popular genres and tropes of your fandom (like Crack, Angst, Enemies to Lovers, and so on)

10. Is there anything else you'd like to say about tagging on AO3?
This could be about your opinions on the tagging system as a whole, how you think when you choose tags for works, or comments on the questions above, for example. You can write as much or as little as you like. (free text field)

Page 4

About you

11. When did you first start using AO3?
Please include any interactions with the web site before you set up an account as well as after. If you're unsure, pick your best guess.

(single-choice)

- Less than 1 year ago
- 1-5 years ago
- More than 5 years ago

12. How often would you say you use AO3? (single-choice)

- Several times per week

- About every week
- Once or twice a month
- Less often than every month

13. Have you ever done any of the following?

a. Studied librarianship or information science (whether at university/college or as a vocational course) (yes/no)
b. Worked in an information profession (librarian, archivist, or another profession directly related to library & information science) (yes/no)
c. Volunteered as a tag wrangler or developer for AO3
d. Studied IT or computer science (whether at university/college or as a vocational course) (yes/no)
e. Worked in an IT or computer science profession (such as developer, software engineer, or systems architect) (yes/no)

Follow-up study

You've reached the end of this survey!

When the responses to the survey have been analysed, I hope to conduct a follow-up study with a small number of the respondents. This follow-up will take the form of a one-on-one video call with me at a time that we agree on, most likely during April. You would be asked to perform a search on AO3 and discuss the process.

Participation is completely voluntary and you can say 'no' at any point, even if you opt in now.

If you're interested in participating in the follow-up study, please let me know via e-mail at S2412240@student.hb.se! If you need a suggestion for what to write, please click the **'View'** button below for an example message that you can copy and paste. You're free to write your message differently, of course.

(shown upon clicking 'View' button)

Important information about online handles and anonymity:

For legal reasons, participants in the follow-up study will be asked to sign a personal data consent form using their government name, not their fandom or online name. If you want to keep your fandom and offline identities separate, please don't write your fandom name in any e-mails to me and consider using a non-fandom e-mail address. If you do share your fandom name with me, it will be treated like personal data, which means (among other things) that I will

handle it according to the GDPR-compliant data policy of the University of Borås[2] and delete any record of it by the end of this thesis project.

Example e-mail

Subject: Follow-up AO3 study

Message: Hi Rickard!
I am interested in participating in the follow-up study. You can contact me about it at this e-mail address.

Regards,
[name]

(shown after submitting the survey)

Thank you for completing the survey!

Your participation really helps – so thank you. Whatever contribution this study ends up making to the research about searching, tagging, classification systems, and online fandom, you've helped make that happen.

Remember that if you have any questions at any point, about the survey you just took or the study in general, you are welcome to contact me at S2412240@student.hb.se. I would love to share my results with you later!

All the best,
Rickard Nilsfred

[2] The words "data policy of the University of Borås" linked to the university's English-language page about its privacy policy, https://www.hb.se/en/shortcuts/about-the-site/privacy-policy/.

Appendix C
Survey invitation message

Hi everyone!

My name is Ricky (he/him) and I'm a student of library and information science at the University of Borås in Sweden and a Good Omens fan. As a part of my master's thesis, I'm conducting a survey that investigates **fans' experiences when searching for fanworks on AO3**, and right now, I'm looking for participants. Would you like to take part? ☺

The survey is meant for anyone who has used AO3 to find fanworks, regardless of how much experience you have using AO3. It will ask a few questions about how you search for fanworks and a couple of questions about your background. It will never ask for any personal information, however, and is set up to be completely anonymous.

You can fill out the survey at
https://sunet.artologik.net/hb/ao3-search-survey
It should take about 10-15 minutes to complete and will be available until March 19th, or until enough people have responded.

If you have any questions about this survey or my thesis project in general, you're welcome to contact me at S2412240@student.hb.se, or to comment below. (E-mail is the most reliable option, though, as Reddit notifications can sometimes go missing.)

