



City Research Online

City, University of London Institutional Repository

Citation: Price, L. (2019). Fandom, Folksonomies and Creativity: the case of the Archive of Our Own. In: The Human Position in an Artificial World: Creativity, Ethics and AI in Knowledge Organization. (pp. 11-37). Germany: Ergon Verlag. ISBN 9783956505508 doi: 10.5771/9783956505508-11

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/25151/>

Link to published version: <https://doi.org/10.5771/9783956505508-11>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Fandom, Folksonomies and Creativity: the case of the Archive of Our Own

Ludi Price, City, University of London

Abstract

Over recent years Web 2.0 has brought information into the hands of the public, and we are increasingly seeing non-professionals doing sophisticated information tasks not merely for work, research or personal interest, but also for leisure – and even pleasure. This paper looks at an online fanfiction repository, Archive of Our Own (AO3), and investigates the ways that media fans have co-opted new technologies to build a ‘curated folksonomy’ (Bullard 2014), in order to organise the fanworks (fan-created creative works) uploaded by fans to the website. Run by volunteers, the site is a fascinating example of how passion, and even obsession, can bring amateur knowledge workers together collaboratively with users to build an intricate ‘democratic indexing’ system (Hidderley and Rafferty 1997; Rafferty and Hidderley 2007).

Through methods of tag analysis and interviews, the paper explores how *Archive of Our Own’s* curated folksonomy allows fans to make full and creative use of their own original, freeform tags, while also building a highly granular and sophisticated taxonomy which, though highly labour-intensive to maintain, serves the community by maintaining a high degree of accuracy while also preserving the folksonomic properties of freeform tagging. As well as building a functioning taxonomy, through standardising its nomenclature, and facilitating the discoverability of AO3’s collections to its users, these amateur knowledge workers see their domain expertise and knowledge organisation labour as a type of fanwork that ‘gives back to the community’, in lieu of other creative works such as fanfiction and fanart.

1. Introduction

Archive of Our Own (AO3) is a fanfiction archive which was developed and is run by fans, for fans – in this case, fans can be defined as followers of media franchises, products or series, such as seen in books, movies, TV, videogames, comics, and so on. AO3 is run by the non-profit Organization for Transformative Works and is maintained wholly by volunteers and donations. The entire archive was recently nominated for a Hugo Award in the Related Works category (Cole 2019). Nomination for such a prestigious award has brought wider recognition for the site, and for its efforts to preserve the many artistic and creative works that fans produce, which are termed ‘fanworks’. AO3 mainly hosts the textual format of fanfiction, but it can also host fanart, videos, songs and lyrics, and games, among other creative works.

This paper, based on the doctoral thesis of the author (Price 2017), details the ‘curated folksonomy’ (Bullard 2014), which is the system used on AO3 for knowledge organisation. This is done using tag analysis, and supplementary interviews with ‘tag wranglers’ – i.e. volunteer subject experts, who link user-generated tags to synonymous, standardised forms. The purpose of the paper is to highlight several points: a very successful use of Hidderley

and Rafferty's (1997) 'democratic indexing'; the sophisticated practices of the volunteer 'tag wranglers' who perform these indexing and classification tasks; and the ways in which such work can be driven by passion and pleasure. It also presents a fan tag taxonomy, based upon the tag analysis section of the study.

The study detailed in this paper describes a portion of a wider comparative case study of three different online platforms used by fans. AO3 was studied in conjunction with Tumblr and Etsy. To read the entirety of the comparative case study, see Price (2017, chapter 5).

1.1 Tagging on AO3





AO3 allows users to organise and categorise their work using tags. However, in a system dubbed a 'curated folksonomy' by Bullard (2014), volunteers called 'tag wranglers' filter these tags by associating them with established synonyms. This flexible system allows for both individual idiosyncrasies in user tagging behaviour to remain intact, while also enabling efficient search retrieval. It does however require considerable effort on the part of the volunteers to combine user tags with established synonyms.

AO3 implements a combined self-tagging and automanual system. Pre-defined tags are suggested when filling them in, although users are also free to choose whatever terms they wish. Tags do not take a hashtag format, and there are no restrictions on spaces, length or characters. Non-Roman script is also allowed. During this study, an interviewed tag wrangler, Participant D, described their work thus:

When a user creates a new, never-before-used tag, it shows up in what we call the "unwrangled bins" of every wrangler assigned to the fandoms tagged on the work. What wranglers such as myself do is look at those incoming tags, and determine, based on the Wrangling Guidelines, if the tag should be marked as canonical (the form of that concept that will show in the drop down menus and autocompletes), made a synonym of any existing canonical, or left unfilterable as a tag that is too unique to be useful for other users to filter with. As a general rule, any character who exists in canon, and any relationship that involves at least one canonical character, will be canonized on the first usage. More general concepts (such as "Alternate Universe" or "Angst") will generally need to be used by multiple users before being canonized.

Figure 1 shows a random search result of works archived under the 'Remy LeBeau/Rogue' tag. The tags are displayed after the archive warning (in this case, the author chose not to employ any archive warnings, e.g. violence, non-consensual sex, etc.).

Figure 1. Archive entry for an X-Men fanfiction on AO3, "Confessions of a Train Wreck", 8th October 2018. Source: <http://archiveofourown.org/> (retrieved 22 April 2019).

Confessions of a Train Wreck by [redacted]

X-Men - All Media Types

08 Oct 2018

Creator Chose Not To Use Archive Warnings, Remy LeBeau/Rogue, Rogue (X-Men), Remy LeBeau, I wrote this story while drunk, Drunk Secrets, Tacos, secret sharing, Hippies, Valle Soleada, Rogue getting drunk, everyone gets drunk, Sandra Hill: romance novelist, i make jokes, this is my headcanon, My First Fanfic, Mentions of My Little Ponies, and Carebears, train wreck, party all night, booze hounds, hangovers, know your limits

There is no way Rogue's parents drove a VW Bug with flowers on it and did all those drugs and nonsense in an attempt to reach the "Far Banks" and named her something like "Anna Marie". My biggest headcanon brought to life. I say the "F-word" a bunch.

Language: English Words: 2,530 Chapters: 1/1 Comments: 2 Kudos: 10 Bookmarks: 1 Hits: 147

In this case greyed out tags denote either a romantic pairing of two characters. Character tags follow, and then descriptive tags. Normally, pre-defined tags start with a capital (e.g. “Hippies”, “Drunk Secrets”, “Tacos”). There are several tags that the author herself has applied freely to the fanfic. These express themes (“i make jokes”, “know your limits”) and story elements (“party all night”, “hangovers”). These tags are typical examples of how users choose to tag their works on AO3, i.e. using a mixture of pre-defined tags and free ones. If a free tag becomes popular enough, it will be merged by a tag wrangler with a standard, pre-defined one; or a new tag will be entered into the system’s taxonomy to accommodate it. For example, in Figure 1, the free tag, “i make jokes” has been made equivalent to the standard tag “Jokes” – if the tag is clicked on, it will lead to all works on the site that have used the tag “Jokes” or their equivalents. Tag equivalencies, once determined by a tag wrangler, are saved automatically in the site’s database.

1.2 Fandom, democratic indexing, and curated folksonomies

Since the rise of Web 2.0. during the 2000’s, increasingly dynamic information technologies have allowed for a more bottom-up or heterarchical (i.e. unranked) system of online, digital content creation. This more democratic approach is exemplified by cases such as Wikipedia, citizen journalism, open source software programming, amateur videogame development,

and basement-made music projects. In terms of fan culture, online communities form around certain franchises or media texts to engage in fantasy and the exchange of common interests, fanworks, and, of course, information (Lee et al 2013). Fans have been prolific users of information technology to create, share and disseminate both information and fanworks (Jenkins 2006). This also extends to knowledge organisation. As Hart et al. (1999) suggest in an early study, fans engage in sophisticated bibliographic control of their creations. Recent work has begun to explore this in more detail. In particular, Julia Bullard (2014, 47), who has been conducting a long-term ethnographical study of an online fanfiction repository, presents the notion of the ‘curated folksonomy’, which she describes as “a system of tag synonyms and tag relationships that addresses some of the major shortcomings of a pure, unregulated folksonomy”. The curated folksonomy involves a degree of structure that mitigates some of the problems associated with pure folksonomies (e.g. the oft-quoted example of tagging photos of oneself with ‘me’). This is particularly of note in terms of fandom, because many fanwork repositories, such as Livejournal and Wattpad, use folksonomies as a way of organising documents, and this can be seen most clearly in AO3.

The curated folksonomy is not a new concept and is very similar to the idea of ‘democratic indexing’ (Rafferty and Hilderley 2007; Hilderley and Rafferty 1997). Here, subject experts evaluate and formalise the indexing choices of a systems’ users to create a taxonomy/ontology. Democratic indexing, as Rafferty (2010, 260) explains, “examines the terms or tags attached to each field and creates a collective interpretation for each field based on counting terms”. As she also notes, such processes have now been made much more achievable with Web 2.0 technologies, which “could potentially allow for the development of interesting approaches to the retrieval of cultural documentation including fiction” (260). This is, indeed, exactly what takes place on AO3. In their previous study of indexing methods on Flickr, Rafferty and Hilderley (2007, 408) note:

The discourse of user-based indexing is one of democracy, organic growth, and of user emancipation, but there are hints throughout the literature of the need for *post hoc* disciplining of some sort. This suggests that, despite Shirky’s claim of philosophical paradigm shifting for social tagging, there is a residing doubt amongst information professionals that self-organising systems can work without there being some element of control and some form of “representative authority”. Perhaps all that social tagging heralds is a shift towards user warrant.

