

# Towards Better Visual Tools for Exploring Wikipedia Article Development – The Use Case of “Gamergate Controversy”

**Fabian Flöck**

Computational  
Social Science Group  
GESIS - Leibniz Institute  
for the Social Sciences  
fabian.floeck@gesis.org

**David Laniado**

Social Media Research Group  
Barcelona Media  
david.laniado@  
barcelonamedia.org

**Felix Stadthaus**

Institute AIFB  
Karlsruhe Institute  
of Technology  
felix.stadthaus@student.kit.edu

**Maribel Acosta**

Institute AIFB  
Karlsruhe Institute  
of Technology  
maribel.acosta@kit.edu

## Abstract

We present a comparative analysis of three tools for visually exploring the revision history of a Wikipedia article. We do so on the use case of “Gamergate Controversy”, an article that has been the setting of a major editor dispute in the last half of 2014 and early 2015, resulting in multiple editor bans and gathering news media attention. We show how our tools can be used to visually explore editor disagreement interactions and networks, as well as conflicted content and provenance, and present some of the results. We argue that these individual tools could benefit from synergies by combining them and lay out a possible architecture for doing so.

## 1 Introduction

Numerous scientific approaches exist to analyse Wikipedia articles in terms of their content, editors and social dynamics. Some visual tools have been provided, both by researchers and by the Wikipedia community. Yet, understanding the collaborative writing history of an individual article as a casual user, editor or even researcher in an easy, intuitive way (i.e., without relying on elaborate statistical analysis) is still a hard task. There is a lack of transparency regarding the editing process on Wikipedia: it is fully documented in the revision history, but not in a way that is straightforward to browse, inspect and analyze by humans in all its intricacy. For instance, one cannot easily discover which words were contributed by what author or what specific dynamics governed the rise of disagreement between editors on particular content in the article. This information would be key to enable *accountability and social transparency*, as has been argued by Suh et al. (2008), but is effectively hidden from the user due to the innate complexity.

Some related visual interfaces as, e.g., “Wikidashboard” (Suh et al. 2008), “Wikitrust” (Adler, de Alfaro, and Pye 2010), “History Flow” (Viégas, Wattenberg, and Dave 2004) and community solutions have been proposed;<sup>1</sup> but most have since been discontinued as a service and further, so we argue here, provide only solutions to specific subproblems of the complex phenomenon that is understanding the collaborative writing of a Wiki article.

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

<sup>1</sup>For community tools see <http://en.wikipedia.org/wiki/Wikipedia:Tools>

We hence argue in this paper that better software tools that allow end-users to visually explore the dynamic, collaborative process of adding, deleting and restoring specific content have to be made available for the purpose of accurately reflecting all relevant interactions of editors with each other and the emergence of content; and that those should be *integrated* to allow a seamless exploration of all relevant editing activity. We also make the case why such tools would be key to (i) enable more transparency and thus reduce the complexity that is inherent to the writing processes of (especially controversial) articles on Wikipedia; and (ii) why such transparency empowers readers, editors and researchers to better comprehend the context of an article’s emergence, and to interpret its content accordingly (e.g., by acknowledging opinion camps and biased behaviors, ownership or edit warring behavior of single authors, etc. and the effect they have on the eventual content presented).

**Structure of the paper** We first motivate the need to make understanding content emergence and collaboration patterns more intuitive in section 2. We do so specifically on the example of the article “Gamergate controversy” that has gained recent media attention as a major battleground of Wikipedia editors.

We then shortly present two tools we recently developed and published in this regard in Section 3: Contropedia and whoVIS; and we introduce whoCOLOR, a userscript for highlighting provenance and controversy of words in Wikipedia articles. Further we briefly describe the above mentioned older legacy tools and one Wikipedia community solution called “Wikireplay”, which we believe could also provide a useful approach to our goal.

Thirdly we demonstrate how our tools can help understand the dynamics of conflict and content development on the example of the “Gamergate Controversy”<sup>2</sup> article, hence enabling a reader of the article to see how the content readable today evolved into its current form, and what content should be interpreted with specific care considering its background. In the process we show the individual strengths of the three tools and how they can complement each other.

Finally, we outline how the individual tools could be combined to better tell “the whole story” of the article develop-

<sup>2</sup>[https://en.wikipedia.org/wiki/Gamergate\\_controversy](https://en.wikipedia.org/wiki/Gamergate_controversy)

## History [ edit ]

As video game production developed into a burgeoning industry, games became an increasingly consumer-oriented product focused on appealing to gamers with satisfying solitary experiences. People who had grown up playing these games developed a "gamer" identity that was associated with these early experiences. As early gamers were predominantly male this is also seen as having contributed to gendered interpretations of the identity. The emergence of the industry also gave rise to numerous publications specializing in the coverage of video games and catering for the interests of gaming enthusiasts. **Some early publications were even owned by manufacturers themselves.** Such publications were seen by industry leaders as a means of promotion for their products rather than sources for honest critical assessment and there was recurring criticism of the close relationship between gaming journalists and the major gaming companies.<sup>[1][2]</sup> **The growing popularity of games among casual consumers led to more accessible technologies such as the handhelds and smartphones expanded the audience for the industry to include many who did not fit the mold of the traditional hardcore gamer. As games had come to be seen as an art form rather than a product, games centring on social issues grew in popularity. This increasing perception of games as an promoted gaming publications to show towards cultural criticism of the games. Independent video game development, which allows developers to release titles without publisher interference, has made these games more common.**<sup>[14][35][7]</sup>

