

Spring 5-2015

# Ask us Anything for Reddit AMA

Chris Arrowood  
*Carnegie Mellon University*

Follow this and additional works at: <http://repository.cmu.edu/theses>

---

## Recommended Citation

Arrowood, Chris, "Ask us Anything for Reddit AMA" (2015). *Theses*. Paper 81.

This Thesis is brought to you for free and open access by the Theses and Dissertations at Research Showcase @ CMU. It has been accepted for inclusion in Theses by an authorized administrator of Research Showcase @ CMU. For more information, please contact [research-showcase@andrew.cmu.edu](mailto:research-showcase@andrew.cmu.edu).

# ask us anything

## for reddit ama

Investigating the role online social networks play in the development and sharing of personal stories and experience.

Chris Arrowood  
Advised by Aisling Kelliher  
Thesis, Master of Design, Interaction Design  
Carnegie Mellon University  
School of Design

---

## Table of Contents

Abstract	4
Introduction	6
Following the Journey	9
Telling of the Journey	11
Past Work on Digital Story & Identity	16
Methods of Digital Story Telling	22
Role of Digital Story	25
Towards a Design Process	28
Data and Identity	29
Digital Social Systems and Online Knowledge Sharing Platforms	31
Principles of Data Visualization	38
First Research Structure	41
First Research Findings	44
Designing for an Improved AMA Experience	46
First Version Application Structure	47
First User Testing & Findings	50
Second Version Improvements	52
Second Version Application Structure	58
Second User Testing	59
Second User Testing Findings	62
Implications for Online Knowledge Sharing Platforms	65
Implications for Broader Online Social Networks	68
Future Work	70
Discussion	72
Bibliography	73
Appendix	77

---

## Abstract

In an eminently connected, networked and anonymous digital world, the individual sense of identity becomes a composition of personal and increasingly public layers of influences, likes, follows, and shares. To the digital native, this multi-layer composition is, in some cases, complimentary, and in others, entirely opposing manifestations of interests, tastes, preferences and experiences, canvased across multiple systems, apps and social networks.

With an interest in enabling a digital archaeology of this information – exploring and understanding it in a more complete sense – this thesis will explore both the theoretical and philosophical basis on which modern identity is considered in the context of user experience design and online social networks.

With a theoretical basis in storytelling, curation, and myth, building towards a technical and research basis in information assessability, natural language processing and data visualization, it will explore the notion of

“personal digital identity” – both in the form of personal experience and personal story – under the context of online social network platforms that are prevalent today.

Finally, it will propose and demonstrate a fully functional tool for the above mentioned “archaeology” of personal digital identity, within the context of a specific online social network – Reddit.com “Ask Me Anything” interview.

As a design research project, the focus of the project will be on realizing the full potential for online knowledge sharing platforms in helping shape, develop, and share personal stories and experiences.

---

## Introduction

“A hero ventures forth from the world of common day into a region of supernatural wonder: fabulous forces are there encountered and a decisive victory is won: the hero comes back from this mysterious adventure with the power to bestow boons on his fellow man.” (Campbell, 1949)

The development of an individual identity in both the digital and the physical world is a process one does not ever truly abandon. It is an inherently introspective process of developing values and beliefs on which to structure your worldview – your internal identity, which evolves into a framework by which the world perceives, judges, and engages you – your external identity.

In today's digitally saturated society, internet-based platforms of communications, sharing, and collaboration have become an important component in many individual's daily lives. Such systems begin to shape the communication and practices of participating individuals, thus helping to give shape to a kind of identity, one

manifested through these platforms and presented to the outside world.

Such systems enable a curation and crafting of an individual's identity story and message in a way past generations did not have. Along with the ability to publish in such a manner, there is often times opportunity for acknowledgement and feedback on one's publications. With much of your personal life potentially being open to the broader audience of your social network circles, the implications for such systems in influencing your personal self-image and identity are numerous.

This accountability to the world is an aspect of the socially connected age in which the prospect of individual identity becomes a compelling area of research. As the plurality of social systems seeps into aspects of daily life previously thought unremarkable and uninteresting, individuals open themselves up to critique and accountability, interest and inquiry for actions and decisions in ways unimagined.

Here we look to explore the heuristics by which social systems have influenced and enabled such a development of individual identity. Our focus is in exploring the implications of use of such systems, the restrictions and bounding participation binds you to, and how design with such implications in mind might better improve the

experience of using such systems.

