

# Game Postmortems vs. Developer Reddit AMAs: Computational Analysis of Developer Communication

Chien Lu  
Tampere University

Jaakko Peltonen  
Tampere University

Timo Nummenmaa  
Tampere University

## ABSTRACT

Postmortems and Reddit Ask Me Anything (AMA) threads represent communications of game developers through two different channels about their game development experiences, culture, processes, and practices. We carry out a quantitative text mining based comprehensive analysis of online available postmortems and AMA threads from game developers over multiple years. We find and analyze underlying topics from the postmortems and AMAs as well as their variation among the data sources and over time. The analysis is done based on structural topic modeling, a probabilistic modeling technique for text mining. The extracted topics reveal differing and common interests as well as their evolution of prevalence over time in the two text sources. We have found that postmortems put more emphasis on detail-oriented development aspects as well as technically-oriented game design problems whereas AMAs feature a wider variety of discussion topics that are related to a more general game development process, game-play and game-play experience related game design. The prevalences of the topics also evolve differently over time in postmortems versus AMAs.

## CCS CONCEPTS

• **Information systems** → *Social networking sites; Document topic models*; • **Computing methodologies** → *Topic modeling*; • **Applied computing** → *Computer games*; • **Software and its engineering** → *Software creation and management*.

## KEYWORDS

Reddit, game development, postmortem analysis, text mining, literature analysis

### ACM Reference Format:

Chien Lu, Jaakko Peltonen, and Timo Nummenmaa. 2019. Game Postmortems vs. Developer Reddit AMAs: Computational Analysis of Developer Communication. In *The Fourteenth International Conference on the Foundations of Digital Games (FDG '19)*, August 26–30, 2019, San Luis Obispo, CA, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3337722.3337727>

## 1 INTRODUCTION

Game development has attracted much academic attention, both because of the important economic and cultural role of the fast

growing game industry as well as the unique nature of game development compared to traditional software development [15]. Game development projects can be analyzed from several perspectives from their role as creative and artistic tasks through their software project management to their role as business ventures. Online descriptions of game development communicated by developers provide a big data archive of game development culture and practices and their trends and variation, suitable for qualitative and quantitative analysis. As the amount of records grows, computer-assisted quantitative analyses become ever more necessary to gain a sufficiently comprehensive understanding of development phenomena including weak but rising trends that might not be noticed in a small manually inspected sample. We carry out such an analysis.

This paper analyzes two different types of post-released game developer-generated documents: postmortems published on online websites and AMA (Ask Me Anything) threads on the Reddit social media platform. Postmortems are a common game industry practice where developers reflect on completed game development projects and their problems and challenges that affected the project outcome. Gamasutra.org's instructions [19] state that each postmortem should include some aspects of the project that went right in the project and some that went wrong, and that the aspects should be unique to the project, and that the postmortem should offer concrete thoughts for other developers to learn from. In contrast, AMA threads are a marketing effort where developers attract attention to their studios and games by participating in an interactive crowdsourced interview on the Reddit platform, where questions are posted and upvoted by the audience and developers answer highly-rated questions. AMAs are usually arranged near the launch of a game title, but questions are not pre-constrained to a particular topic and can thus concern details of the latest game, previous games, and other aspects of the developer. In general, the main difference between the Postmortems and AMAs is that the AMAs can be considered as products of the developers interacting with people outside of the game development project whereas the postmortems mainly represent retrospectives from the developer's point of view.

The benefit of analyzing multiple types of developer communication is twofold: firstly, it can provide a more comprehensive picture of issues and focus areas in game development that are not covered sufficiently in an individual source; and secondly, the communication of game developers to different audiences is itself a game development phenomenon worth studying. The time and style in which these communications take place also differs. Postmortems may only be written after the development process, and AMAs may also be conducted during the development process. Our hypothesis is that both the timing and style of the different sources will differ between the communities, so that we will be able to find both shared focus areas and ones strong in one community only, as well as differing temporal trends in these focus areas.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

FDG '19, August 26–30, 2019, San Luis Obispo, CA, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7217-6/19/08...\$15.00

<https://doi.org/10.1145/3337722.3337727>

Postmortems have been used as a source for e.g. researching development practices [16, 17, 22]. However, making use of more information beyond postmortems can be beneficial. In particular, gaming AMAs are where developers interact mainly with players or potential players of their games. As such, AMAs can be considered as a complement of postmortems since they contain more first-hand feedback of the user-experience which might have been neglected during the game development and evaluation process. Thus, it is important to understand differences and commonalities of the two text sources. Previous developer-oriented research works have focused on using designed surveys [10] or interviews [12] and have put less attention on how developers interact with game players.