The growth of the gaming audience also brought in many female gamers whose primary gaming interests did not conform to those of the male-oriented gamer identity, and who began to question some assumptions and tropes that were historically used by game developers. In light of the growing female audience for games, and growing female representation in the gaming industry, outlets became increasingly interested in detailing issues of gender representation in video games.<sup>[12]</sup> One prominent feminist critic of the representation of women in gaming is Anita Sarkeesian, whose *Tropes vs. Women in Video Games* project is devoted to criticism of female stereotypes in games. Her initial Kickstarter to raise funds for the series and her subsequent videos have all been met with hostile commentary and harassment from hardcore gamers. Further incidents, such as those concerning Jennifer Hepler raised

(a) Before the merging of the draft article and editor bans put into effect (22.11.2014) we see (i) a clear overall dominance by a small group of authors (cf. section 4, esp. Figure 4a) in written words and (ii) in some sections, like "History", an almost complete dominance of these authors' content.

## Author List

Ryulong	27.1%
Masem	21.7%
NorthBySouthBaranof	17.4%
The Devil's Advocate	11.1%
Hahnchen	7.5%
Starship.paint	1.5%
Kenc0618	1.4%
Willhesucceed	1.4%
Loganmac	1.3%
Diego Moya	1.1%
Halfhat	0.9%
TheRedPenOfDoom	0.7%
Sookenen	0.7%
Koncorde	0.5%
Bilby	0.4%
ChrisGualtieri	0.4%
Torga	0.4%
GamerProb64	0.3%
MyMolobooaccount	0.3%

## History [ edit ]

In February 2013, independent game developer Zoe Quinn released *Depression Quest*, an interactive fiction browser game, through the game's website. Though the game was met positively by critics, a backlash developed among some gamers who believed that the game received an undue amount of attention, especially after a planned Steam distribution platform release. Quinn began to receive hate mail over the game upon its release and criticism from some parts of the Steam user community receiving enough harassment to cause her to change her phone number.<sup>[1][2][3][4][5]</sup> By September 2014, Quinn had been the target of eighteen months of increasing harassment, which had created "an ambient hum of menace in her life; albeit one that she has mostly been able to ignore."<sup>[6]</sup>

Shortly after the Steam release of *Depression Quest* in August 2014, Quinn's former boyfriend, Eric "Eron" Eron, wrote a blog post, described by *The New York Times* as a "rambling online essay,"<sup>[7]</sup> alleging among other things, that Quinn had an affair with Kotaku journalist Nathan Grayson.<sup>[8]</sup> Quinn's detractors falsely alleged that the relationship had induced Grayson to publish a favorable review of the game.<sup>[9][10][11][12][13][14]</sup> Kotaku editors-in-chief Stephen Totilo, affirmed the existence of a relationship, but clarified that Grayson had not written anything about Quinn after the relationship had commenced, and had never reviewed her games, though he did acknowledge a piece written before the two began their relationship.<sup>[10][9]</sup> Commentators in and outside the gaming industry denounced the attack on Quinn as misogynistic and unfounded.<sup>[5][15][17]</sup>

As a result of these allegations, Quinn and her family were subjected to a virulent, often misogynistic harassment campaign, including doxing, threats of rape, hacks of her Tumblr, Dropbox and Skype accounts, and death threats.<sup>[2][9][8][18]</sup> One threat read, as reported by *The New Yorker*, "Next time she shows up at a conference we'll give her a crippling injury that's never going to fully heal. It's good solid injury to the knees. I'd say a brain damage, but we don't want to make it so she ends up too retarded to



## Author List

Masem	17.5%
Ryulong	16.4%
Tony Sidaway	13.2%
NorthBySouthBaranof	8.3%
The Devil's Advocate	6.3%
Strongjam	4.8%
Krano	3.8%
Aquillon	2.9%
Hahnchen	2.8%
Carwil	2.7%
Bosstopher	2.4%
TheRedPenOfDoom	2.4%
Starship.paint	2.3%
Kenc0618	1.7%
PresN	1.1%
Torchiest	0.9%
Torga	0.8%
Rhokar	0.6%
Koncorde	0.5%

(b) Today, much of the content of the previously dominating editors has been replaced (as of 30.03.2015), as well as content by other, banned editors. The content of the previously dominating editors has (i) decreased overall, as the comparison to Figure 4a shows. (ii) Some sections, like "History", are now much more diverse in terms of authorship (non-marked words in the above screenshot were, e.g., written by 22 distinct editors)

Figure 1: Screenshots of the "History" section of "Gamergate controversy" with whoCOLOR markup on the text (on the left sides, respectively) and the author lists (right sides, respectively), ordered by percentage of written words. The same editors are marked in both screenshots. Except Masem and Torga, only editors that were later banned are marked.

ment and problematic content, and suggest a concrete plan for integration of these approaches.

**Contributions** In this this paper we (i) give an overview of software tools directly related to the problem of visually explaining the editing history of an article; (ii) show how the tools we are currently developing can help in tackling the issue of insufficient transparency of the editing process and content emergence; (iii) show how an integration and further development of tools can provide a user with insights each of those tools could not provide individually; (iv) "on the side", conduct a first scientific, although preliminary, qualitative analysis of the evolution and disputed content of the "Gamergate controversy" article, which we employ as a use case.