Furthermore, as we explore one specific type of online social platform, our aspiration is to identify aspects of design for the improved communication and reception of knowledge and information online. This serves to benefit many prospective designers and researchers, who look to build on the base of knowledge and content already existing on the social web, designing systems that allow further exploration and understanding of socially generated content – our digital stories.

---

## Following the Journey

“In his life-form the individual is necessarily only a fraction and distortion of the total image of man.”  
(Campbell, 1949)

Joseph Campbell’s account of the Hero’s Journey serves as the first cornerstone of this paper’s attempt to structure an exploration of identity in the modern, digital world. The Hero’s Journey, in Campbell’s understanding, is a structured and shared framework which can be transposed, post hoc, onto nearly all of the common mythologies and epic stories of man. Mythology, the stories passed down from times and cultures long past, are a form of storytelling which are regarded in many ways by many different audiences – religious individuals and scholars may view myth very differently.

Considered by some “as a production of poetical fantasy from prehistoric times, misunderstood by succeeding ages (Muller)” or, more optimistically “as the profoundest metaphysical insights (Coomaraswamy)” (Campbell, 1949)(p.330), it is Campbell who harmonizes these

competing views, describing mythology as “all of these” (Campbell, 1949) (p.330).

Respective of its parts, mythology is no single individual’s story to tell. It is a multivariate, layered and, most importantly, representative understanding of human life. Campbell talks of humanity as a living body - its individual components, it’s cells, being the human. From this human body, no one cell can remove itself, for it relies on that body to live. In the formalities and ceremony of life – the weddings, the funerals, the births and the birthdays – we see a natural cadence to life. From understanding this cadence one sees a broader form, and in this form, we see the commonalities of life and the human spirit.

Campbell describes this form as the “super individual” – the representative body for all man kind. It is through the exploration into the space of the myth and the legend that we see what makes our super story so common, and our individual journeys through it so unique. As Campbell, quite poetically describes: “By an enlargement of vision to embrace this super-individual, each discovers himself enhanced, enriched, supported, and magnified. His role, however unimpressive, is seen to be intrinsic to the beautiful festival-image of man – the image, potential yet necessarily inhibited, within himself.” (Campbell, 1949) (p.330)

---

## Telling of the Journey

Throughout time, stories of immense scale and meaning have relied on the works of artists, craftsman, authors and witnesses to pass along the great tale of the journey undertaken by man. Campbell provides us with many great examples, including that of Homer’s Odyssey.



Fig.1: The Shield of Achilles.

As Campbell describes, it is Homer who, failing the language to comprehend and convey the staggering size of the Trojan armies landing at the beachhead before him, resorts to a process of simplification through hierarchy to tell of the army's size. By listing the names of only the highest rank Captains, each with leagues of soldiers behind them, the size of the force can be left to the reader's imagination.

Such list making is exercised in myth, art, performance, authoring, and indeed nearly all practices through which an idea is to be communicated or shared. Eco takes on a census of list making in cultures past and present. Using an example of the shield of Achilles (FIG.1), Eco values what represents to him the perfection of form and story, as one. (Eco, 2009)

The shield, being perfectly round, contained, and certain, becomes the perfect vessel for visual storytelling. On Achilles' shield, the epic tale forged into it is artfully consolidated and abstracted only to its most essential elements. It is the finitude and limit of the human hand which Eco uses to demonstrate the bounds of storytelling, for it is not the medium or the story that runs out, as the shield could contain as much detail as could be fit on it, at such minute scale so as to nearly capture, theoretically, the entire text of Homer many times over. It is instead the human hand's ability to work at any

smaller or more exact a scale, and further, the human eye for being able only to read and interpret what is large enough to be read. What Achilles' shield represents is not one story, but a life's – if not a whole world's journey – captured, contained, and communicated within a single artifact.

Where the shield falls short, in Eco's eye, is in the lack of representation for the "etcetera". When communicating stories or concepts beyond the scale of one object or form, the artist must find forms that enable such an "etcetera". Moving past the physical, it is in the image, and by extension, written language, that one finds more true presentation and acknowledgment of this information too broad, too challenging or, just simply too unknown to communicate.