AMAs are an untapped resource compared to postmortems. This is partly since they are a newer phenomenon than postmortems. However, analysis of a large set of game development related AMAs also takes more effort, as analysing them by simply reading is cumbersome and time consuming. The topic modeling approach in this paper is instrumental in increasing our understanding of AMA content. We demonstrate similarities and dissimilarities between topics in both postmortems and AMAs, revealing information that can be gained from AMAs. One identified reason why postmortems are seldom analysed is the lack of postmortems [16]. Instead of submitting to this issue, we propose AMAs as an important source of data, which can give complementing insights to game development.

Our research questions are: (1) Are AMAs useful for game development analysis? (2) Can the two text sources be used in a large scale analysis? (3) What are the differences in topics discussed between the two sources?

## 2 RELATED WORK

Analysis of game postmortems has been used in research. Petrillo et al. [17] analyzed 20 postmortems to survey problems in the development process of electronic games. Petrillo & Pimenta [16] analysed the same 20 postmortems in another study using the Postmortem Analysis technique, to investigate the adoption of agile methods in game development. The technique is a manual technique, requiring the papers to be read and sentences highlighted. Washburn et al. [22] have analyzed 155 postmortems qualitatively to investigate characteristics of development in order to identify best practices and pitfalls in game development. Tschang [21] analysed 76 game postmortems for use as secondary information in a study about forces that influence creativity in video game development. The analysis was done following the work of Miles & Huberman [13], by coding and classifying using established protocols.

Interestingly, the Reddit AMAs haven't received very much attention in the game research community. AMAs seem to have been used very little in general as a source in research, and e.g. in a study by Bergstrom [3], AMAs are used as a source due to the events that unfold relating to a specific Reddit community member, instead of AMAs being a primary source of information.

## 3 METHOD

### 3.1 Data collection

We collected game postmortems and Reddit AMA threads between the years 2009 and 2018. The starting year 2009 was chosen as the subreddit */r/iAMA* was launched on that year.

Postmortems exist over many online sources from industry focused websites such as Gamasutra and GameCareerGuide to developer conferences such as the GDC conference to blogs of individual developers or companies. Moreover, recently the itch.io platform has launched a postmortem blogging platform focused on indie development and game jams [9]. Out of these several postmortem sources we chose to focus on the industry related websites Gamasutra.com and GameCareerGuide.com. We collected the postmortems from the two above-mentioned websites with a crawler based on Python. In total, 163 postmortems are collected from 2009 to 2018 where 84 of them are from Gamasutra.com and 79 of them are from GameCareerGuide.com. The postmortems are curated by the publishing websites and entirely developer-written, and thus required only technical cleanup to e.g. remove HTML tags.

The AMAs are collected via an open software pushshift API<sup>1</sup>. We first collected the AMAs from the IAmA subreddit with the "Gaming" flair-tag. However, the "Gaming" flair-tag first appeared in 2014; thus, from 2009 to 2014 we collected the AMAs with title containing "game developer". This yielded 668 AMA threads.

To focus on the game developers, we keep only the developer-written content. Within each collected AMA thread, we kept only the posts written by the developer (submitter of the AMA thread), to avoid confusing developers' views with opinions of the audience. As the AMA submitter can choose the question to ask, this filtering enables us to keep the text that reveals more of developers interests.

We noticed that some AMA threads with the "Gaming" flair-tag were not about game development but concerned discussions about esports competitions and other related gaming topics. Therefore, the 668 collected AMAs were further manually inspected to remove AMAs that were not by game developers, or were about something other than game development; ultimately 78 AMAs were removed containing, e.g., AMAs of individual gamers, game streamers on YouTube or Twitch, game store employees, and developers of products related to games but not games themselves such as developers of streaming platforms or gaming devices. After this, 590 game developer AMAs remained that were kept for the analysis.

Finally, the 163 postmortems and 590 AMAs were analyzed with text mining methods described next.

### 3.2 Text Mining

This research employs a text mining technique called topic modeling [1, 2, 5, 18] to analyze the collected texts of postmortems and Reddit AMAs. A topic model is a probabilistic model that represents each document as arising out of several latent topics, which represent semantic themes present over the whole data collection. The topics are not provided by a human but are inductively learned by an algorithm. Topic models are inherently designed to fit well discrete data like text documents, and are better suited to model such data than models like principal component analysis that are designed for continuous vector-valued data. The classical topic model, Latent Dirichlet Allocation has been used in many different research domains including game studies (e.g. [6]).