## 2 Motivation

We aim to provide end-users with tools to understand the collaborative writing process in its complexity. A more specific goal is to enable even casual readers to understand the social collaboration patterns that governed the emergence of the content – leading it to be as it is at the time of reading – and to exercise care especially with controversial text. This supplies context for the interpretation of said content, linking it to an explicit, understandable history.

### "Gamergate Controversy"

A specific example of an article for which it is hard to attain the full picture of all parties involved in its creation, and which content they have been arguing about, is "Gamergate Controversy".

It is a highly controversial article on the English Wikipedia that has recently garnered even the attention of prominent media outlets, when 11 editors were sanctioned by Wikipedia's "Arbitration Committee" (ArbCom), mainly including 1-year topic-bans on the article and related topics with most participants disciplined for "uncivilized" and "battleground" behavior.<sup>3</sup> The sole involvement of the ArbCom, arguably Wikipedia's "Supreme Court" when it comes to quarrels between editors, plus the scope of the sanctions shows the gravity of the dispute this article has been subject to.<sup>4</sup>

The conflict surrounding the "Gamergate" phenomenon – going far beyond Wikipedia – had already become of significant societal relevance, as seen by the coverage in many major media outlets like *The Washington Post*, *The New Yorker* and *The New York Times* (Kaplan 2014; Parkin 2014; Wingfield 2014). After "Gamergate" was discussed in social media and news media, the Wikipedia article "Gamergate Controversy" was created and specifically the related editing dispute and resulting editor bans came into focus of the news media as well, as indicated by the coverage in media outlets like *The Guardian*, *The Washington Post* and *Slate* (Hern 2015; Dewey 2015; Marcotte 2015).

The article deals mainly with alleged corruption in gaming journalism and the following reported sexism and harassment

<sup>3</sup>Find the full ArbCom case here: <https://en.wikipedia.org/wiki/Wikipedia:Arbitration/Requests/Case/Gamergate>

<sup>4</sup>Usually users in the "administrator" role police articles and implement sanctions. ArbCom is only appealed to if those sanctions fail repeatedly to restore order to the editors' behavior.

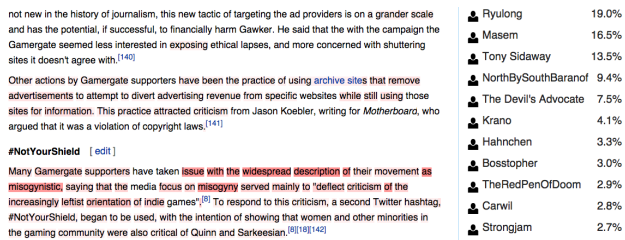


Figure 2: A paragraph with heavily controversial content in a recent version of the article, as seen in the whoCOLOR Conflict View (as of 09.02.2015). Shades of red are relative to each other, being more intense the more conflicted a word is.

of (mostly female) individuals through the “anti-corruption” or “pro-Gamergate” faction.<sup>5</sup>

All in all, the article is a good example of how Wikipedia coverage of a topic can achieve (i) high societal relevance and how it is therefore important for readers to understand the underlying motivations, agendas and editors fighting for control over the article *and* the effect of those disagreements on the eventual output in form of the content presented to the readers. Only this transparency enables a critical and informed consumption of the information therein. In a case like this, moreover, (ii) the number of editors involved and the amount of content changes they have been applying over time is so vast that the resulting patterns of editor interactions and content development are not viable to understand in their entirety for an average reader (or, e.g., a journalist) just by looking at the revision history provided in the MediaWiki software or reading the talk page. As journalist Amanda Marcotte puts it in an article for “Slate” in direct reference to the Wikipedia article: “As is generally the case with Gamergate, piecing together the story of what really happened amid the cacophony of finger-pointing and recrimination is nearly impossible [...]” (Marcotte 2015).

Providing suited visual tools to explore the article history in terms of content and editor interactions is therefore essential to assure *fully informed readers*.

### 3 Related work

The following visualization and exploration software tools can potentially be used to shed light on the development of the article by an end-user.

- **whoVIS** (Flöck and Acosta 2015) visualizes editor-editor disagreement networks over time, derived from the collaborative editing actions on word level in an article. It features a main network graph view, navigable over time, and allowing to inspect individual edges for the disagreed-on content.<sup>6</sup> It also provides views for aggregated metrics

<sup>5</sup>The authors want to stress that they aim to depict this topic from a scientific perspective, as neutrally as possible, and have absolutely no intention to take any position in this matter.

<sup>6</sup>whoVIS is available as web-based demo at: <http://km.aifb.kit.edu/sites/whovis/index.html>.



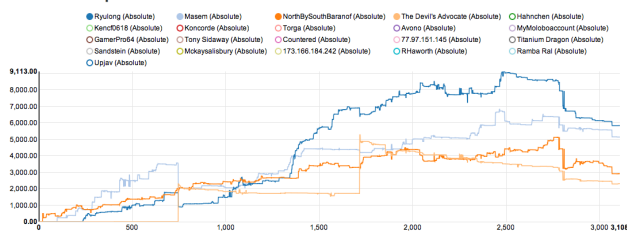
Figure 3: The “Word History” feature of whoCOLOR is used to inspect the most controversial words from Figure 2. Shown: a time period (descending from most recent to older) where the marked-up words were heavily contested. On the left, one can see the proponents (names under blue bars) and opponents for inclusion (names under white gaps), next to information about revisions and time duration the words were ex-/included.

over time and tracking most dominant editors in terms of written words.