I would contend that this is exactly what has happened on Archive of Our Own, where the tag wrangling system enables domain experts to discipline user tags while, as Rafferty (2010, 260) describes it, “still allowing for user interpretation and the recording of historical shifts in our understanding of generic [and, in this case, fan] history”. What is perhaps different here, is that the “*post hoc* disciplining” is not performed by traditional domain experts, but

by amateur domain experts, i.e. fans who are conversant enough in a certain fandom to be judged competent enough to become that “representative authority”. They are not professionals, and hold no qualifications other than being self-taught, and passionate about the domain they are interested in.

To return to the concept of warrant, as mentioned in Rafferty and Hilderley’s (2007) quote above, Bullard’s research into knowledge organisation in fanfiction repositories has also yielded some fascinating insights into how a curated folksonomy works in practice, and how this collaborative process involves the application of different warrants to create an effective daily classification system that is in constant use. Classification design is always, to some extent, reliant on the concept of warrant. As Bullard (2017, 76) explains, “classification designers express their allegiance with particular theories of classification through their appeals to warrant – the body of evidence and terminology taken as authoritative in the design of a classification system”; or, by Beghtol’s (1986, 110) definition, warrant is “the authority a classificationist invokes first to justify and subsequently to verify decisions” in their choice of terms. Different warrants include: 1) literary warrant (classification derived from the field of scholarship that is being classified); 2) scientific or consensus warrant (classification based on current scientific conclusions and consensus between relevant fields); 3) user warrant (classification based on user needs and/or expectations), and; 4) ethical warrant (classification based on ethical considerations regarding users, e.g. minority groups, discriminatory language in current classification systems, potential divergence from consensus terms).

In practice, classification design is more complex, and several types of warrant may be used at any given time. Bullard’s (2017) work expands on this by giving examples of how warrant works in an online fanfiction repository where its folksonomy is highly specialised and constantly expanding as users add to it. The volunteers who curate this folksonomy communicate behind the scene to discuss controversial or problematic terms that have been entered by users. The scenarios Bullard describes succinctly indicates the tension classification designers encounter regarding different warrants, and that these tensions are not easily surmountable, especially when a classification system is being collaboratively designed, and even more so when it is being developed on-the-fly. Despite the fluid nature of collaborative and democratic classification design, it can work successfully. Unlike official classification systems, created by professional bodies, the curated folksonomy described in Bullard’s work, and seen on sites such as AO3, is not monolithic and does not take years to implement change. Its workers are passionate, expert volunteers. Interviews with AO3’s tag wranglers, discussed in section 6.0., show that while the curated folksonomy

system is under strain, it is nevertheless a successful one that generally works well, simultaneously both preserving and standardising the terms created by its users (i.e. the fan community). When one considers the vast size and granularity of the folksonomy, AO3 is a stunning achievement, blending all four warrants to build a classification system that both serves its community well and generally describes content accurately.

Here it is also important to note that Bullard's research into her fanfiction repository posits classification work as fun, pleasurable, and recognised by the wider fan community as the efforts of volunteers with a particular expertise. There is still a relative dearth of research acknowledging the important role that fun, passion, obsession and play have in motivating volunteers to take part in collaborative knowledge organisation projects. Activities such as classifying galaxies by shape (GalaxyZoo), editing a wiki article (Wikipedia), or standardising an obscure fandom term (AO3) can be monotonous in the extreme – so why are so many people doing it? Scholarship should perhaps move away from regarding these activities from the sole standpoint of something which constitutes labour, and instead consider creativity, passion and play as a way to understand why people such as amateur experts, enthusiasts and fans engage in this type of activity.

2. Empirical Study

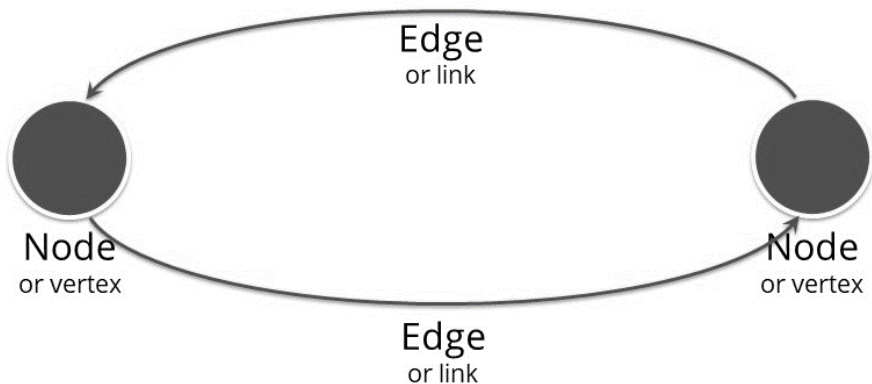
The aim of the empirical section of the study was to understand how the curated folksonomy used on AO3 was being used – how fans chose to tag their work; how tag wranglers controlled the taxonomy, and whether they were effectively preserving the meaning of the original tags when standardising them 'behind the scenes'. This involved: a) ascertaining the meaning of tags, as they had been input by users; b) ascertaining the user's intention in assigning that tag (e.g. was it being used to describe the content of the work, or something else); and c) ascertaining whether the meaning of the original tags was being preserved by tag wranglers. A secondary aim was to engage with the tag wranglers about their work and their opinions on AO3's tagging system. Two methods were used to achieve this: firstly, tag analysis; and secondly, structured interviews.

2.1 Methods

Tag analysis as a research tool has its roots in hyperlink network analysis, or link analysis (Thelwall 2004), which in turn has its roots in social network analysis (Park and Thelwall 2003). Social network analysis (SNA) is a research approach and technique that has been widely used in the social sciences for many decades (Carrington and Scott 2011, 1). It takes as its premise the idea that, as Marin and Wellman (2011, 11) explain, "social life is created primarily and most importantly by relations and the patterns formed by these relations".

When people form networks, they bring with them and exchange resources. These resources can be tangible in form: money, goods, and services; or they may be intangible, such as information, expertise, and influence (Haythornthwaite 1996, 323). People become ‘nodes’ (or vertices) in the network, network members or actors who are linked through relationships, or ‘edges’ (or links – see Figure 2). Social network analysis studies these relationships within a network for meaningful patterns that can tell us about the nature of the network, such as how connected each node in the network is, which node is the most connected, and through which relationships resource exchange works most efficiently.

Figure 2: The basics of a social network (Source: Hawksey 2017).



There are several ways in which sense can be made of a network, which are as follows:

- ▪ Degree: the number of connections that a node has. The more connections, the higher the degree.
- ▪ Betweenness centrality: how often a node appears on the shortest path between two other nodes. The higher the betweenness centrality, the higher the importance of that node in exchanging a resource within the network – thus nodes with a higher betweenness centrality can be considered a more efficient carrier of information between other nodes.
- ▪ Clustering: groups of highly interconnected nodes within a network. A cluster denotes nodes that can reach one another in only one step. This is a group of highly influential nodes.

- ▪ Density: the degree to which a node is connected to all other nodes in the network.

Social networks are not merely restricted to people. In fact, Marin and Wellman (2011, 11) note that “any units that can be connected to other units can be studied as nodes”, and networks can be used to describe the relationships between units of information as well. It is thus not surprising that it has found applications within the field of LIS (Bawden and Robinson 2012, 174; Otte and Rousseau 2002). The potential usefulness of the method in the discipline appears to have first been explicitly suggested by Haythornthwaite (1996). Haythornthwaite (1996, 338-339) noted five aspects of information exchange that SNA is well able to shed light on. These are:

- ▪ Information needs: information exchange between certain group members, and the type of information being exchanged, can tell the information provider how best to serve users.
- ▪ Information exposure: relationships with highly influential network members can illustrate a person’s level of exposure to information.
- ▪ Information legitimization: measuring the strength of ties between network individuals can show how information is being passed on to others. The stronger the tie, the more legitimised the information (and therefore its source).
- ▪ Information routes: establishing the routes of information exchange within the network is useful not only for describing information flows, but also which routes are most efficient.
- ▪ Information opportunities: influential people in the network can control information flows between other individuals within the network, thus becoming information brokers or gatekeepers, regulating both information sources and outlets.