The frame of a painting, for example, indicates only just that – a frame, or window into the world the artist is portraying. When considering the written or spoken word, specifically written lists, Eco reintroduces Homer's "topos of ineffability". "Faced with something that is immensely large, or unknown, of which we still do not know enough or of which we shall never know, the author proposes a list as a specimen, example or indication, leaving the reader to imagine the rest." (Eco, 2009)

Behind Campbell's heroic journey, Homer's attempts to

describe the immensity common in such stories, and in Eco's exploration of such attempts, we see the challenging of reality that Campbell himself was aware of all along: "There is no final system for the interpretation of myths, and there will never be any such thing". (Campbell, 1949) (p.329) Recalling the wisdom of Campbell prior, one can understand myth as both the shared and timeless stories of eons, as well as the personal and introspective stories of the daily here-and-now. With this extension, one must ask what indeed is the outlook for any such attempt to understand and interpret a personal mythology, such as the journey towards self-identity.

ing, one must only understand the composition of the form to see the entirety of the story. In other words, one only needs a tool to map or translate the form into story and value to the viewer.

Even the most complex storytelling artifacts can be interpreted and understood with such aid, such as his visual map of the shield of Achilles (FIG.2). Our question now is, what is an equivalent aid to interpret or map our modern, digital stories and identities?

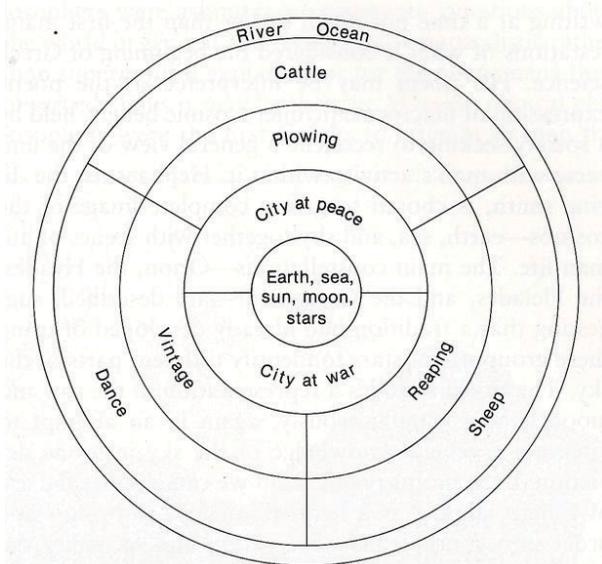


Fig.2: Interpretive diagram of the Shield of Achilles.

It is Eco's argument that with complex forms of storytell-

---

## Past Work on Digital Story & Identity

Viegas (Viegas, boyd, Nguyen, Potter, & Donath, 2004), Feinberg (Feinberg, 2013) and Harris (Kamvar & Harris, 2011) each delve into digital identity in unique ways. Viegas, et al. along with Boyd begin with research into the personal correspondence that is electronic mail – looking quantitatively at the contacts, relationships, and associations present in such an archive of digitally enabled conversations. Feinberg explores the notion of curation within the modern digital world, and asks why such practices exist, who participates in this digital collecting and sharing, and what are the methods to doing so. Finally Harris (Kamvar & Harris, 2011) turns the medium outward, exploring where the digital system acts as tool and enabler to tell stories that would otherwise go untold.

What intrigues researchers about the digital world is its nature as a linked, shared, iterative space. Where Homer found Trojans, the modern world sees data points; for Homer's mythical Sirens, there exist open APIs. The connectivity of activity and content in the digital world

affords understanding and establishment of relations unlike content of any other period. The challenge lies in the rapid pace at which this information is created, meaning the design of any system to understand such information must account for both its enormity and its ever-changing nature.



Fig.3: (Viegas, boyd, Nguyen, Potter, & Donath, 2004) *PostHistory* and *Social Network Fragments*.

Viegas, et al. constructed digital platforms which allowed a visualization of user email history (FIG.3), one of the faster accruing collections of content one individual will

develop. In developing these platforms, the designer's aim was to identify relations between individual users and groups, activities, and topics; the occurrences and frequency of communication between these groups and the user; inferences on the roles and associations these communicating members play within the users world; etcetera.

Through their platforms, Viegas, et al. show that one's archive of digital correspondence can serve as an engaging exploration of personal past, with opportunity to provide a more thorough and meaningful understanding of both one's sense of self and the meaning of their relationships, as viewed through their digital activity.

Work of Feinberg (Feinberg, 2013) on digital collections and curation compliment and validate the type of explorative design prototyping Viegas and others have undertaken. With an interest in defining the role curation plays in communication in the digital age, Feinberg arrived at a conclusion that the impact and efficacy of curated content, in communicating a theme or opinion, is vastly determined by the manner in which it's information is written and presented. Aspects central to the object-as-story – such as descriptions, placement, ordering, relation and theme – each impact the efficacy of the final artifact when considering the list as either a personal, introspective thing, or as an artifact and discus-

sion point for others to view.