- **Contropedia** (Borra et al. 2015b) highlights most controversial elements in an article, and when and why there was dispute about them. Two main views are the entry point to inspect activity around a specific topic: the *layer view*, providing an overlay for the original article, highlighting controversial elements, and the *dashboard view*, that presents a ranking of the most controversial elements together with a timeline, showing when each element underwent most dispute activity and the users involved.
- **WikiDashboard** (Suh et al. 2008) visualizes edits over time by contributors to an article in a graph above the article content, but does not track content changes or conflicts. The service is no longer officially available.
- **WikiTrust** (Adler, de Alfaro, and Pye 2010) provides a browser add-on that adds an overlay to Wikipedia articles to display estimated trustworthiness of content, according to the longevity of introduced changes. It provides word provenance information, but not interactions of editors. The API providing the trust mark-up was discontinued.
- **History Flow** (Viégas, Wattenberg, and Dave 2004) creates a layer-like visualization of the different parts of the article written by distinct editors, over the revision history. In this way it helps to follow content changes and moves over time, although the concrete content or disputes are not visible and content attribution is performed on sentence, not word granularity.
- **Wikireplay**<sup>7</sup> (or “re Edit”) is a community-built web application that allows the user to select a Wikipedia article

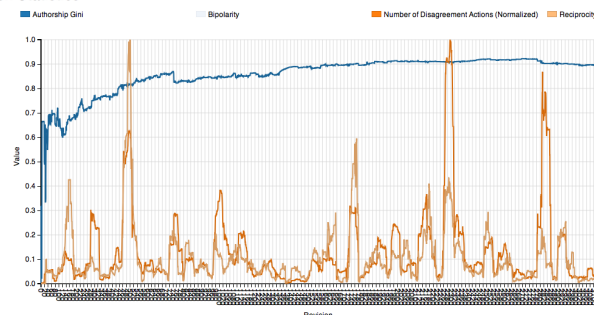
<sup>7</sup><http://cosmiclattes.github.io/wikireplay/player.html>, by Wikipedia user Jeph Paul

Word Ownership Distributions



(a) Line graphs of the amount of words owned in a specific revision, over all revisions, by the top four contributors of content to “Gamergate controversy”. All the other editors’ own word shares are notably lower than those depicted.<sup>13</sup>

Graph Statistics



(b) The Gini coefficient for authorship, the relative disagreement between editors and the relative mutual disagreement, for “Gamergate controversy”, over the whole revision history. We can see a constant increase of the authorship Gini (= increase in concentration) and clear spikes of disagreement at several time points.

Figure 4: Additional metrics provided by whoVIS for word ownership of the top editors (left) and conflict development and word ownership Gini coefficient (right).

and a starting revision. It then displays the look of the HTML view of the article at the given time and sequentially visualizes all single additions and deletions that took place in a video-like animation.

### Introducing whoCOLOR

whoCOLOR consists of a userscript meant to be loaded in a client’s browser with the Tamper-/Greasemonkey extensions.<sup>8</sup> It is activated when an article page in the English Wikipedia is loaded and queries the WikiWho API for data about the authorship of each individual word in the text, based on the WikiWho algorithm (Flöck and Acosta 2014).<sup>9</sup> The provenance data is sent to a server-side service of whoCOLOR which generates colored markup as an overlay on the actual article text, which is then displayed by the userscript (See examples in Figure 1). By hovering of words in the text, the user is notified of the author of the selected sequence through a highlight-effect in the author-list that is inserted on the right-hand side of the article content. By clicking on words or authors, the user can permanently highlight the author and all her written words in a distinct color. The approach is heavily inspired by very similar work done for the community solution “WikiPraise” by Wikipedia User Ne-taction, that was based on the now-defunct Wikitrust API.<sup>10</sup>

whoCOLOR also features a “Conflict View”. It colors those words in the article in a stronger tone of red the more deletes and reintroductions (hence: disputes) they were subjected to in the past (see Figure 2). A “Word History” is available in both views (see Figure 3). It can be used to mark

<sup>8</sup><https://tampermonkey.net>, <https://www.greasemonkey.net>

<sup>9</sup>The WikiWho API can be queried for the provenance information of a revision range from live article data. Beta tester version/s/access to whoCOLOR and the WikiWho API can be requested by email to the first author.

<sup>10</sup>See <https://de.wikipedia.org/wiki/Benutzer:NetAction/WikiTrust>.

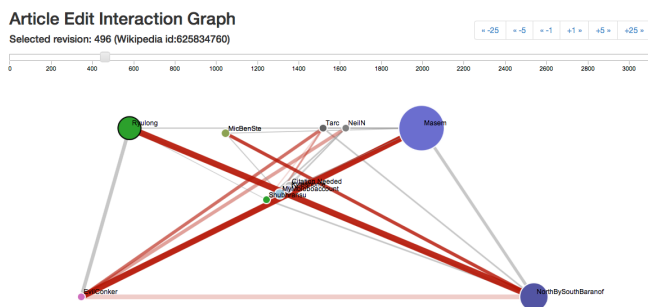
up a sequence of words with the mouse to show and inspect the periods of time when the selected words were present (blue background) in the article or when they were not (white background). It also shows who removed the content, for how long (see Figure 3), and which user reintroduced it. It can hence aid in understanding who the antagonists were in possible disputes indicated by the “Conflict View”.