In a topic model, each topic occurs with different strength in each document, so that some documents may be composed mostly of a particular topic whereas others are a mixture of several topics;

<sup>1</sup><https://github.com/pushshift/api>

we expect that each postmortem and Reddit AMAs will naturally cover several topics, so this representation is natural. The words present in a document then arise out of its chosen topics, as samples of typical words used in each topic. Topic models are fitted to a collection of data, so that the prevalences (strengths) of the topics over documents and the word content of the topics are optimized to represent the available data as well as possible.

Mathematically, a topic model represents the probability to generate a word  $w$  into a document  $d$  as a multi-step process where each document is first generated a multinomial distribution  $\theta_d$  containing probabilities of  $K$  different topics, and each topic  $k = 1, \dots, K$  in turn is generated a multinomial distribution  $\beta_k$  over possible words in the vocabulary. Individual words are then generated into a document by sampling, for each word, a topic  $k$  from the topic distribution  $\theta_d$  and sampling the word  $w$  from the word distribution  $\beta_k$ . The training of a topic model is an optimization process that finds the topic distributions  $\theta_d$  and word distributions  $\beta_k$  which best explain the observed collection of documents.

*Influence of document covariates.* It is expected that interests of developers and the content of their communication will vary, both based on the different postmortem and AMA audiences as well as over time. To analyze such effects, a naive approach would perform separate analyses in different divisions of data and try to relate the results, however, this would suffer both from limited data within each data division and from difficulty of manually identifying trends over multiple groups. Instead, we directly take the audience and time into account in the modeling itself. To do so, in this paper, we go beyond the simple Latent Dirichlet Allocation and choose a more advanced approach, Structural Topic Model (STM, [18]). The STM is an advanced topic model where the strength of topics within a document, and the words that are chosen from each topic, can depend on document-level covariates, that is, annotations that are known about the document. It has been widely used in different research fields such as policy research [8], climate change [7] and travel research [20]. In this work we will use the type of the document (postmortem or AMA) and its publication time as document-level covariates. To avoid overfitting the models, we allow the document-level covariates to affect the topic strengths but not the word probabilities within topics. Modeling the relationships of the topic strengths to the document-level covariates allows us to gain better understanding of how the text source and time influence the developer discussion.

*Choosing the best model.* In STM, user has to assign the number of underlying topics  $K$ , here, we use the held-out likelihood as the criteria decided the number of topics. To do so, a small subset of the collected documents is kept separate ('held out') and is not used to build the topic distributions, and the STM models are evaluated by their likelihood on this held-out subset, representing ability of the models to represent previously unseen documents. From  $K = 5$  to 20, under each setting we built 50 STM models each with different initialization and the  $K$  that yielded the highest average value of held-out likelihood is determined to be the optimal number of topics. The optimal  $K$  turns out to be 16.

After determining the  $K$ , we further build 50 STM models each with different initialization with  $K = 16$  and the one with the best average semantic coherence [14] over topics is selected as the final

model. The average semantic coherence measures how strongly the top words in each topic co-occur over documents, and can be employed to choose the best-performing topic model among several comparable models, as we do here.

When analyzing the results of a topic model, it is typical to list topics, their proportions in the data collection, and top words for each topic. However, listing simply the most frequent words in each topic is not sufficient, since some very common and uninformative words may be shared by several topics. Instead, an alternative scoring metric FREX [4] is employed in this paper. Using FREX balances the frequency of a word within the topic with the specificity of the word to that topic. The FREX score is defined as:

$$FREX_{k,v} = \left( \frac{\omega}{ECDF(\beta_{k,v} / \sum_{j=1}^K \beta_{j,v})} + \frac{1 - \omega}{ECDF(\beta_{k,v})} \right)^{-1}$$

where  $ECDF(\cdot)$  denotes the empirical cumulative distribution function (proportion of items that are smaller than the value in side of the parentheses) and the  $\omega$  is the weight balancing word frequency and exclusivity which is set to 0.5. We use top FREX words within a topic to represent content that particularly characterizes each topic in contrast to other topics.

In the STM model, in addition to the above overall results we are further able to extract and analyze the relationship of the topic proportions to the document-level covariates.