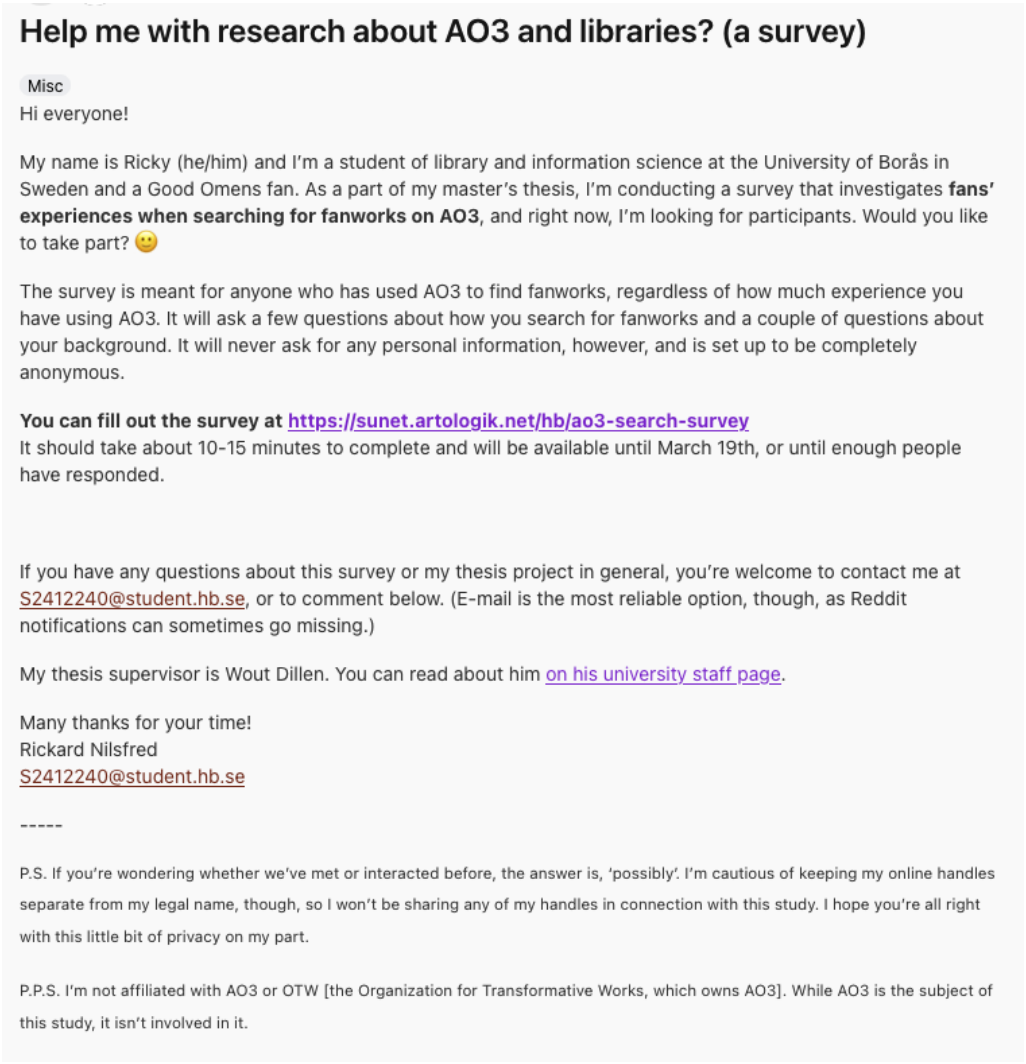
My thesis supervisor is Wout Dillen. You can read about him on his university staff page[1].

Many thanks for your time!
Rickard Nilsfred
S2412240@student.hb.se

P.S. If you're wondering whether we've met or interacted before, the answer is, 'possibly'. I'm cautious of keeping my online handles separate from my legal name, though, so I won't be sharing any of my handles in connection with this study. I hope you're all right with this little bit of privacy on my part.

P.P.S. I'm not affiliated with AO3 or OTW (the Organization for Transformative Works, which owns AO3). While AO3 is the subject of this study, it isn't involved in it.

[1] The words "on his university staff page" were a link to the staff web page in question.



Help me with research about AO3 and libraries? (a survey)

Misc

Hi everyone!

My name is Ricky (he/him) and I'm a student of library and information science at the University of Borås in Sweden and a Good Omens fan. As a part of my master's thesis, I'm conducting a survey that investigates **fans' experiences when searching for fanworks on AO3**, and right now, I'm looking for participants. Would you like to take part? 🙂

The survey is meant for anyone who has used AO3 to find fanworks, regardless of how much experience you have using AO3. It will ask a few questions about how you search for fanworks and a couple of questions about your background. It will never ask for any personal information, however, and is set up to be completely anonymous.

You can fill out the survey at https://sunet.artologik.net/hb/ao3-search-survey
It should take about 10-15 minutes to complete and will be available until March 19th, or until enough people have responded.

If you have any questions about this survey or my thesis project in general, you're welcome to contact me at S2412240@student.hb.se, or to comment below. (E-mail is the most reliable option, though, as Reddit notifications can sometimes go missing.)

My thesis supervisor is Wout Dillen. You can read about him on his university staff page.