## 4 Use Case: Exploring “Gamergate controversy”

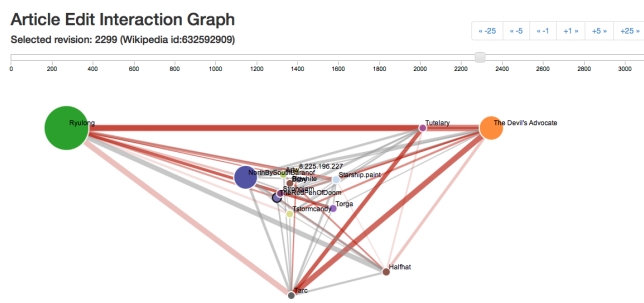
When reading about the dispute concerning “Gamergate controversy” in news media, on Wikipedia meta pages or other external sources, it is routinely portrayed as “pro-gamergate” against “feminists”, or at least the situation is outlined as a clear-cut, two-camps edit war.<sup>11</sup> The lecture of the ArbCom page on the case gives a vague impression of who the opponents in the dispute were: looking at the list of banned editors and consulting third party websites and the article talk page, one is prone to believe that the “pro-feminist” or “anti-gamergate” faction comprised 5 now-banned editors (sometimes even called “The Five Horsemen”).<sup>11</sup> On the other side we seem to have another – although even less clearly defined – “pro-gamergate”, “anti-feminist” group of 6-9 now-banned editors and several unnamed users. Yet, this vague picture is most likely a very strong simplification of the actual editing dynamics and actors in the article.

The specific *disputed contents* of the conflict are even harder to pinpoint, apparently ranging from wording disputes over using expressions like “misogyny”, “harassment” etc., to arguing about whether certain factual claims are correct, to disagreement about whether certain statements, quotes or sources belong in the article at all. But which exact formulations in the article are changed from what to what, which ones

<sup>11</sup>Cf., e.g., the news sources cited in Section 2, or <http://thinkprogress.org/culture/2015/03/06/3629086/wikipedia-gamergate-war/>



(a) Disagreements between four main editors at wrev 496. Masem, NorthBySouthBaranof, Ryulong and EvilConker show major disagreements, alongside various minor disagreements. Editor EvilConker’s content changes will be undone before the conflict ends.



(b) A spike of disagreements towards the later third of the article (cf. Fig. 4b) involves The Devil’s Advocate in disagreement with one of the dominant editors, Ryulong . Also shown: disagreements between these two editors and Tarc and Tutelary, respectively, two editors that were also later banned.

Figure 5: whoVIS disagreement network graphs for “Gamergate controversy”

are most disputed and between which editors these disagreements actually took place is very hard to discern only from the article itself, the associated talk page or third sources. To get an unbiased, first-hand picture ex-post, one would have to manually sift through a vast amount of revision-history and talk page data.

### Exploring with whoVIS

The “Word Ownership” view of whoVIS (“original word author” equals “owner”) provides first of all insights into who were the *most influential editors in terms of written words*. As shown in Figure 4a, four editors mainly coined the narrative of the article by having authored most of the text, especially in the hot phases of the debate (cf. disagreement spikes in Figure 4b): Ryulong, Masem, NorthBySouthBaranof and The Devil’s Advocate. Three of those editors were later banned.<sup>3</sup> The “Additional Metrics” view of whoVIS (not depicted) also shows that the article was at least semi-protected almost its entire life and has seen frequent full-protection periods.

We can glean from whoVIS that these users wrote significantly more content on the page than other editors (not shown in Figure 4a) and did so increasingly over time. The Gini coefficients measuring (i) the concentration of ownership of words (shown as the blue line in Figure 4b) and (ii) how concentrated the distribution of edit actions is over all active editors (not depicted, very similar trend) are in accordance with this *apparently increasing dominance of just a few editors*. Both curves are showing a steady incline of words owned and edits applied, that levels off in the last third of the article at a high value.

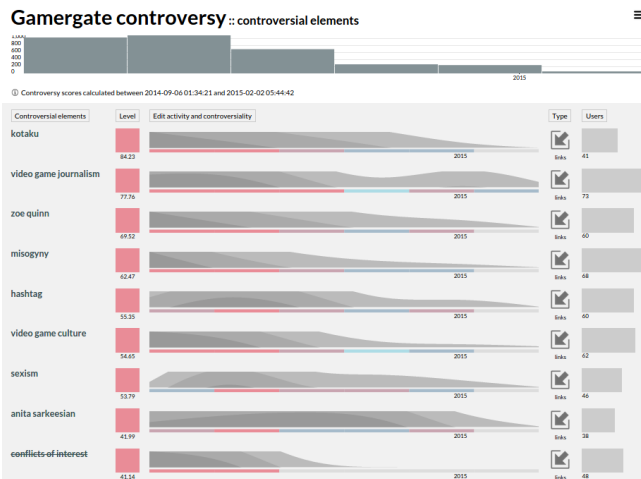
Just towards the end of the recorded revision history, we see unusual, distinct and synchronized drops in the amount of words owned for all four individuals (Figure 4a); upon inspection of the Wikipedia diffs corresponding to those revisions and associated comments, we learn that apparently, *the community has started a separate draft article* which was at these revisions merged into the original (hence removing or replacing much of the original content). The need for the page to be fully protected at times and Wikipedians to

start a parallel article draft *at all* is a strong indicator that the article climate up to this point was too unwelcoming for many editors. A reason for this could be the dominance of some authors in the article as well as the ongoing conflicts, depicted by the number of spikes in disagreed-on words and mutual disagreements in Figure 4b.