## 4 RESULTS

### 4.1 Overall Data Volume

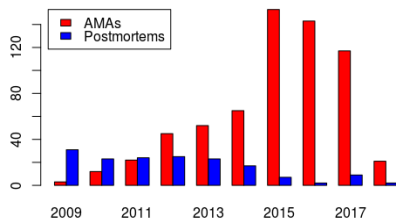
The overall number of published postmortems versus AMAs in our collection over time is shown by year in Figure 1. The amount of published postmortems has started to sharply decrease from around 2013 onwards, whereas the amount of published AMAs has been steadily growing since 2009, with especially sharp growth in 2015, where at the same time the number of postmortems has dropped to below 10. Our postmortem collection arises from two sources representing industry focused websites. Thus, we note that postmortems on such websites seem to be a decreasing trend; this could be due to postmortem publication moving to other sources such as developer conferences, or due to postmortems being replaced in part by AMAs, or a combination of these effects. The potential trend of postmortems being in part replaced by AMAs could be due to several reasons, e.g. because AMAs are available to a wider pool of developers without requiring publication permission from a site such as Gamasutra, or because AMAs offer immediate interaction with fans, or because developers prefer the freeform question answering of AMAs in contrast to more formalized writing of postmortems. The impact of the rise of AMAs, compared to the effect of postmortems moving to other publication venues from the industry websites in our collection, can be a topic of follow-up studies.

### 4.2 Extracted Topics

The overall result of the topic modeling is shown in Table 1, showing the extracted topics, their overall proportions in the postmortems and AMAs together, and the top words in each topic; each topic has been assigned a name based on its top words.

**Table 1: Topic modeling results: topics, their proportions (Pr), and top 10 FREX words**

Topic	Pr (%)	Top 10 FREX Words
Programming	11.34	unity, program, indie, language, java, code, tutorial, learn, pixel, python
Wants and Likes	9.50	thing, say, want, lot, sure, like, still, day, see, tell
Inducing Emotions	8.93	make, feel, seem, way, people, enough, quite, idea, good, right
Music	8.78	cat, song, guitar, music, lego, album, soundtrack, composer, wii, youtube
Team Work	6.39	student, member, morale, copyright, communication, task, igf, asset, project, team
Storyline	6.07	film, bioshock, porn, perception, path, horror, child, woman, fiction, writer
Maintaining Online Communities	6.04	account, ban, dota, community, steam, boost, contact, customer, mod, manager
Mobile Game Development	5.68	ios, app, android, device, iphone, mobile, ipad, store, user, html
Humor and Subjectivity	5.45	haha, funny, guess, damn, joke, terrible, hate, weird, suck, dude
Design of Dangers	5.45	zombie, gun, survival, weapon, pvp, quest, server, behemoth, npcs, planet
Strategy Games	5.03	rts, civ, faction, dragon, realm, paradox, war, city, unit, age
Gameplay Design	4.90	mechanic, experiment, prototype, player, board, monster, avatar, iterate, king, hero
Bugtesting	4.79	percent, xbox, postmortem, magazine, microsoft, external, bug, localization, skull, publisher
Cinematic & Environmental Design	4.68	camera, dialogue, uncharted, trine, enemy, light, deus, environment, sequence, texture
Funding	4.67	kickstarter, fund, backer, salary, manufacture, profit, japanese, pretzel, business, japan
Competitive, Card & Esports Games	2.31	starcraft, deck, pokemon, tournament, card, esports, pro, yea, hearthstone, korean

**Figure 1: Number of collected postmortems and Reddit AMAs per year, after removing nonrelevant AMA threads.**

The extracted topics all correspond to meaningful aspects of game development. The most prominent overall topic is **Programming**, showing that game development is regarded primarily as a technical challenge; the Unity platform and the Java and Python languages are prominent words in the topic, reflecting their prevalent position in current game development. A particular aspect of the programming discussion is providing starting points to audience members interested in game development, shown e.g. in occurrence of words ‘tutorial’ and ‘learn’ in the topic. Other topics related to technical challenges such as **Mobile Game Development** and **Bugtesting** are also extracted. Mobile game development aspects overall took 5.68% proportion of the content; notably, non-mobile systems such as consoles and PC did not arise as their own topic. Besides, ‘store’ as a prominent word reflects the importance of the App Store and Google Play mobile platforms, whereas the non-mobile platform Steam is mentioned in context of **Maintaining Online Communities** rather than as a development concern.