In the context of designing for interaction with such information, Feinberg serves to help orient what future work on identity, through the curation of storytelling and representative digital artifact, might become. Might such platforms serve as reflective, self-edifying experiences for individual users themselves, or as outwardly communicative and expressive means of developing a social-facing identity?

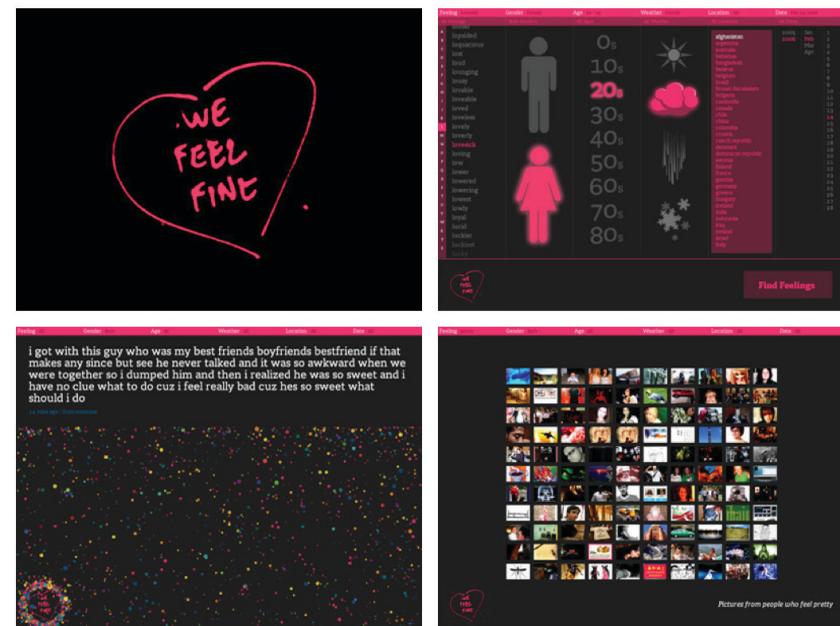


Fig.4: (Harris, 2011) We Feel Fine.

Harris (Kamvar & Harris, 2011)(FIG.4) approaches work

on individuality in the digital age with an interest in the outwardly facing, shared human experience. This is not work pertaining to personal email, but rather of publicly shared expressions of opinion (Kamvar & Harris, 2011), or of experiences and stories shared with the whole (Harris, 2011). Harris, more so than Feinberg or Viegas, interrogates what the notion of private or personal story truly is. Cowbird (Harris, 2011), a custom-designed platform, provides a medium for users to share short accounts of moments, days, or even years of their lives, with other users. From this sharing generates any number of discussions, similar stories, and inspiration.

There is not one true problem or brief Harris is looking to solve for his users. It is only the users essence, their journeys as humans, that brings them together and engages them as a community, within itself. Cowbird itself represents a kind of digital ceremony, such as Campbell defines for us, where individual members contribute to something that takes on an importance and value greater than their individual parts.

Here we will mention some areas of similar work our audience may be familiar with, but that we will not focus on in this project. We will not be bringing a “big data” approach to our theme. Our aim is to look at the more personalized and qualitative forms our digital stories take, through modern social networks.

Similarly, we will not be applying a “quantified self” approach of harnessing all the data points generated by an individual. We look towards digital social systems as the modern platforms on which individual storytelling takes place. This storytelling eschews much of the data a user might generate, and instead represents a heavily curated message chain that would run counter to the quantitative nature of the “quantified self”.

---

## Methods of Digital Storytelling

Previous work mentioned to this point has considered personal data, existing on other networks, such as email and social network posts. Harris takes a different approach with Cowbird (Harris, 2011)(FIG.5), which is designed and built as a platform for authoring stories and experience from the lived world. The platforms express purpose was to publish and spread this content in and amongst its community of users.

As an open-ended and free system, the community defines for itself what the platform means and represents, to the community itself, and to the world outwardly. Other networks similarly self-define, whether through content curation and themes, or through user personality/experience types, to differentiate themselves from other communities which would otherwise seem similar or related.