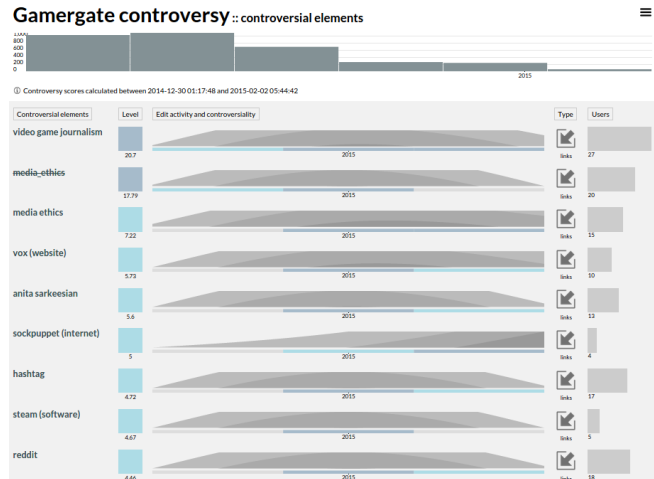
We therefore take a look at the disagreement network graph provided by whoVIS. The basic pattern visible starts at an early stage, from approximately whoVIS-revision (wrev) 400 (of approximately 3100 as of writing).<sup>12</sup> *Three main actors seem to dominate the stage, often strongly disagreeing with other editors and each other*: Ryulong, Masem and NorthBySouthBaranof (cf. Figure 5a). Frequently, other editors are involved in these disagreements, but never for equally long periods as these main actors, who are almost constantly in disagreeing relations. One example is an intense mutual disagreement of user EvilConker with Masem (Figure 5a) at around wrev 480 about how the introducing “Background” section should be written (indicated as well by the first major spike in Figure 4b). Eventually, the content by EvilConker is reverted back to the version before his intervention and the user ceases the conflicting interaction. Several of these short-lived, intense conflicts with various editors can be observed. Yet, some distinct antagonists emerge – although often only active temporarily – as e.g. users Torga and Diego Moya at around wrev 510, or user Titanium Dragon at wrev 640, to give just a few of many examples.

Although we cannot dive into the finer details extractable with the whoVIS tool here, certain patterns in these interactions become salient. (i) *There are constant challenges of the content written by the main three authors.* (ii) *The main three authors challenge each other significantly as well*, especially Ryulong and NorthBySouthBaranof. (iii) *The challenges of content by less-dominant editors seem rarely to result in their own content to be accepted in the article*, while the main three actors increase the amount of owned words, as we have

<sup>12</sup>wrev = whoVIS’s internal sequential numbering from the first revision of the article to the last, starting at 0. The respective Wikipedia revision-IDs can be gleaned from the whoVIS tool.



(a) Contropedia Dashboard View based on the complete edit history of the article, from September 6, 2014 to February 2nd, 2015.



(b) Contropedia Dashboard View for the last period of edits between December 30, 2014 and February 2nd, 2015.

Figure 6: Contropedia’s dashboard view for article “Gamergate controversy”: a ranking of the most controversial elements in the article, for the complete edit history (left) and after the draft article was merged twice, including the period when editor bans were declared and enforced (right).

seen in Figure 4a.

The only major exception to this rule seems to be editor The Devil’s Advocate, who, starting at around wrev 740, begins rewriting and adding much content in the article, consequently entering into major disagreements with other editors about his/her changes. The Devil’s Advocate is henceforth very active in disputes, as for example a major dispute starting at wrev 2250 (cf. Figure 4b, second-to-last major spike). Nonetheless, the editor can gradually increase influence in form of originally written words in the article, as seen in Figure 4a.<sup>13</sup>

As first conclusions of this very preliminary analysis, one could infer that (i) while certain camp-like behavior exists – e.g. many edit comments, especially in mutual disagreements, are trivializing edits and content by putting them into the dichotomous categories of either “pro-gamergate” or “anti-gamergate” – the reality of the edit interactions is much more complex than one would think from reading about clear-cut pro-gamergate and anti-gamergate editor camps. Somewhat clearer is the insight that (ii) it presumably early on became very hard and unwelcoming for “average” editors to sustainably contribute to the article without possessing a high degree of boldness and endurance, such as editor The Devil’s Advocate. Yet having only a group of such very bold (and possibly strongly opinionated) editors be the main actors and writers of an article might deter more moderate contributions, and moderating voices, and might have self-reinforced the climate that eventually caused the banning of many participants. As a conclusion for the article reader, the revisions of the article up to January 2015 should be read with the

<sup>13</sup>Hahnchen is the only other editor who contributed a sizable amount of content, but did so early on and without notable conflict (the amount stayed almost constant at around 1,000 words throughout the entire article – curve omitted from Figure 4a )

clear awareness about remnants of these – sometimes very intense and possible biased – editor disputes in the article. Even today a lot of content as a result from these disputes still persists, as we will see in the following section.

### Exploring with whoCOLOR

While with whoVIS we can explore the editor interactions and the contested content attached to them, this approach might be too abstract or complicated for a casual reader who is simply interested in which content is controversial or where it is coming from in the current article revision s/he is reading.

With whoCOLOR, the reader can retrieve information about the author and provenance of a word easily in the browser as an intuitive annotation of the text while reading the article.