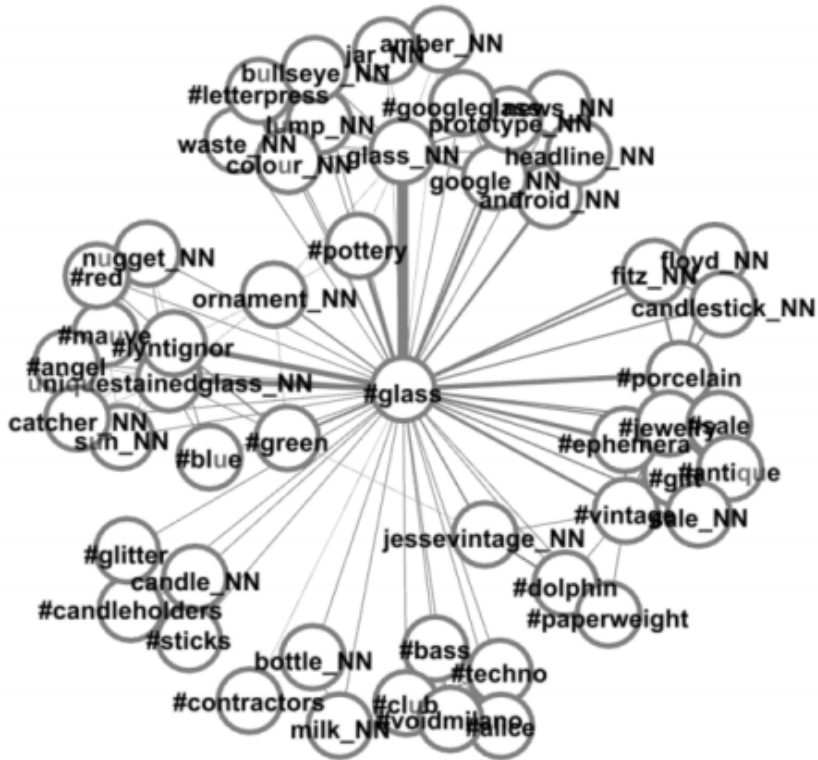
Since then, SNA has been used in a variety of papers within the field of LIS. Johnson (2019) used the method to explore collaborative information seeking between healthcare teams; Jiang, Zhang and Liu (2014) to map the relationships between the editors of LIS journals in China; Jalalimanesh and Yaghoubi (2013) examined an Iranian interlibrary loan service to map the transfer of knowledge between institutions; and Johnson (2004) used SNA to investigate how a group of residents in Ulaanbaatar, Mongolia, searched for information, finding that they often went to people they didn’t know very well, but who had higher social

capital. This is but a small sample of works in this area – at the time of writing, LISTA lists 65 papers that use or mention social network analysis as a research method.

The growth of social network analysis over the past couple of decades is not surprising, considering the rise of the internet and the fact that it is, in effect, a vast social network in and of itself (Otte and Rousseau 2002, 441). This has led to the method of hyperlink analysis (Park and Thelwall 2003), or simply link analysis (Thelwall 2004), which “casts hyperlinks between Web sites (or Web pages) as social and communicational ties, applying standard techniques from Social Network Analysis to this new data source” (Park and Thelwall 2003, n.p.). Here, the website is the node, and the hyperlink is the edge that connects websites. By analysing a network of hyperlinks, one can discern patterns between individuals, organisations, companies, and even nation states through their website links, much as one would by analysing offline social networks.

Not only can social network analysis methods can be applied to people, organisations and websites, it can also be applied to metadata stored within the Web. One of the ways in which this has taken shape over the past decade or so is in the form of *tag analysis*, where the network properties of tags are analysed. In this case, the nodes in the network are not people or organisations, but tags (or hashtags), for example on Twitter, Flickr, or Delicious. The edges between nodes in a tag network demonstrate when a tag is used in conjunction with another tag in the same post (see Figure 3).

Figure 3: A co-occurrence graph of the hashtag #glass. Source: Wang and Iwaihara (2015).



Tag analysis can be used to examine many facets of online phenomena, such as political sentiment on Twitter (Small 2011), usage patterns of bookmarking tags on Del.icio.us (Golder and Huberman 2006), and the semantic information in Flickr tags (Bolognesi 2016). Tag analysis is particularly prevalent in the context of Twitter hashtags, of which there is much literature – recent research includes Malik et al. (2018), Rossi and Giglietto (2016), Wang, Liu and Gao (2016), and Wang and Iwaihara (2015). A growing area of related research involves the merging of tag analysis and social network analysis, where the latter is applied to the analysis of tags in order to visualise and thus better understand the network-type properties of social media folksonomies (Cattuto et al. 2007; Ma and Li 2014). Such analyses are presented in graph form, usually depicting a base tag as a central node in a network, connected to co-occurring tags – these graphs are called *co-occurrence graphs* (see Figure 3). A central node (in this case #glass) represents the base tag; tags that are co-occurring (i.e. that occur in the same post, or tweet) are joined to the central node by an edge.

More complicated relationships between tags, such as group clustering etc., can be visualised by the application of various algorithms, which can depict tag usage amongst different communities, thus elucidating how patterns of tag usage differ between different groups and networks. This gives some insight into information exchange in online and social media settings.

Tag analysis has also been used to describe the tagging behaviours of users within a given information domain. Trant (2009, 23) gave an early overview of LIS and computer science articles on tagging and folksonomy, noting that tools “such as vocabulary analysis and classification, user interaction theory, and social network theory are used to describe and analyse the nature of tagging and folksonomy”. Examples of LIS papers which use tag analysis are Ådland and Lykke (2012), Chen and Ke (2013), Vaidya and Harinarayana (2016), and Estrada et al. (2017). Within fan contexts, Johnson (2014), Rose (2013), and Gursoy (2015) have examined tag usage in fanfiction repositories, although these studies did not use social media data/tag analysis specifically.

While tag analysis comprised the bulk of the empirical research presented in this paper, it was felt that a qualitative element would be needed to shed some more light on the tag wranglers’ activities. Therefore, some small-scale, structured email interviews were conducted to complement and test the results of the tag analysis.

3. Tag analysis

For this study, one particular tag which is used on AO3 – ‘Remy LeBeau/Rogue’ – was crawled. This tag describes the romantic pairing between two characters from a multimedia franchise called *X-Men*, which was created by Marvel. The tag was chosen as the author was familiar with these characters, and with the X-Men franchise at large. Since time and resources were scarce for this doctoral project, there was not the luxury of training up a team of assistants in the requisite domain expertise to parse and code the tags. Thus, it was deemed more expedient to perform the analysis on a domain that the author was already familiar with. This subject knowledge allowed the author to more easily navigate the co-occurring tags associated with the ‘Remy LeBeau/Rogue’ tag. AO3 does not use tags in the same way that they are used on social media platforms such as Twitter or Instagram. Instead, tags are based on their own ‘home page’, which has its own URL, similar to a traditional authority file (see Figure 4). Tags therefore cannot be harvested in the same way that they might be on Twitter or Instagram. They can only be retrieved via a static URL address.

Figure 4: AO3's 'Remy LeBeau/Rogue' tag homepage. User tags that have been designated synonymous by tag wranglers are shown under 'Tags with the same meaning'. Source: https://archiveofourown.org/tags/Remy%20LeBeau*s*Rogue/ (retrieved 22nd April 2019).



The screenshot shows the AO3 tag page for 'Remy LeBeau/Rogue'. At the top left is a red square with a white lowercase 't'. To the right, the tag name 'Remy LeBeau/Rogue' is displayed in a large, serif font. Below the tag name are two buttons: 'Works' and 'Bookmarks'. A line of text below the buttons reads: 'This tag belongs to the Relationship Category. It's a common tag. You can use it to [filter works](#) and to [filter bookmarks](#).' Below this is a section titled 'Parent tags (more general):' containing a list of related tags: 'All New X-Factor', 'Gambit (Comic)', 'Marvel', 'Marvel (Comics)', 'Marvel 616', 'Marvel Noir', 'Marvel Ultimates', 'Remy LeBeau', 'Rogue (X-Men)', 'Wolverine (Movies)', 'X-Men (Comicverse)', 'X-Men (Original Timeline Movies)', 'X-Men (Ultimateverse)', 'X-Men - All Media Types', 'X-Men Evolution', and 'X-Men: The Animated Series'. The next section is titled 'Tags with the same meaning:' and contains a long list of synonymous tags, including 'Anna Maria LeBeau/RemyLeBeau', 'Anna Marie/Remy LeBeau', 'anna-marie/remylebeau', 'former Remy LeBeau/Rogue', 'Gambit / Rogue', 'Gambit x Rogue', 'Gambit/Man!Rogue', 'Gambit/Rogue', 'Gambito/Rougue', 'implied Remy Lebeau/Rogue', 'implied Rogue/Gambit', 'MAYBE Remy LeBeau/Rogue', 'mentions of Rogue/Remy LeBeau', 'mr. and mrs. lebeau', 'past Rogue/Remy LeBeau', 'Post Remy LeBeau/Rogue', 'referenced Remy LeBeau/Rogue', 'Remy LeBeau | Gambit/Marie D'Ancanto | Rogue', 'Remy LeBeau/Anna Marie', 'Remy LeBeau/Anna Marie LeBeau', 'Remy LeBeau/Anna-Marie', 'Remy LeBeau/Anna-Marie LeBeau', 'Remy LeBeau/Rogue (X-Men)', 'Remy/Rogue', 'Rogue | Marie/Remy LeBeau (Gambit)', 'Rogue/Gambit', 'Rogue/Gambit (mentioned)', 'Rogue/Gambit (past)', 'rogue/remylebeau', 'Rogue/Remy LeBeau', 'Rogue/Remy LeBeau - Pairing', 'Rogue/Remy LeBeau - Relationship', 'romylebeau', 'Rogue/Gambit', 'slight Gambit/Rogue', and 'slight Remy LeBeau/Rogue'.

The crawl was done using SocSciBot 4.1. SocSciBot is a free crawler programme developed by Mike Thelwall specifically for use in the social sciences and humanities. It has seen wide use within information science, especially in webometrics, altmetrics, and link analysis research (recent examples include Thelwall 2017; Hendrikx et al 2016; Saha and Mukhopadhyay 2016). Because the tag is based on a homepage, SocSciBot was easily able to run a crawl, in a way that would have been much more complicated on, for example, Twitter or Instagram, which are hashtag-based.