Many thanks for your time!
Rickard Nilsfred
S2412240@student.hb.se

P.S. If you're wondering whether we've met or interacted before, the answer is, 'possibly'. I'm cautious of keeping my online handles separate from my legal name, though, so I won't be sharing any of my handles in connection with this study. I hope you're all right with this little bit of privacy on my part.

P.P.S. I'm not affiliated with AO3 or OTW [the Organization for Transformative Works, which owns AO3]. While AO3 is the subject of this study, it isn't involved in it.

One of the two invitation posts posted on Reddit.

Appendix D
CCI session plan

1. Introductions, planned duration, agenda for the session.
2. Repeat key information from GDPR consent form (recordings are personal data; who can access the data; how long data will be retained). Clarify that no sensitive personal information can be handled.
3. Purpose of the session and the study: to look into how people search on AO3, *not* to test participant's ability.
4. Explanation of master–apprentice model.
5. **Start recording.** State name of interviewer and participant and confirm consent to record.
6. General/warm-up questions[1]
    a. Have you ever worked as a librarian or archivist or something like that or studied something related to it? (If yes, clarify what.)
    b. Have you ever worked in IT or computing or studied a related subject? (If yes, clarify what.)
    c. Roughly how often would you say you use AO3?
    d. About how long have you used AO3 for?
    e. Normally when you use AO3, what sort of things do you use it for? (Probe for: How often do they *actively search* for things? Look through bookmarks? Just read and/or comment? Post their own works?)
    f. When you're looking for something to read on AO3, how do you usually prefer to go about finding that? (Search/browse directly? Look through recommendation lists? Ask a friend? And so on.)
7. Main observation period (participant to **start screen sharing**).

    If needed, remind participant about master–apprentice model and to keep a running commentary. Prompt if needed.
8. Any follow-up questions noted down during observation.
9. Words game/card sorting exercise. Participant to choose as many words as they like about the search experience and categorise them as they find appropriate. Ask participant to explain/discuss choices afterwards.
10. Request participant's feedback about the CCI experience.
11. Thank the participant! Offer to keep them updated. Wrap up.

[1] No question about tag wrangling experience was included, as no survey respondents said yes to this question.

Appendix E
Reaction cards

Although this list is alphabetical, for the exercise, card order was randomised.

Accurate
Advanced
Annoying
Appealing
Approachable
Boring
Broad
Busy
Calm
Clear
Comfortable
Complex
Confident
Confusing
Consistent
Controllable
Creative
Dated
Difficult
Distracting
Easy to use
Effective
Efficient
Enjoyable
Exciting
Familiar
Fast
Fiddly
Flexible
Friendly
Frustrating
Hard to use
Helpful
Impersonal
Impressive
Inconsistent
Incorrect
Ineffective
Informative
Intimidating
Inviting
Motivating
Narrow
Necessary
Ordinary
Organised
Overwhelming
Patronising
Powerful
Precise
Predictable
Relevant
Rigid
Satisfying
Slow
Specific
Stable
Straight-forward
Stressful
Time-consuming
Tiring
Too little
Too much
Unapproachable
Unattractive
Unclear
Unconventional
Understandable
Unpredictable
Unsure
Usable
Useful
Valuable
Wrong

Appendix F
Breakdowns

Description	Type
Too many results to look through manually	bad search results
Too few or no results	bad search results
Tag fields use ‘AND’ logic, so can't enter multiple options	functionality
Tagging is inconsistent; impossible to know what metadata was used	functionality
Keyword search in tags only handles *some* synonymy	functionality
Keyword search in summary handles no synonymy	functionality
Belief that Fandoms field uses ‘OR’ logic	misunderstanding
Belief search terms in Any Field are comma separated	misunderstanding
Belief spaces in search bar are equivalent to ‘AND’	misunderstanding
Belief Exclude filters exist on Work Search page	misunderstanding
Belief Exclude filters exist on result pages from Work Search	misunderstanding
Unclear how to ‘undo’ selection in Rating field	discoverability
Unclear how “All Media Types” fandom tags work	discoverability
Unable to find browse-by-tag page	discoverability
Googles keyword to check spelling	workaround
Must clear search fields manually	workaround; functionality
Looks for Exclude filters on other types of search pages	workaround; misunderstanding
Compares interface to the same page on a different device (in order to find Exclude filters)	workaround; misunderstanding
Misspells name of fandom	other
No suitable suggestions in auto-complete	other
‘Asker’/‘informant’ doesn't know answer to follow-up question	other
Prior information is too vague to formulate a search	other

Note. ‘Functionality’ denotes breakdowns caused by functionality which was intended by the developers, where the user also understood the functionality. ‘Discoverability’ denotes issues caused by a lack of clarity in the interface.