We see in the provenance view, as shown exemplarily in Figure 1, that the merging of the draft article, bans and the activity of new editors seem to have had a diversifying effect. While before the imposed topic bans and intervention of new editors, some sections were written almost entirely by the previously discussed dominant editors (Figure 1a), currently<sup>14</sup>, (i) the overall share of words written by these users has dropped dramatically (although still high) and (ii) sections like “History” contain now content written by many different users (Figure 1b). While this is not necessarily a sign of quality in Wikipedia, it might be interpreted as such here, as it can be presumed that more points of view on the topic now found their way into the article.

Via the “Conflict View”, the user can also explore which the most contested parts of the content have been so far. In Figure 2 we see an example of a paragraph that was heavily disputed in the article. It involves a statement, to paraphrase,

<sup>14</sup>As of 30.03.2015

about “what description of their movement Gamergate supporters have taken issues with”. The inspection of this word sequence via the “Word History” feature further shows when the main dispute about those words happened, how long it lasted and who was involved (cf. Figure 3). We see that some previously mentioned main actors in the article were most involved in the dispute about whether or not to include those exact words in the article. This is also the case for many other disputes still remaining. By clicking on the arrow icons, one can navigate to the Wikipedia diffs of those edits to also see which words were proposed as an alternative.

Examining the whole article, the main controversial words still present are concentrated in individual paragraphs, mainly in the sections “History” and “Misogyny and antifeminism”, and distributed over the whole document in sequences about harassment, threads and alleged statements of individuals. A reader using whoCOLOR is alerted to those controversies and can hence interpret them accordingly.

### Exploring with Contropedia

While whoCOLOR provides a very detailed tracking of each single word, Contropedia focuses on *substantial*, disagreeing edits involving wiki links (Borra et al. 2015b). As a relevant concept or entity should be linked to the corresponding article, such linked elements are used as focal points around which content and activity are aggregated. The controversy score of an element is roughly based on the number of times a sentence including the corresponding wiki link was substantially edited in a disagreeing fashion. This approach reduces the information the user has to process by giving a condensed picture of the most central disputes. Contropedia’s dashboard view (Figure 6) presents a ranking of these most controversial elements (wiki links) in the article.

Figure 6a reveals that over the whole history of the article the elements that were subject to most dispute were “Kotaku”, “Video game journalism”, “Zoe Quinn” and “Misogyny”. The dashboard also unveils elements that were highly controversial in the history of the article but are currently not included as wiki links; such elements are shown struck-through. This is the case of “conflicts of interest” (visible in Figure 6a), “Doxxing”, “Christina Hoff Sommers” and “Harassment”. By expanding an element it is possible to see all the edits involving it, in a detailed table view. It becomes clear that among other editors, the principal editors previously discussed as well as many later-banned users were part of these main disputes.

The dashboard view also presents a timeline for each element, showing in which periods it was increasingly edited (gray horizon chart) and became more controversial (colored bar). It can thus be seen that some elements, such as “conflicts of interest” or “gamer”, were disputed only in the first two months of the article’s life, while others, such as “Sexism”, “Culture war” and “Milo Yiannopoulos” started to be heavily edited and disputed after the first month. Overall, conflict seems to have cooled down over the last period of the article’s life. This fact can also be seen more clearly when restricting the exploration to specific intervals of time; Figure 6b shows the dashboard view for the last period, in which editor bans were declared and after the draft article was

merged twice. Elements such as “Kotaku” and “Zoe Quinn”, at the center of the dispute when considering the whole history of the article, are no more among the most controversial. Only “Video game journalism” and “Hashtag” remain in the top of the ranking, while negotiations now take place around other topics, such as “Media ethics”, “Vox (website)” and “Sockpuppet (internet)”. Yet, the overall controversy has died down considerably.

By aggregating content and activity around wiki links, used as the entry points to explore controversies, the platform offers an easy access to inspect the development of single topics within an article, and helps the users to make sense of the edit history. On the other hand, this approach could in some case represent a limitation, for example in case of controversies in sentences that do not contain any wiki link, and therefore are not captured by the tool.

## 5 Envisioned Tool Synergies

An optimal visual tool aimed at enabling article development transparency should make several aspects available to the user: **(i)** the interaction patterns of the editors *with the content* (e.g. edit sprints or words added by one editor) and *with each other*, in terms of disagreement and how it is resolved (e.g. conflicts/edit wars and resulting controversial content); **(ii)** the development of the content over time (which content was there first, deleted/reintroduced, replaced by which other content, when was it disputed); **(iii)** a pre-selection and focus on the most important (e.g., controversial) (a) content, (b) users, and (c) time periods, so that the end-user does not have to explore the complete potential space of edit information; **(iv)** the overall “climate” of the article, given by aggregate metrics about editor behavior; **(v)** an interface to explore all these aspects over the revision history in one consolidated environment.

In Section 4 we have seen how the tools presented previously each allow distinct but complementary insights into the analyzed article. Some of their main features could hence be combined to benefit from this complementarity and offer one integrated solution.

### Integrating whoCOLOR, whoVIS and Contropedia

We envision an integrated platform that brings together the functionalities of all three tools. Such a platform could offer three main views, namely a “Controversy Mode”, a “Provenance Mode” and a “Editor Network Mode” for exploration of the article.

**Controversy Mode** In the Controversy Mode, the user could first explore the current version of the article with the Contropedia Layer View, for investigating the *most important* disputed concepts in the current article and in past revisions and explore their individual conflict history. For a less focused, but more fine grained analysis, the wikiCOLOR Conflict View mark-up of the text would be provided in the same fashion (not as userscript), highlighting *all* controversial words and allowing to study their individual history.