The crawl of the 'Remy LeBeau/Rogue' tag took place on 29 April, 2016. The maximum number of pages to crawl was 1000, and the max crawl depth was set to 1. This was to limit the crawl only to pages which co-occurred with those using the 'Remy LeBeau/Rogue' tag. To ensure that only web addresses for tags were returned (rather than for the actual works themselves), the crawl was set to collect only pages that began with the URL

<http://archiveofourown.org/tags/>. The difficulty with this crawl is that SocSciBot does not crawl a URL that is composed of special characters. In this case, the URL that needed to be crawled includes asterisks, which SocSciBot does not recognise. Therefore, a workaround was used, wherein the URL for the synonym, ‘Romy’ (<http://archiveofourown.org/tags/Romy>) was used to start the crawl, as it automatically loads the ‘Remy LeBeau/Rogue’ tag homepage in any case.

The resulting data was saved as a Pajek file (.net). The Pajek format is a widely used standard within network science, and is interoperable with many programmes, such as SocSciBot, NodeXL, Gephi, and others. The resulting Pajek file was imported into NodeXL. Here the data was cleaned, and the URLs rendered in their plain tag form (e.g. ‘archiveofourown.org/tags/kitty*s*kurt’ became ‘kitty/kurt’). During this process several problems were encountered, as listed below:

- ▪ Due to SocSciBot’s limitations, some URLs were truncated. For example, all instances of ‘in a URL were rendered as &, and the rest of the URL was not rendered. For example, archiveofourown.org/tags/darcy%20likes%20&. In cases such as these, the original tag was often considered irretrievable. Since these examples were now rendered useless, they were removed from the dataset. Some, however, could be reconstructed by searching for the tag via Google (e.g. archiveofourown.org/tags/dracy%20and%20logan%20aren&; contained a spelling error (‘dracy’ instead of ‘darcy’), and could easily be found through a Google search (the final tag was ‘dracy and logan aren't normal’).
- ▪ Some tags can be deleted, presumably if the work it was attached to is removed by the author. In such cases, the tag was left in the dataset.
- ▪ Due to the dynamic nature of AO3, tags are always being merged with standard synonyms. The dataset therefore does not reflect changes made to tags post-analysis.

The final dataset included a total of 8182 individual tags, with a total of 4368 tag names. The next stage was to create a separate dataset by merging synonyms according to their tag wrangled version (i.e. their standardised form as determined by AO3’s tag wranglers). For example, all incidences of ‘Romy’ were merged with ‘Remy LeBeau/Rogue’, ‘ultimate x-men’ was merged with ‘x-men (ultimateverse)’, ‘logan – oc’ was merged with ‘logan (x-men)/original character’, and so on. After merging all synonyms with the standardised form, tags that were not popular enough to have been ‘tag wrangled’ yet remained. These comprised the ‘long tail’ of the dataset. The wrangled dataset came to a total of 4946

individual tags, with a total of 2752 tag names. This indicated that 63% of tags that co-occur with ‘Romy’ had been wrangled.

There were now two datasets – one comprising pre-wrangled tags, and one comprising wrangled tags (if, indeed, a tag was popular enough to have been wrangled), allowing for both sets to be compared. Each tag was then manually coded to a tag type, using an iterative, inductive process. This process was used to develop a fan tag taxonomy (see Table 1), which is described further in the author’s doctoral research (Price 2017, chapter 5). This taxonomy was based on a simple, generalised tag taxonomy used by Smith (2008, 67), but has added categories that can be implemented in fan-specific contexts.

Table 1. Fan-tag taxonomy as developed during Price’s (2017) doctoral thesis, based on Smith (2008, 67).

CODE	TAG TYPE/SUB- TYPE	DEFINITION	EXAMPLES
1	Descriptive	Describes content	vintage; commission; black and white; regram
1.1	Fandom	Describes fandom	X-Men; Marvel; Avengers; Harry Potter
1.2	Ship	Describes characters in a romantic relationship	Romy; Erik Lehnsherr/Charles Xavier; Ioroki
1.3	Character	Describes characters	Gambit; Rogue; Thor; Wade Wilson
1.4	Genre	Describes genre of resource	drabble; fluff; angst; slash; steampunk
1.5	Event	Describes a ‘real world’ event	Christmas; Valentines Day; dragoncon
1.6	Person	Describes a ‘real world’ person	Channing Tatum
1.7	Friendship	Describes characters in a friendship	kitty pryde & kurt wagner; darcy and logan
1.8	Organisation/Team/Group	Describes a group of people	witches; Hydra; X-Men; Illuminati
1.9	Location	Describes a location or setting	Alkali Lake; Xavier Institute; Wakanda
1.10	Plot	Describes a fictional story element	M-Day; Crimson Gem of Cyttorak
1.11	Warning	Describes sensitive content	spoilers; swearing; rape/non-con
2	Resource	Type of resource	comics; drawing; photo; video
2.1	Fanwork	Type of fan resource	fanfic; fanart; cosplay; fanfic rec list
2.2	Title of fanwork	Title of fan resource	In Between; Loki and the Loon

CODE	TAG TYPE/SUB- TYPE	DEFINITION	EXAMPLES
2.3	Citation	Citation of fan resource	Episode: Shadowed Past; X-Men Legacy 272
3	Ownership	Ownership of tagger	mike draws
3.1	Creator/source	Name of fan resource creator	Jim Lee; toyscomics; bbrae; ishandahalf
3.2	Recipient	Name of intended recipient of fan resource	txpeppa
4	Opinion	Opinion on resource	sexy; geeky; quirky; badass; epic
4.1	Communication	Communicates thoughts	I blame Tumblr; I need this shirt; great gift idea
4.2	Explanatory	Explains resource content	this is how I vent; iron fist is shameless
4.3	Affective	Explains emotional reaction	poor Pietro; ineedhelp; theyre so cute omg
4.4	Conversational & enunciative	Instigates or responds to a dialogue	why?; ask me stuff; leah shut up; askbox
4.5	Emoticon	Visual communication	XD; :D; 0:)
5	Self-reference	Reference to tagger/self	personal post; my art; self; my life
6	Task organising	Personal organisation of resource	work in progress; other character tags to be added; queueballs
7	Play & performance	Resource is part of an event, or has some performative aspect	prompt fill; fangirl challenge; frostiron month

4. Supplementary interviews

AO3 interviewees were recruited from tag wranglers. It was decided to interview two tag wranglers who worked specifically on works in the Marvel fandom. This was because: a) they would likely have wrangled the tags in the analysed dataset, and; b) they might shed some light on the tag wrangling process, how it impacts the site, and how they perceived their role in the organisation of fan-related information. Contact with tag wranglers must be made through the Organization for Transformative Works (OTW), who runs AO3. Therefore, a private message was sent to the Communications Team via a contact form on the AO3 site. In order to be approved for scholarly research on the site, information sheets, ethics checklists, consent forms and interview guides were emailed to the Communication Team Co-chair. Once approval was granted, interviews with the tag wranglers (Participants C and D) were mediated electronically by the co-chair. For this reason, interviews were by necessity structured. This was due to the considerable time and negotiation between

mediations, and the questions being mediated through a third party. The questions asked were:

1. What do you do as a tag wrangler?
2. What is important about tag-wrangling?
3. Tag-wrangling might be considered a monotonous task – why do you do it?
4. Do you see yourself as a gatekeeper of your fandom, and if so, how?
5. What do you think of AO3’s tagging system? Do you think it could be improved?

Interviews were conducted online via email, once consent forms had been returned. When responses highlighted concepts that required further investigation, a follow-up email was sent with further questions.

5. Analysis

There were two different datasets created from the AO3 Romy tag crawl. The first was designated the ‘pre-wrangled’ dataset – that is, all the tags were the original versions that had been input by the user. The second was designated the ‘wrangled’ dataset – that is, all tags that had been filtered by a tag-wrangler and merged with a standard, synonymous tag (e.g. all instances of “aggressive flirting” are merged with the standard tag “Flirting”; “team fic” with “Team”, etc.). Therefore, the ‘wrangled’ dataset showed higher levels of homogeneity and much lower tag name counts than the ‘pre-wrangled’ set. The ‘pre-wrangled’ dataset comprised a total of 8182 individual tags, and 4638 different tag names. The ‘wrangled’ dataset comprised a total of 4946 individual tags, and 2752 different tag names. This indicated that well over half of the tags in the data set (63%) had been processed by tag-wrangers.

Figure 5. Co-occurrence graph for the ‘Rogue/Remy LeBeau’ tag on AO3 (from the wrangled dataset).