Questions we might ask of these networks and their communities and systems could include: What are the practical uses of such a system? What types of users

are inclined to use such a system? How does the system grow and expand, or does it become a silo, an independent entity, one of more selective curation of content and members?

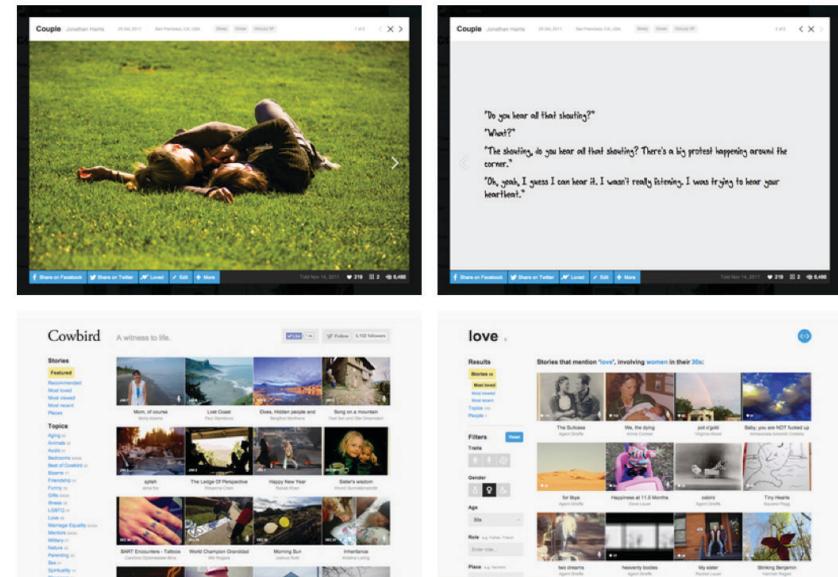


Fig.5: (Harris, 2011) Cowbird.

Heer & Boyd, introduced above, explore the social domain from the opposite perspective of Harris. Their interest is in “better facilitating the discovery of people, connections and communities to promote increased awareness of community structure and information exposure...” (Heer & boyd, 2005). The authors set out to visualize networks of individuals on the already existing social network Friendster. Their work serves as a design

case study in the heuristics and requirements such a mapping and visualization system offer and require of both its authors and of its users.

Referring back to Campbell and his analogy of Achilles' shield, the shield of today might be, or be made with the use of tools and libraries for data visualization and data analysis that we now have access to.

Whether bespoke and proprietary, such as tools built by Viegas and Harris, or open-sourced community maintained tools such as D3 (Bostock, 2011) and Processing (Fry, 2001), these tools allow for the experimentation and composition of stories that have been left untold, in our own data, for so long.

---

## Role of Digital Story

“We are insecure in our understanding of ourselves, and this insecurity breeds a new preoccupation with the question of who we are. [...] The computer is a new mirror...” (Turkle, 1984)

Turkle has written extensively on the role digital technology plays in the lives of widely different user groups, now and in the future. In 1984 she published (Turkle, 1984) about technology and its capacity to fill a void within, or to provide a structure for an individual's life, world, and identity. Within this, she describes the computer as a machine “that lets you see yourself differently, as in control, as ‘smart enough to do science,’ as a more-full participant in the future.” The sense Turkle has for the role the computer has as an instrument to participate in the future is even more significant today than it was at the time of her writing.

Turkle interviews individual users who have developed an affinity for computer programming, and have flourished as individuals because of such applications of

technology. What these users are doing is not always a professional endeavor, as most are in fact students. The machine, given its bounds and logic, its natural constraints, serves to instill the sense of structure these individuals are unknowingly seeking as still-developing, independent young adults.

“With adolescence, there is a return to reflection, but this time reflection is insistently about the self. The questions of the first stage, What is this machine?, and of the second, What can I do with it?, give way to Who am I?” (Turkle, 1984)

Turkle has introduced a third, or perhaps a hybrid approach to understandings of technology and its role in the development of the individual as the two mentioned prior in this paper. To her, technology serves a more macro purpose to some individuals, as a filament on which life and individuality is lit and tested. For these users, technology provides a missing piece to the sense of their whole as individuals. With this new sense of wholeness, they are able to embark on life’s journey with better sense of individual north.

Following this thread Turkle provides into the role of technology as a means of personal development, boyd picks up by exploring the development of youth growing up with social networks as a primary aspect of their lives

and social culture. In (boyd, 2014) develops a perspective on the dangers such systems currently present, as static and permanent forms of publishing provided to a user at a tenuous and naturally iterative phase of life – youth using social media are still defining for themselves their sense of identity, individuality and place in their world, yet communicating and sharing such a journey in, often times, a very public, and almost always, a very permanent platform.