The topic **Wants and Likes** is the second most prevalent, thus people indicating what they like and want in games and their development is common. In the topic **Humor and Subjectivity**, both positive words such as funny and negative ones such as terrible and hate occur, showing that attitudes and preferences to this type of

content vary. The third most prevalent topic, **Inducing Emotions** shows a considerable amount of content of this kind of communication is emotion related. This topic may represent both emotions felt by the game developers themselves during the development process, and emotions that they wish to evoke in the game audiences.

As game design has been always an important issue, related topics such as **Music**, **Storyline**, **Design of Dangers**, **Gameplay Design**, and **Cinematic and Environmental Design** are extracted. The occurrence of cinematic and environmental design may be relevant particularly to high-budget game development, whereas the other mentioned aspects are prevalent both in indie and high-budget games; however, there are also examples in indie games, for example a postmortem for the game “Amnesia: A Machine for Pigs” describes that “The game environment has far greater potential for [creating poignancy] than a spoken or written dialogue describing that environment.” The topic **Design of Dangers** is mainly concerned with weapons (guns in particular) and opponents such as zombies, but also deals with player versus player danger. Gameplay mechanics, prototyping and experiments emerged as their own topic; however, notably, tactics, multiplayer strategies and other team based play did not emerge as a prominent topic, perhaps suggesting that the experience of individual players (even if they are members of a team) is still more prominent in developer discussion.

The topic **Maintaining Online Communities** arises as the 7th most prominent overall; enforcement of conduct through banning players (‘ban’ is the 2nd most prominent word) arises as an important concern for developers, whereas boosting communities, contacts and manager roles also arise although less prominently.

Topics **Team Work** and **Funding** are crucial in the process of game development. Team work is an aspect where game development overlaps with other software development, whereas funding emphasizes crowdsourced funding such as Kickstarter prevalent especially in indie projects. Lastly, the **Strategy Games** and **Competitive, Card and Esports Games** are associated to discussions

where other games are referenced, but also on game mechanics and features. In contrast, several other known game types (e.g. adventure games, puzzle games, role-playing games, sports games, simulator games) did not arise as prominent topics of their own.

### 4.3 Overall Comparison Between Corpuses

The topic proportions in different text sources can be found in Figure 2. Based on the occurrence proportions of the topics in the postmortem and AMA corpuses, we note that both Postmortems and AMAs share some common interests such as **Inducing Emotions** and **Mobile Game Development**. The former may be due to the key role of achieving players' emotional engagement in game design. The latter, in turn, reflects that mobile games now span a wide range of project types from small indie projects to high-budget productions, hence discussion of their development also occurs in multiple venues and formats. However, the sources also differ with respect to their topical content. According to the results, topics **Team Work**, **Bugtesting**, **Cinematic Environmental Design** and **Gameplay Design** are clearly more prevalent in Postmortems. Team work being prominent in postmortems may be due the longer-form descriptions making it easier to mention team roles. The writer of the postmortem, having been a part of a development team and writing the text alone, might feel compelled to talk about the team. The style of AMAs is based more on answering questions, which will affect what is discussed. On the other hand, articles in venues like Gamasutra and GameCareerGuide may also act as promotion of the authoring companies to prospective employees, and describing team work aspects may benefit such promotion. Prominence of bugtesting may also be due to the role of postmortems in describing the overall development process whereas AMAs may not concentrate on early development as much. The higher prominence of cinematic and environmental design and gameplay design in postmortems is somewhat surprising; it may again be partly due to the role of postmortems in describing the overall process. In contrast, topics **Programming**, **Wants and Likes**, **Music**, **Storyline**, **Maintaining Online Communities**, **Humor and Subjectivity**, **Design of Dangers**, **Strategy Games**, **Funding** and **Competitive, Card and Esports Games** are overall found more prevalent in AMAs. The fact that programming discussion is more prevalent in AMAs may be in part due to smaller projects being discussed, and partly due to the earlier mentioned interest by prospective game developers being part of the audience. Wants and likes being more prominent in AMAs may be due to the interactive discussion with the audience bringing out more expressions of preference.

### 4.4 Temporal Trends

The topic proportion in different text sources over time are shown in Figure 3. The prevalence of the different topics varies strongly over time, and varies differently in the different sources.

Among the AMAs, the prevalence of topics **Programming** and **Mobile Game Development** are decreasing over time; decrease of Programming discussion might be because of maturing toolsets and their existing online tutorials, whereas decrease of Mobile Game Development discussion might be because in the highly competitive mobile game development market it has become harder to

attract attention to an individual developer's AMA. In contrast, the prevalence of the topic **Music** has been slightly growing.