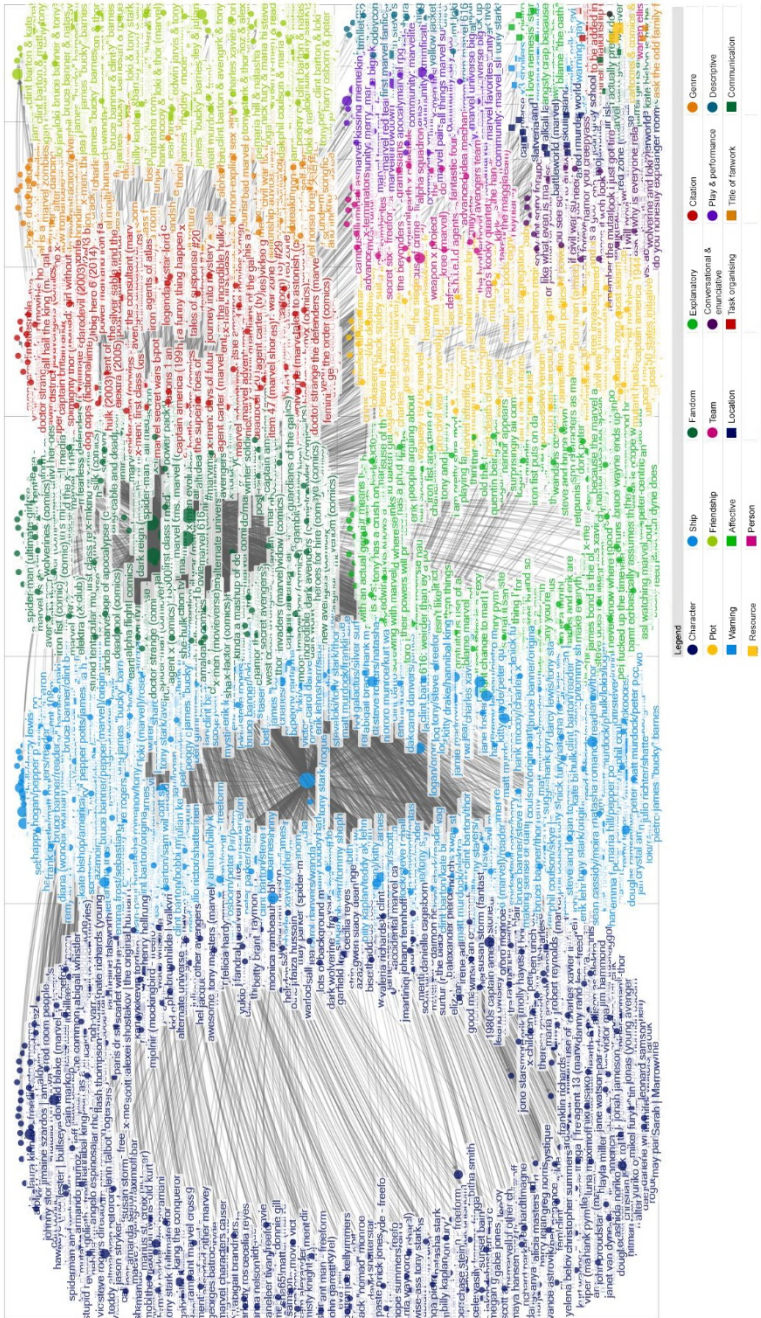
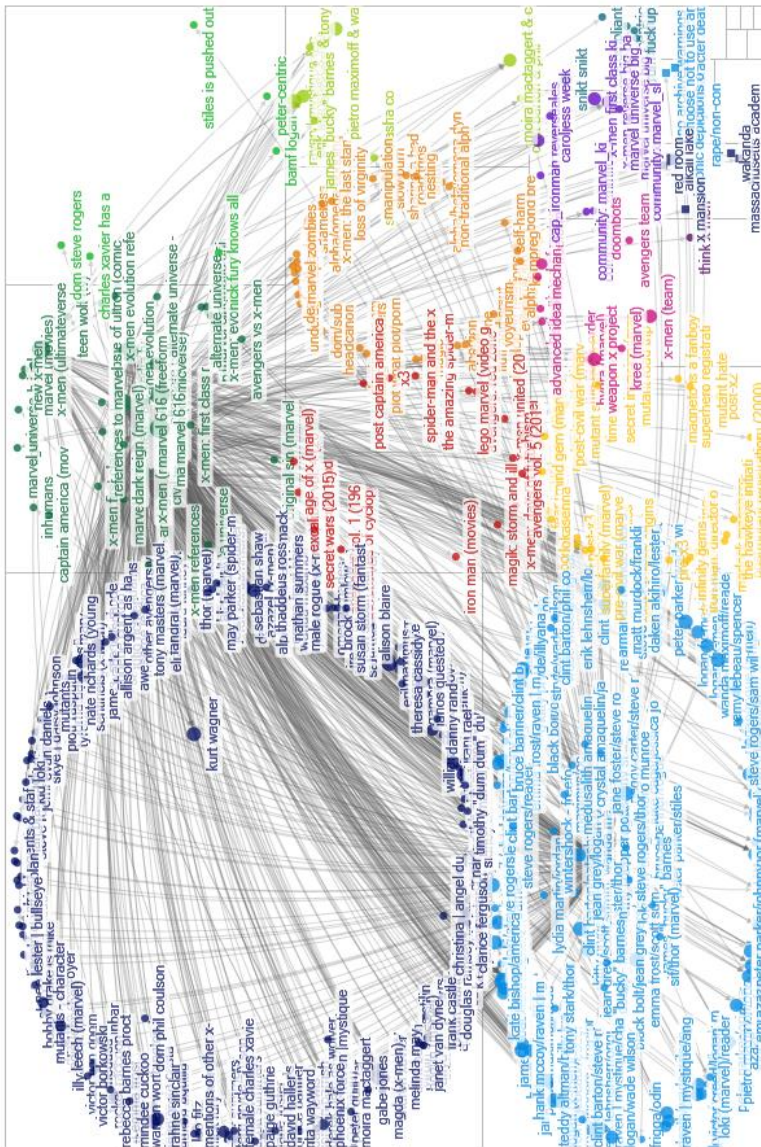


Figure 5 shows a directed co-occurrence graph for the ‘Remy LeBeau/Rogue’ tag on AO3, grouped by tag type, laid out using the Fruchterman-Reingold algorithm, as this had the most visual clarity. Edge opacity is denoted by the edge weight; vertex size by betweenness centrality; tag type by colour and vertex shape. Figure 6 depicts all tags with a betweenness centrality of 1 or above (that is, tags that appear to be more efficient carriers of information content within the network). Three hundred and nineteen (11.6%) tags reached this value. Each tag was categorised according to the fan tag taxonomy on Table 1. Most of the tags with a betweenness centrality of 1 or above either described the franchise or property associated with the post (Fandom); relationships between two characters (Ship); the characters themselves (Character); or simply described the content of the story (Descriptive). These figures were similar in both ‘pre-wrangled’ and ‘wrangled’ datasets.

This was an important finding. It had been expected that the effect of the tag wrangling process would be evident on the ‘wrangled’ dataset, thus implying that some form of gatekeeping or inaccurate/biased bibliographical control was being exerted by the tag wranglers. In fact, there was very little difference between pre-wrangled and post-wrangled tag usage. Figures 7 and 8 compare the number of tag names in both pre-wrangled and wrangled datasets – they show very similar patterns, despite the standardisation of the wrangled set.

Figure 6: Co-occurrence graph for the ‘Rogue/Remy LeBeau’ tag on AO3 (from the wrangled dataset). All displayed tags have a betweenness centrality of 1+.



Legend					
● Character	● Ship	● Fandom	● Explanatory	● Citation	● Genre
● Plot	● Friendship	● Team	● Conversational & enunciative	● Play & performance	● Descriptive
■ Warning	■ Location				

Figure 7: Comparison of pre-wrangled and post-wrangled tag names, by type (see Table 1 for tag types). There is little appreciable difference.