Similarly, also Contropedia’s Dashboard View of the most controversial elements could be complemented with a more

fine-grained view, showing the words that have been more frequently deleted and re-added. This view would also unveil words that are not currently present in the article but were important objects of dispute during its previous history.

As a constant top-navigation element in this Controversy Mode, a graph over time and (optionally) revisions would provide several article-wide conflict scores as now present in the Contropedia timeline and in the whoVIS disagreement and mutual disagreement line graphs. This graph would also be used to select specific time frames to restrict the calculation of the controversy scores of both views to certain time or revision frames.

Additionally, “Word Profiles” for users could be provided, highlighting the words most added and disputed by a user in a Tag-Cloud and their main antagonists.

**Provenance Mode** Switching to Provenance Mode, the user would see a view akin to the whoCOLOR authorship mark-up (with article content and mark-up stored on the platform), alongside the whoCOLOR authorship list. As a secondary option, the user could display the “Word Ownership” line graph of whoVIS to track the main contributors to the article and the Gini coefficients for word ownership and edits per editor.

Thirdly, an “Age View” could be provided, showing the oldest and most recently added words.

**Editor Network Mode** The Editor Network Mode would offer two network graph views: the whoVIS main editor network and the network view offered by Contropedia (Borra et al. 2015a). Both visualizations represent disagreement networks, but in two complementary ways: while the former represents mutual disagreements as distance in the graph, to show emerging user camps as clusters of users, the latter highlights actual interactions by placing closer editors that interacted more with one another, but uses colours to convey user camps, assigning the same colour to users who do not interact (i.e. disagree) with each other, but interact with the same other users. A third part of this mode would be the option to display additional metrics for the network per revision, such as those provided by whoVIS.

**Integration of Modes** The user would be able to jump between different modes, keeping focus on a specific element, or on a revision or timestamp or window of these. E.g., inspecting the editor network of revision 222, the user could directly switch to the controversy mark-up of revision 222 to learn about the most conflicted words at this revision.

Or, from an element or word in the dashboard view in the controversy mode, the user could inspect the social dynamics behind it by switching to the editor network and to the provenance mode for the element.

### Possible further additions

**Wikidashboard** A functionality to visually track edit activity as implemented in Wikidashboard could augment the Provenance and Controversy Modes.

**HistoryFlow** The graphical representation of word ownership by HistoryFlow would be an optimal extension of the

Provenance Mode for inspection over time, given the underlying calculation of authorship is adapted to word instead of sentence level, as provided by the WikiWho algorithm.

**Wikireplay** When inspecting a certain revision or window of revisions, each mode could be linked to a Wikireplay sequence of only that revision window or a window of 50 revisions leading up to the specified revision. In this way, users could spot interesting parts of the article history in one of the modes to then watch the replay of that specific editing period.

## 6 Conclusions

We have presented a comparative analysis of three tools for visually exploring the revision history of a Wikipedia article. We showed the results of applying our tools to the “Gamergate Controversy” article, exploring editor disagreement interactions and networks, as well as conflicted content and provenance. We argued that these individual tools could benefit from synergies by combining them, and we outlined a possible solution for doing so and for extending them further with other previously proposed tools. The envisioned combination of the features offered by these tools would improve transparency in Wikipedia, augmenting the experience of both readers and editors.

## References

- Adler, B.; de Alfaro, L.; and Pye, I. 2010. Detecting wikipedia vandalism using wikitrust. *Notebook Papers of CLEF* 1:22–23.
- Borra, E.; Laniado, D.; Weltevrede, E.; Mauri, M.; Magni, G.; Venturini, T.; Ciuccarelli, P.; Rogers, R.; and Kaltenbrunner, A. 2015a. A Platform for Visually Exploring the Development of Wikipedia Articles. In *Proc. ICWSM*.
- Borra, E.; Weltevrede, E.; Ciuccarelli, P.; Kaltenbrunner, A.; Laniado, D.; Magni, G.; Mauri, M.; Rogers, R.; and Venturini, T. 2015b. Societal Controversies in Wikipedia Articles. In *Proc. CHI*.
- Dewey, C. 2015. Gamergate, wikipedia and the limits of ‘human knowledge’, <http://goo.gl/T1VhTh>.
- Flöck, F., and Acosta, M. 2014. WikiWho: Precise and efficient attribution of authorship of revisioned content. In *Proc. WWW*.
- Flöck, F., and Acosta, M. 2015. whoVIS: Visualizing editor interactions and dynamics in collaborative writing over time. In *Proc. WWW (to be published)*.
- Hern, A. 2015. Wikipedia votes to ban some editors from gender-related articles, <http://goo.gl/xDxwWI>.
- Kaplan, S. 2014. With gamergate, the video-game industry’s growing pains go viral, <http://goo.gl/F2xtwj>.
- Marcotte, A. 2015. On wikipedia, gamergate refuses to die, <http://goo.gl/n73dQx>.
- Parkin, S. 2014. Zoe quinn’s depression quest, <http://goo.gl/u5Z6GD>.
- Suh, B.; Chi, E. H.; Kittur, A.; and Pendleton, B. A. 2008. Lifting the veil: Improving accountability and social transparency in Wikipedia with Wikidashboard. In *Proc. CHI*.
- Viégas, F. B.; Wattenberg, M.; and Dave, K. 2004. Studying cooperation and conflict between authors with history flow visualizations. In *Proc. CHI*.
- Wingfield, N. 2014. Feminist critics of video games facing threats in ‘gamergate’ campaign, <http://goo.gl/zvw8EZ>.