For postmortems, the topic **Team Work** has been a dominant topic in terms of prevalence, and has a rising trend, perhaps due to growing budgets [11] and needed team sizes. Topics **Bugtesting** and **Cinematic and Environmental Design** have had decreasing prevalence. The former may be due to maturing technology.

Other topics such as **Wants and Likes**, **Inducing Emotions**, **Storyline**, **Maintaining Online Communities**, **Humor and Subjectivity**, **Design of Dangers**, **Strategy Games**, **Gameplay Design**, **Funding** and **Competitive, Card and Esports Game** do not display obvious changes of prevalence over time.

## 5 DISCUSSION

One major finding from the analysis is that the content of the developer communications depends on who they interact with. Take technical related topics for instance, shown in Figure 2: both text sources share the same interest in **Mobile Game Development**, but reddit users are more interested in relatively general concepts (topic **Programming**), whereas in postmortems, more emphasis is put on detailed-oriented discussions (topic **Bugtesting**). Another fact of the technical related topics is that, as shown in Figure 3, they have been less and less prominent over time. The prevalence of **Programming** and the topic **Mobile Game Development** are found decreasing and slightly decreasing in AMAs and a similar trend of **Bugtesting** is found in postmortems.

Concerning game design, AMAs focus more on content creation or player experience such as **Music**, **Storyline** and **Design of Dangers** whereas in postmortems more technical-oriented topics, **Gameplay Design** and **Cinematic and Environmental Design** are more frequently discussed. Besides, game-play related topics such as **Maintaining Online Communities**, **Strategy Games** and **Competitive, Card and Esports Games** are shown more prevalent in AMAs instead of postmortems. The proportion of the **Programming** topic is higher in AMAs compared to postmortems, while the opposite is the case for **Gameplay Design**. This may point to AMAs as a timely venue for learning and discussing about software development aspects of game creation, while postmortems allow developers to explore design decisions thoroughly.

Another phenomenon is prominence of **Team Work** in postmortems. This may be due to increased development cost especially in large-budget games [11] which need large teams and hence lead to more team effort related discussion in developer communications. In contrast, it has not drawn as much attention in AMAs.

Some of these differences can be due to how AMAs and postmortems differ from each other in style and content. AMAs are driven by the questions of community members, while postmortems are written by a developer without interaction with the community. The point in the game's life-cycle at which these communications take place is also different. This means that the content of postmortems is determined by what the developer feels is important to describe about the development process after development has ended, while AMAs are created in an interactive fashion, without such a limitation on time. The way AMAs can interactively respond to, and raise consumer interest in, game development aspects merits further research.