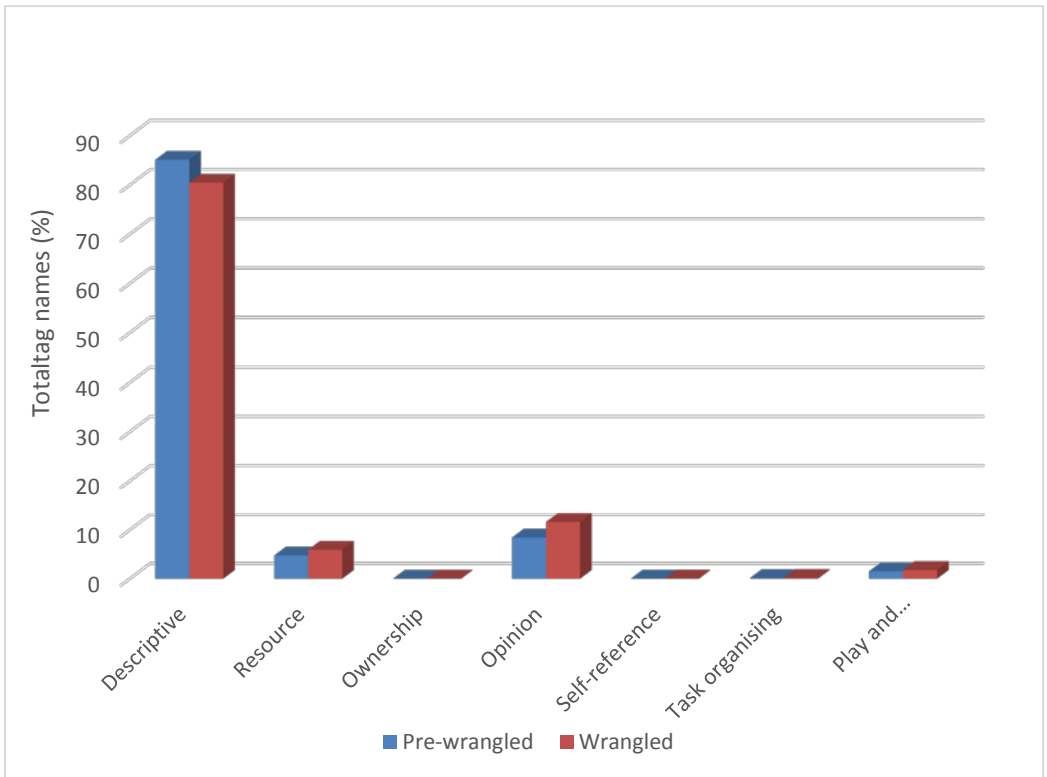
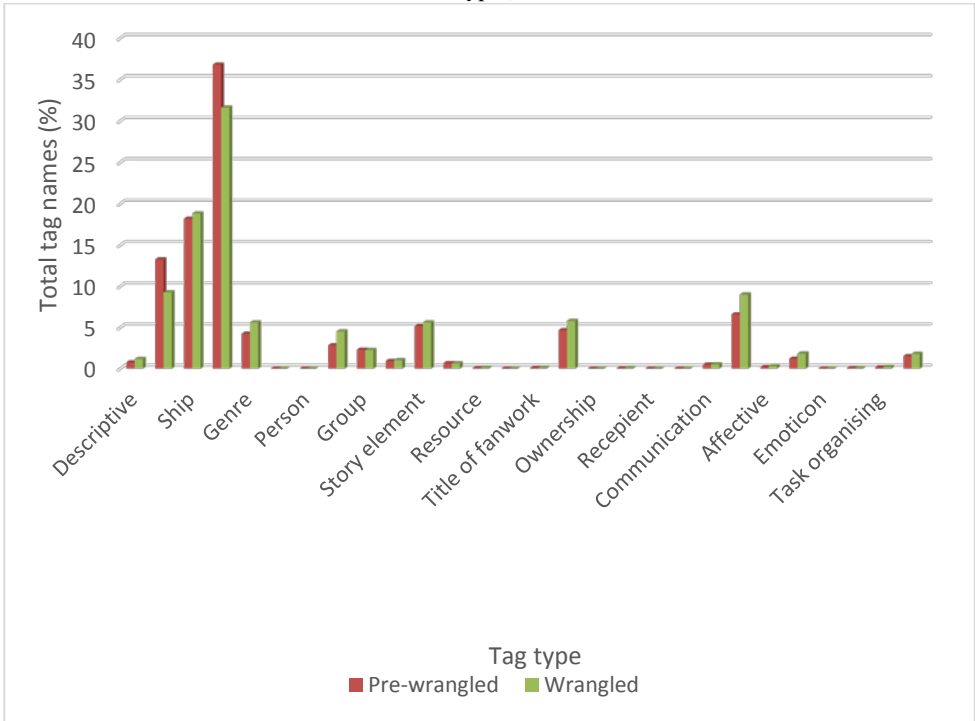


Figure 8: Comparison of pre-wrangled and post-wrangled tag names, by sub-type (see Table 1 for tag types).



This would suggest that tag wrangling is not a form of gatekeeping of the vernacular or the taxonomy used in the Romy or wider Marvel fandoms. Nor is it gatekeeping in terms of the bibliographical control of fanworks. Indeed, during the interviews, the tag wranglers did not see themselves as gatekeepers, and confirmed that they tried to follow the original tagger's meaning and intent as closely as possible. Examples from their interviews show considerable expertise in their chosen area (i.e. the Marvel Universe, and therefore it may be concluded that, in order to do their task properly, tag wranglers pride themselves on being able to recognise the obscure references in certain tags, preserve them in the wrangling process, and standardise them if warranted. As Participant D said:

What I do see myself as providing is a chance to make too many years reading a lot of comic books useful. Marvel has a very, shall we say, dense, history. But if you think there aren't users out there who will tag for characters who appeared in one issue of *Fantastic Four* back in 1973, I want to assure you: you are wrong.

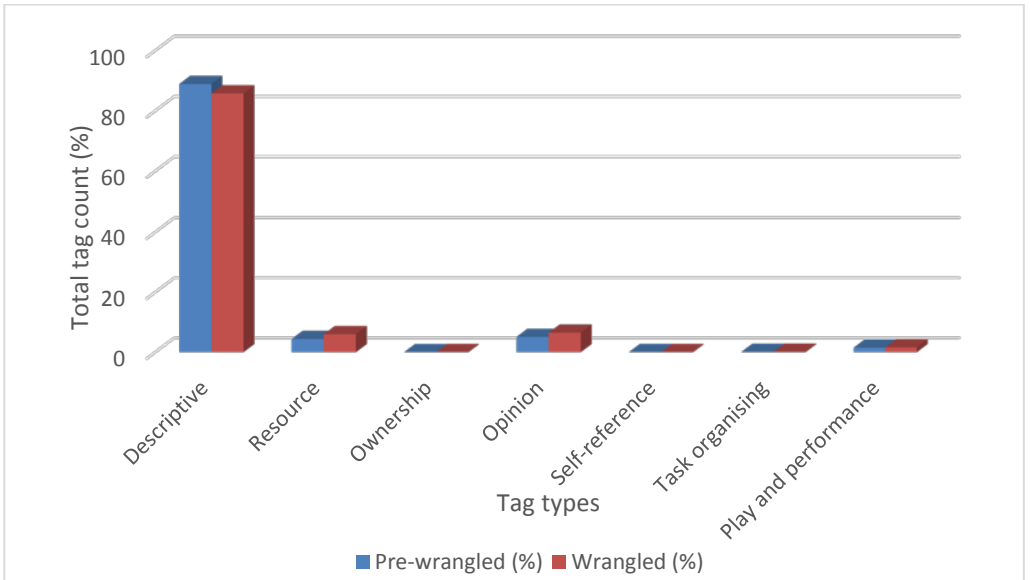
This knowledge capital could equate to some prestige in the fan community, but it is important to note that tag wranglers do not benefit from this, as they are unseen and anonymous.

Ownership type tags (see Table 1) were seldom employed as ownership is inherent in the post itself (i.e. the author of the story is the poster of the content). The Descriptive type was highest, particularly the Character sub-type, which indicated the characters present in the story (arguably characters are the most important aspect of a story, enabling readers to easily find the characters they want to read about). Communication was the second highest tag type employed, and surprisingly this type showed a marked percentage rise post-wrangling. This may be because Communication (sub)-type tags are idiosyncratic, and unlikely to be used more than once. Since their usage is so low, this means that they are rarely tag wrangled and merged with other tags. Therefore, their percentage of the total Communication type tag count tends to remain static, while other tag types, particularly Descriptive and Resource tags, tend to be readily merged with already-existing synonyms, which therefore reduces their percentage of the total tags within the post-wrangled dataset.

As with Ownership, there was negligible use of Self-reference or Task organising tags (less than 1%) in both datasets. There was some slight use of the Play and Performance tag type, and this was used in very specific contexts (indicating stories written as part of events, competitions, contests, challenges or games, which were mostly hosted on other social media sites). Examples of these were 'i accepted a few prompts', 'community: xmen15', 'secret mutant ficathon 2014' and 'x-men big bang challenge'.

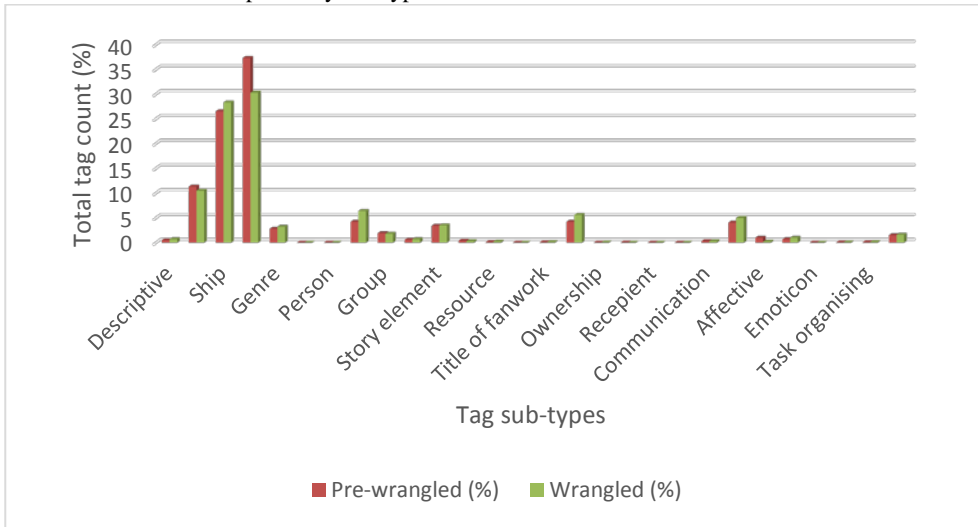
Figures 9 and 10 show a comparison of the total tag count used in the 'pre-wrangled' and 'wrangled' datasets. Figure 9 demonstrates that on the level of tag type, there is an almost negligible difference between the two sets. Figure 10, however, shows some significant disparities at the level of tag sub-type. These are at the following sub-types: Ship, Character, Friendship, Citation and Explanatory Communication.

Figure 9: Comparison of the percentage of the total tag count in the pre- and post-wrangled datasets, arranged by tag type (see Table 1 for tag types). There is little appreciable difference.



It seems that this is where the long tail of tags manifests itself most clearly. This long tail is made up of all the tags that are not popular enough to have been wrangled. However, it was noticed during the merging of pre- and post-wrangled tags that several more obscure tags had been ‘shoehorned’ into a standardised tag that did not encapsulate the specificity of its original meaning. To take the Citation sub-type as an example, the tag ‘Star-Lord and Kitty Pryde’, which is the title of a comic series, and of low popularity, has been tag-wrangled into a synonym of the much broader Fandom sub-type, ‘Marvel’. Likewise, the Ship tag ‘loroki’ (indicating a romantic pairing between the characters of Loki and Storm) has also been made a synonym of ‘Marvel’. In both cases the precise meaning of the original tags have been lost in the process of tag-wrangling, and both have also been classified under an inaccurate sub-type. While examples of this are in the minority, they are still frequent enough that it would suggest that, at the tag sub-type level, tag wrangling is slightly less successful than it is at the tag type level. Such funnelling of less popular tags into inaccurate tag sub-types (even if they are still in the same overall tag type) might account for the unusual spikes in the Friendship, Citation and Explanatory sub-types in the ‘wrangled’ dataset, and in the Character sub-type in the ‘pre-wrangled’ dataset.

Figure 10: Comparison of the percentage of the total tag count in the pre- and post-wrangled datasets, arranged by tag sub-type (see Table 1 for tag types). Ship, Character, Friendship, Citation and Explanatory sub-types show the most marked difference.



6. Discussion

AO3 tagging practices show high density and granularity, as authors attempt to convey the minute particulars of their fandom, as well as the plots of their stories. Fans are known to be particular about the types of fanworks they will engage with (Driscoll 2006), showing preference according to characters, ships, genres and kinks (i.e. the sexual predilections depicted in fanfic). All these elements and more are of primary importance, both for the reader, who wishes to find a fic that matches her preferences as precisely as possible, and for the creator, who wishes to draw as large an audience as possible to her work. Because of this, tagging – on AO3 in particular – becomes an important finding aid, similar to the subject headings found in library catalogues, except that they are far more granular and far more numerous in scope.

While there is the unique practice of ‘tag wrangling’ on AO3, this did not generally seem to affect the overall meaning or sense of the original tags used. Tag wranglers in the Marvel fandom appeared to have an in-depth expertise in their area, and, judging by the marked similarity between the pre- and post-wrangled AO3 datasets, they were for the most part correctly able to interpret and maintain the sense of the original tags. Indeed, the tag

wranglers appeared to be very dedicated to this mission, and to preserving the authenticity of the original tags, as Participants C and D opined respectively:

The AO3 Terms and Conditions and the Wrangling First Principles both strictly prevent us from being gatekeepery. We can't change tags, we can't tell users how to tag in any official capacity ("describe not proscribe"). Our goal is to organize tags in a way that fans will be able to find what they're looking for. To do that, we have to speak their language and use the words they use.

One of the most important principles of tag wrangling is that we don't alter a user's tags. The beauty of the AO3's system is that everyone can tag for whatever they want, in exactly the format they want. As well, most large fandoms have multiple wranglers assigned to them, and that means that there has to be a general consensus on how to handle any given tag that is for some reason challenging, or requires a judgement call of some kind.

In fact, these tag wranglers strongly felt that their work was a way of giving back to the community, in lieu of more traditional types of fan production, such as writing fanfiction and drawing fanart. Participant C and D respectively offered that:

Tag wrangling is a way I can contribute to a community that I love. I like this kind of work and, with the decline of livejournal, I felt less connected to the community and less like I was pulling my own weight. Wrangling both lets me meet people from across fandom and help out.

I consume a great many fanworks in my day to day life, but I don't really create that many. Tag wrangling is a way that I can feel as though I give something back to the community that has brought me so much joy.

Both interviewees rejected the idea that they were gatekeepers within their fandom. Participant C even went so far as to say "I don't think of myself as a gatekeeper, mostly because I hate that word". Despite this, I would contend that tag wranglers are information gatekeepers in the sense that they are, as Case (2012, 339) says, "shaping, emphasizing, or withholding" information, or the flow of information. This is with the caveat that they do not appear to be actively or intentionally withholding or emphasizing certain aspects of information within their fan community. Rather, they are shaping it in the sense of streamlining its flow, and facilitating greater access to it. In fact, Participant C noted that the reason why tag wrangling was important was that it facilitated greater access over a broad community:

Fan writing is increasingly centralized at AO3, while our day-to-day fannish expressions are ever more decentralized. I think fan writing is amazing and important, but there are sometimes some disconnects in how different parts of a fandom talk about a topic or a character. That shouldn't keep them from being able to see each others' work. For example, tagging your fic as "Romy" would keep it from being seen by people who weren't familiar with that smushname [i.e. a portmanteau of two character's names] unless a wrangler hooked them together on the backend.

The idea that tag wrangling assists in the streamlining of information is supported by the fact that far more co-occurring tags in this dataset had a higher betweenness centrality when compared to the other sites studied in the original doctoral study, Tumblr and Etsy (Price 2017, chapter 5). This means that more tags on AO3 acted as points of information exchange than either of the other two. In essence, tags on AO3 were more effective bearers of information. One might deduce, therefore, that both AO3's tag wranglers, and its curated folksonomy, are a very effective method for mitigating the less predictable effects of online tagging.

Further investigation, incorporating the views of more tag wranglers, as well as AO3 users, would be interesting. This would allow a better understanding of the tag wrangler's role as a democratic indexer, and would allow us to engage with user perceptions of the system and how it works. While tag wranglers exert a great amount of control over AO3's tagging system when compared to platforms such as Twitter and Tumblr, it is an 'invisible control', as on the surface the tags themselves are not changed, but merged with synonym, and categorised under a parent tag. The tag wranglers interviewed had positive views on the tagging system, especially considering the "insane strain it's under" (Participant C). Participant C felt that the filtering system could be better streamlined in order to increase retrieval accuracy. Participant D was largely satisfied with the system, but felt the "most changes that could be proposed would have more to do with changes in policy":

For example, there are an unfortunate number of tags floating about that can't be wrangled because users entered them in the wrong field, [and] if you put "Tony Stark" in the Fandom field, we can't make it a synonym of Tony Stark the character tag. Changing the type of a given tag is changing what a user entered in a way that we don't do as a matter of policy, and it's a policy I have to agree with.

This is very indicative of the trade-off between 'messiness' and control that is so often seen in folksonomic systems (Smith 2008). While AO3 suffers in some ways from maintaining this balancing act, on the whole it seems to be doing it successfully.

7. Conclusions

AO3 shows us that tagging and folksonomies are used by fans in a variety of ways – not merely for reasons of classification and organisation, but also for creative, affective and dialogic purposes. A tagging system should be flexible enough for fans to use it in any of these ways, but may be labour-intensive to run. AO3's tag wrangling system seems to be largely effective, having achieved a method for linking synonymous tags, which Lu, Zhang and He (2016, 677), Chen and Ke (2013), and Rafferty (2010) have suggested as a desirable function for improving tag retrieval. This curated folksonomy is an innovative solution to the messiness of folksonomies that on the whole successfully standardises fan taxonomies

without losing the original tagger's intended meaning. Such a system may be implemented in wider contexts, and should be of great interest to knowledge organisers and information architects. There is much we can learn from AO3's tagging system about the ways in which both platform creators and users can come together to create knowledge organisation systems which are best-gearred towards user warrant. This can be a monumental task – but AO3's tagging system shows that with passion and even obsession as a driver, much can be achieved with little. In future studies, it would be useful to interview more tag wranglers about their work, as well as users themselves. It would also be interesting to do more research into the challenges wranglers face, both in terms of warrant, and the technological “strain” that Participant C referred to.

References

- Ådland, Marit Kristine, and Lykke, Marianne. 2012. “Social Tagging in Support of Cancer Patients’ Information Interaction”. In *Social Information Research*, eds. Gunilla Widen and Kim Holmberg. Bingley: Emerald Group Publishing, 101-128.
- Bawden, David, and Robinson, Lyn. 2012. “Introduction to Information Science”. London: Facet Publishing.
- Beghtol, Clare. 1986. “Semantic Validity: concepts of warrant in bibliographic classification systems”. *Library Resources & Technical Services* 30, no. 2: 109-125.
<https://utexas.instructure.com/courses/1028478/files/29266064/>
- Bolognesi, Marianna. 2016. “Flickr® Distributional Tagspace: evaluating the semantic spaces emerging from Flickr® tags distributions”. In *Big data in cognitive science: from methods to insights*, ed. Michael. N. Jones. Hove: Psychology Press, 144-173.
- Bullard, Julia. 2014. “Values and Negotiation in Classification Work”. In *CSCW Companion '14: Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing*, 15-19 February 2014, Baltimore, Maryland, USA. New York: ACM, 45-48.
- Bullard, Julia. 2017. “Warrant as a Means to Study Classification System Design”. *Journal of Documentation* 73, no. 1: 75-90.
- Carrington, Peter J., and Scott, John. 2011. “Introduction”. In *The SAGE Handbook of Social Network Analysis*, ed. John Scott and Peter J, Carrington. London: SAGE, 1-10.
- Cattuto, Ciro; Schmitz, Christoph; Baldassarri, Andrea et al. 2007. “Network Properties of Folksonomies”. *AI Communications* 20, no. 4: 245-262.
- Chen, Ya-Ning, and Ke, Hao-Ren. 2013. “An Analysis of Users' Behaviour Patterns in the Organisation of Information: a case study of CiteULike”. *Online information review*, 37, no. 4: 638-656.
- Cole, Samantha. 2019. “An Internet Fan Fiction Archive Is Nominated for a Hugo”. *Motherboard*. April 3 2019. https://motherboard.vice.com/en_us/article/vbw9eb/internet-fan-fiction-archive-ao3-hugo-award