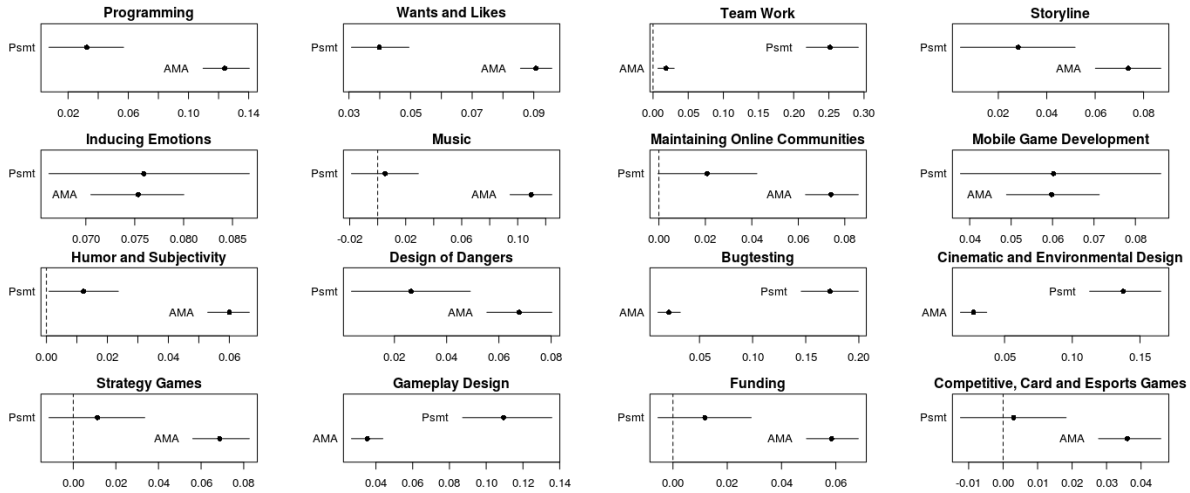


Figure 2: Effect of Source on Topic Proportions

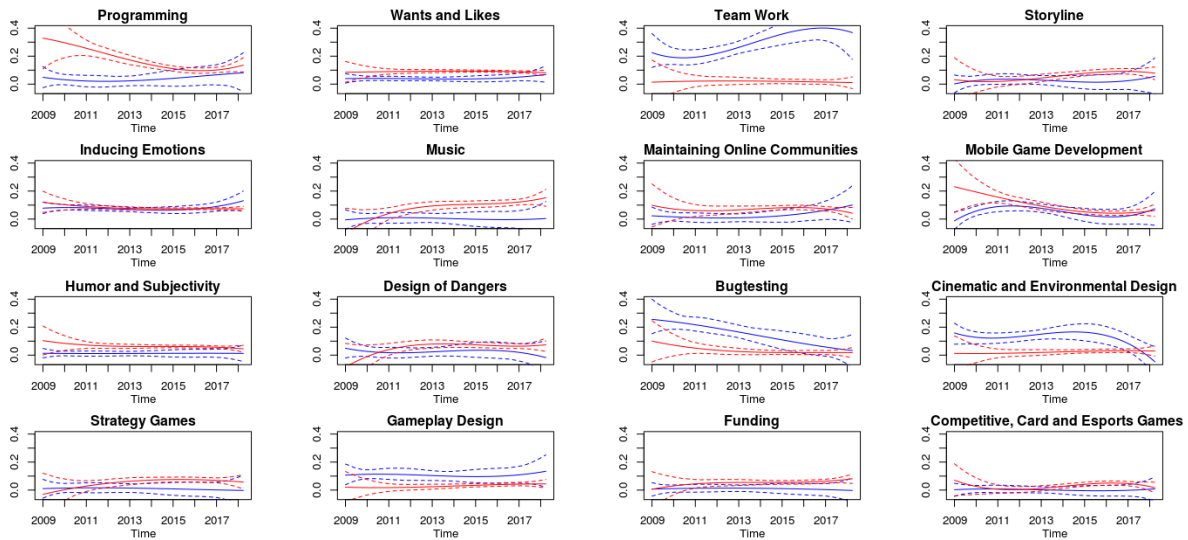


Figure 3: Topic Proportions Over Time, blue line represents the topic proportion among Postmortems and red line represents the topic proportion among AMAs. Dotted lines represent the 95% confidence interval.

## 6 CONCLUSION

We presented a quantitative analysis of game developer communication in two venues, postmortem articles and Reddit AMAs based on probabilistic topic modeling of a large set of postmortems and AMAs. We find that (1) AMAs are useful for game development analysis: they yielded a large amount of data which turned out to contain game development relevant discussion worth analyzing; (2) AMAs and postmortems are compatible for large scale analysis: we were able use both sources quantitatively to infer topics of game development, without manual extraction of compatible parts; (3) both similar interests and different topics were extracted from the sources, revealing prominent topics of interest to developers, different prevalence of the topics in the two developer communication types overall, and different temporal trends of the topics in

postmortems and AMAs. As AMAs proved beneficial, we suggest that, other than postmortems, AMAs are also worth monitoring for better understanding of current issues in game development.

The analysis was carried out in a quantitative fashion employing STM, a topic modeling based text mining technique. The quantitative results were both interesting on their own and supported human interpretation and efficient reading of the documents. We expect that as game development data sets continue to grow, support from such computational methods will become ever more crucial.

## ACKNOWLEDGMENTS

The work was supported by Academy of Finland decisions 312395 and 313748.

## REFERENCES

- [1] Md Hijbul Alam, Jaakko Peltonen, Jyrki Nummenmaa, and Kalervo Järvelin. 2018. Author Tree-structured Hierarchical Dirichlet Process. In *Proceedings of DS 2018, the 21st International Conference on Discovery Science*. Springer, 311–327.
- [2] Md Hijbul Alam, Jaakko Peltonen, Jyrki Nummenmaa, and Kalervo Järvelin. 2019. Tree-structured Hierarchical Dirichlet Process. In *Distributed Computing and Artificial Intelligence, Special Sessions, 15th International Conference (Proceedings of DCAI 2018)*. Springer, 291–299.
- [3] Kelly Bergstrom. 2011. "Don't feed the troll": Shutting down debate about community expectations on Reddit.com. *First Monday* 16, 8 (2011). <https://doi.org/10.5210/fm.v16i8.3498>
- [4] J. M. Bischof and E. M. Airoldi. 2012. Summarizing topical content with word frequency and exclusivity. In *Proceedings of the 29th International Conference on Machine Learning*.
- [5] Ali Faisal, Jussi Gillberg, Gayle Leen, and Jaakko Peltonen. 2013. Transfer Learning using a Nonparametric Sparse Topic Model. *Neurocomputing* 112 (2013), 124–137.
- [6] Ali Faisal and Mirva Peltoniemi. 2018. Establishing video game genres using data-driven modeling and product databases. *Games and Culture* 13, 1 (2018), 20–43.
- [7] Justin Farrell. 2016. Corporate funding and ideological polarization about climate change. *Proceedings of the National Academy of Sciences* 113, 1 (2016), 92–97.
- [8] Fabrizio Gilardi, Charles R Shipan, and Bruno Wüest. 2015. The Diffusion of Policy Perceptions: Evidence from a Structural Topic Model. *University of Zurich and University of Michigan* (2015).
- [9] Spencer Hayes. 2018. Post mortems are here on itch.io! Write them, read them, and support developers! <https://itch.io/blog/54291/post-mortems-are-here-on-itchio-write-them-read-them-and-support-developers>
- [10] Jussi Kasurinen, Maria Palacin-Silva, and Erno Vanhala. 2017. What concerns game developers?: A study on game development processes, sustainability and metrics. In *Proceedings of the 8th Workshop on Emerging Trends in Software Metrics*. IEEE Press, 15–21.
- [11] Raph Koster. 2018. The cost of games. [https://www.gamasutra.com/blogs/RaphKoster/20180117/313211/The\\_cost\\_of\\_games.php](https://www.gamasutra.com/blogs/RaphKoster/20180117/313211/The_cost_of_games.php).
- [12] Annakaisa Kultima. 2015. Developers' perspectives on iteration in game development. In *Proceedings of the 19th International academic Mindtrek conference*. ACM, 26–32.
- [13] M. B. Miles and A. M. Huberman. 1994. *Qualitative data analysis: An expanded sourcebook (2nd ed.)*. Thousand Oaks : Sage Publications.
- [14] David Mimno, Hanna M Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, 262–272.
- [15] Emerson Murphy-Hill, Thomas Zimmermann, and Nachiappan Nagappan. 2014. Cowboys, ankle sprains, and keepers of quality: How is video game development different from software development?. In *Proceedings of the 36th International Conference on Software Engineering*. ACM, 1–11.
- [16] Fabio Petrillo and Marcelo Pimenta. 2010. Is agility out there? Agile practices in game development. In *SIGDOC '10: Proceedings of the 28th ACM International Conference on Design of Communication*. <https://doi.org/10.1145/1878450.1878453>
- [17] Fábio Petrillo, Marcelo Pimenta, Francisco Trindade, and Carlos Dietrich. 2009. What went wrong? A survey of problems in game development. *Computers in Entertainment (CIE)* 7, 1 (2009), 13.
- [18] Margaret E Roberts, Brandon M Stewart, and Edoardo M Airoldi. 2016. A model of text for experimentation in the social sciences. *J. Amer. Statist. Assoc.* 111, 515 (2016), 988–1003.
- [19] Ara Shirinian. 2011. Dissecting The Postmortem: Lessons Learned From Two Years Of Game Development Self-Reportage. [https://www.gamasutra.com/view/feature/134679/dissecting\\_the\\_post-mortem\\_lessons\\_php](https://www.gamasutra.com/view/feature/134679/dissecting_the_post-mortem_lessons_php)
- [20] Panagiotis Stamolampros, Nikolaos Korfiatis, Panos Kourouthanassis, and Efthymia Symitsi. 2018. Flying to Quality: Cultural Influences on Online Reviews. *Journal of Travel Research* (2018), 0047287518764345.
- [21] F. Ted Tschang. 2007. Balancing the Tensions Between Rationalization and Creativity in the Video Games Industry. *Organization Science* 18, 6 (Nov. 2007), 989–1005. <https://doi.org/10.1287/orsc.1070.0299>
- [22] Michael Washburn Jr, Pavithra Sathiyarayanan, Meiyappan Nagappan, Thomas Zimmermann, and Christian Bird. 2016. What went right and what went wrong: an analysis of 155 postmortems from game development. In *Proceedings of the 38th International Conference on Software Engineering Companion*. ACM, 280–289.