- Driscoll, Catherine. 2006. "One True Pairing: the romance of pornography and the pornography of romance". In *Fan fiction and fan communities in the age of the internet*, eds. Karen Hellekson and Kristina Busse. Jefferson: McFarland, 79-96.
- Estrada, Liliana Melgar; Hildebrand, Michiel; de Boer, Victor; van Ossenbruggen, Jacco. 2017. "Time-Based Tags for Fiction Movies: comparing experts to novices using a video labeling game". *Journal of the Association for Information Science and Technology*, 68, no. 2: 348-364. <http://dx.doi.org/10.1002/asi.23656>
- Golder, Scott A., and Huberman, Bernardo A. 2006. "Usage Patterns of Collaborative Tagging Systems". *Journal of Information Science*, 32, no. 2: 198-208. <http://dx.doi.org/10.1177/0165551506062337>
- Gursoy, Ayse. 2015. "Evaluating Fan Fiction Metadata for Preservation Use". In *ASIST '15 Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*, St. Louis, MO, USA, 6-10 November 2015, 52: 1-4. <https://doi.org/10.1002/pr2.2015.1450520100114>
- Hart, Chris; Shoolbred, Michael; Butcher, David and Kane, David. 1999. "The Bibliographical Structure of Fan Information". *Collection Building* 18, no. 2: 81-89.
- Hawksey, Martin. 2017. "Making the Complex Less Complicated: an introduction to social network analysis". MASHe, 3 June. <https://mashe.hawksey.info/2017/06/making-the-complex-less-complicated-an-introduction-to-social-network-analysis/>
- Haythornthwaite, Caroline. 1996. "Social Network Analysis: an approach and technique for the study of information exchange". *Library & Information Science Research*, 18, no. 4: 323-342. [http://dx.doi.org/10.1016/S0740-8188\(96\)90003-1](http://dx.doi.org/10.1016/S0740-8188(96)90003-1)
- Hendriks, Bas; Dormans, Stefan; Legendijk, Arnoud and Thelwall, Mike. 2016. "Understanding the Geographical Development of Social Movements: a web-link analysis of Slow Food". *Global Networks*, 17, no. 1: 47-67. <http://dx.doi.org/10.1111/glob.12153>
- Hidderley, Rob, and Rafferty, Pauline. 1997. "Democratic Indexing: an approach to the retrieval of fiction". *Information Services and Use*, 17 nos. 2/3: 101-109.
- Jalalimanesh, Ammar, and Yaghoubi, Seyyed Majid. 2013. "Application of Social Network Analysis in Interlibrary Loan Services". *Webology*, 10, no. 12: article 108. <http://www.webology.org/2013/v10n1/a108.html>
- Jenkins, Henry. 2006. "Fans, Bloggers, and Gamers: exploring participatory culture". New York: New York University Press.
- Jiang, Chunlin, Zhang, Liwei and Liu, Shengbo. 2014. "Social Network Analysis of the Interlocking Editorship among Chinese Library and Information Science Journals". *Journal of the China Society for Scientific & Technical Information*, 33, no. 5: 481-490.
- Johnson, Catherine A. 2004. "Choosing People: the role of social capital in information seeking behaviour". *Information Research*, 10: paper 201. <http://www.informationr.net/ir/10-1/paper201.html>
- Johnson, J. David. 2019. "Network Analysis Approaches to Collaborative Information Seeking in Inter-Professional Health Care Teams". *Information Research* 24: paper 810. <http://www.informationr.net/ir/24-1/paper810.html>

- Johnson, Shannon Fay. 2014. "Fan Fiction Metadata Creation and Utilization within Fan Fiction Archives: three primary methods". *Transformative Works and Cultures*, 17. <http://dx.doi.org/10.3983/twc.2014.0578>
- Lee, So Young, Kim, Hyang Mi; Chu, Kyounghee and Seo, Jungchi. 2013. "Fandom as a Prosumer: study on information behavior of fandom". *The Journal of Digital Policy and Management*, 11, no. 12: 747-759. <http://dx.doi.org/10.14400/JDPM.2013.11.12.747>
- Lu, Chao, Zhang, Chengzhi and He, Daqing. 2016. "Comparative Analysis of Book Tags: a cross-lingual perspective". *The Electronic Library*, 34, no. 4: 666-682. <http://dx.doi.org/10.1108/EL-03-2015-0042>
- Ma, Feicheng, and Li, Yating. 2014. "Utilising Social Network Analysis to Study the Characteristics and Functions of the Co-Occurrence Network of Online Tags". *Online Information Review*, 38, no. 2: 232-247. <http://dx.doi.org/10.1108/OIR-11-2012-0124>
- Malik, Aqdas; Johri, Aditya; Handa, Rajat; Karbasian, Habib and Purohit, Hemant. 2018. "How Social Media Supports Hashtag Activism Through Multivocality: a case study of #ILookLikeanEngineer". *First Monday*, 23, no. 11. <https://doi.org/10.5210/fm.v23i11.9181>
- Marin, Alexandra, and Wellman, Barry. 2011. "Social Network Analysis: an introduction". In *The SAGE Handbook of Social Network Analysis*, ed. John Scott and Peter J, Carrington. London: SAGE, 11-25.
- Otte, Evelien, and Rousseau, Ronald. 2002. "Social Network Analysis: a powerful strategy, also for the information sciences". *Journal of Information Science*, 28, no. 6: 441-453. <http://dx.doi.org/10.1177/016555150202800601>
- Park, Han Woo, and Thelwall, Mike. 2003. "Hyperlink Analyses of the World Wide Web". *Journal of Computer Mediated Communication*, 8, no. 4. <http://dx.doi.org/10.1111/j.1083-6101.2003.tb00223.x>
- Price, Ludovica Wing Sheun. 2017. "Serious Leisure In The Digital World: exploring the information behaviour of fan communities". Ph.D. thesis, City, University of London.
- Rafferty, Pauline. 2010. "Genre Theory, Knowledge Organisation and Fiction". In *Paradigms and conceptual systems in knowledge organization: Proceedings of the eleventh international ISKO conference*. February 23-26, Rome, Italy, eds. Claudio Gnoli and Fulvio Mazzocchi. Würzburg: Ergon-Verlag, 254-261.
- Rafferty, Pauline, and Hilderley, Rob. 2007. "Flickr and Democratic Indexing: dialogic approaches to indexing". *ASLIB Proceedings* 59, nos. 4/5: 397-410.
- Rose, Sarah Marita. 2013. "There is a Conversation in my Search: differing uses for tags". In *From collections to connections: turning libraries "inside-out"*. The 21st International BOBCATSSS Conference, 23-25 January 2013, Ankara, Turkey, eds. Zehra Taskin et al. Ankara: Hacettepe University Department of Information Management, 28-31.
- Rossi, Luca, and Giglietto, Fabio. 2016. "Twitter Use During TV: a full-season analysis of #serviziopubblico hashtag". *Journal of Broadcasting & Electronic Media*, 60, no. 2: 331-346. <https://doi.org/10.1080/08838151.2016.1164162>
- Saha, Arpita, and Mukhopadhyay, Parthasarathi. 2016. "Ranking OER Providers in India: a webometric analysis". *International Research: Journal of Library & Information Science*, 6, no. 1: 134-143.

- Small, Tamara A. 2011. "WHAT THE HASHTAG? A content analysis of Canadian politics on Twitter". *Information, Communication & Society*, 14, no. 6: 872-895.
<http://dx.doi.org/10.1080/1369118X.2011.554572>
- Smith, Gene. 2008. "Tagging: people-powered metadata for the social web". Berkeley, CA: New Riders.
- Thelwall, Mike. 2004. "Link Analysis: an information science approach". London: Elsevier.
- Thelwall, Mike. 2017. "Book Genre and Author Gender: Romance>Paranormal-Romance to Autobiography>Memoir". *Journal of the Association for Information Science and Technology*, 68, no. 5: 1212-1223. <http://dx.doi.org/10.1002/asi.23768>
- Trant, J. 2009. "Studying Social Tagging and Folksonomy: a review and framework". *Journal of Digital Information*, 10, no. 1: 1-44.
- Vaidya, Praveenkumar, and Harinarayana, N. S.. 2016. "The Comparative and Analytical Study of Librarything Tags with Library of Congress Subject Headings". *Knowledge Organization* 43, no. 1: 35-43.
- Wang, Mengmeng, and Iwaihara, Mizuho. 2015. "Hashtag Sense Induction Based on Co-Occurrence Graphs". In *APWeb 2015: Web Technologies and Applications*, 18-20 September 2015, Guangzhou, China, eds. Reynold Cheng et al. *Lecture Notes in Computer Science*, 9313, Cham: Springer, 154-165.
- Wang, Rong; Liu, Wenlin and Gao, Shuyang. 2016. "Hashtags and Information Virality in Networked Social Movement: examining hashtag co-occurrence patterns". *Online Information Review*, 40, no. 7: 850-866. <https://doi.org/10.1108/OIR-12-2015-0378>