These “arenas for the formation and enactment of social identities” (boyd, 2014) (p.20) present a challenge for both the users (often times unaware of the implications), the authors of such platforms, and of the researchers of such systems. Indeed it is not only youth who may lack the appreciation for the hyper public nature of these platforms, but any potential user of the platform.

---

## Towards a Design Process

This thesis has explored research into various fields of storytelling, curation, and the development of an individual sense of identity. This research of prior work is only one particular and distinct path through the explored domains, hoping to draw distinction between the merits and applicability of each when considering the evolving digital sense of self identity.

While technology continues to evolve and embed itself within our lives in increasingly complex manners, the challenge of understanding this technology's role and impact is ongoing. While one might imagine Campbell would feel it impossible to ever build such a tool to fully understand such complexity and such a grand story, his own words give some value to even a partial understanding of that story: "The aim is not to see, but to realize that one is, that essence; then one is free to wander as that essence in the world".

---

## Data and Identity

The issues facing users of today's digital systems are of a personal variation to many of the issues faced by the creators and maintainers of the systems on which these users find themselves. The age of "big data" has opened to users, institutions, researchers, marketers, and even the ill intentioned, the ability to see at a meta scale the activity and actions of millions of users.

For individuals, the notion of the "quantified self" has developed as a heuristic and approach to understanding and tracking one's individual data. Meanwhile, campaigns for information and data security literacy by many social platforms have hoped to stymie the fears of "hacking" and identity theft that has plagued the digital era.

It is not the express interest of this project to tackle "big data" nor the "quantified self" – but rather, to explore ways in which the timeless models for identity, storytelling, and curation have shaped the modern individual user of digital systems. Within this context, questions

of the design of such systems naturally give way to an interest in the data acquired by and used to inform such systems. Where did the data come from? What does it say about the user/author? How has the platform on which it was created informed or influenced the generated data?

We will consider how best to appropriately engage with this data, with specific interests in helping enable a process of producing, viewing, sharing or otherwise developing an individual digital identity. In doing so, we may eschew completeness in the interest of comprehensibility (as did Homer in describing the Trojan army). We will look for the qualitative within the naturally quantitative world of “big data”. We look towards digital social systems as the modern platforms on which storytelling and the publication of our modern mythologies takes place.

---

## Digital Social Systems and Online Knowledge Sharing Platforms

Online knowledge sharing platforms and communities are a popular and reliable source of information on the social web. Traditionally, online knowledge sharing platforms take the form of either encyclopedic data (Kittur & Kraut, Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination, 2008), or conversational question and answer threads (Wang, Gill, Mohanlal, Zheng, & Zhao, 2013). Other platforms, such as Yahoo! Answers or Patient’s Like Me, enable a more personal and conversational dialogue. As these platforms become more prevalent, the ability of these systems to enable not only participation and high quality discussion (Kittur & Kraut, Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination, 2008), but also higher-level analysis and comprehension (Wang, Gill, Mohanlal, Zheng, & Zhao, 2013) – “information assessability” (Forte, Andalibi, Park, & Willever-Farr, 2014) – is of critical importance.

Over time, new models of knowledge sharing have

developed (Wang, Gill, Mohanlal, Zheng, & Zhao, 2013), including the community which we study and design for here: the increasingly popular “AMA” or “Ask Me Anything” discussion. Ask Me Anything is a crowd-sourced interview popularized by the social website Reddit (Weninger, Zhu, & Han, 2013). In an AMA, an individual (typically of an interesting background or qualification) makes themselves available to the community for an interview, through comment submission, using the websites discussion thread interface.

The AMA represents a form of modern storytelling which embodies aspects of the social web as well as aspects of storytelling as we have explored in this thesis. The social nature of making a request to the community for another user to make themselves available leverages the size complexity of the reddit social network. The model of crowd-sourced questions further embodies the unique capability of digital systems.

Meanwhile, the individual and self-directed nature of the AMA author’s role gives them agency over the discussion, allowing for a very personal story to be told. The final artifact of the AMA discussion thread, as we will look at in-depth later, presents itself as just the dense and content-rich entity in need of interpretive aid as Eco saw in the shield of Achilles. It is with these compelling aspects of the AMA in mind that we have chosen it as

the platform to design around.

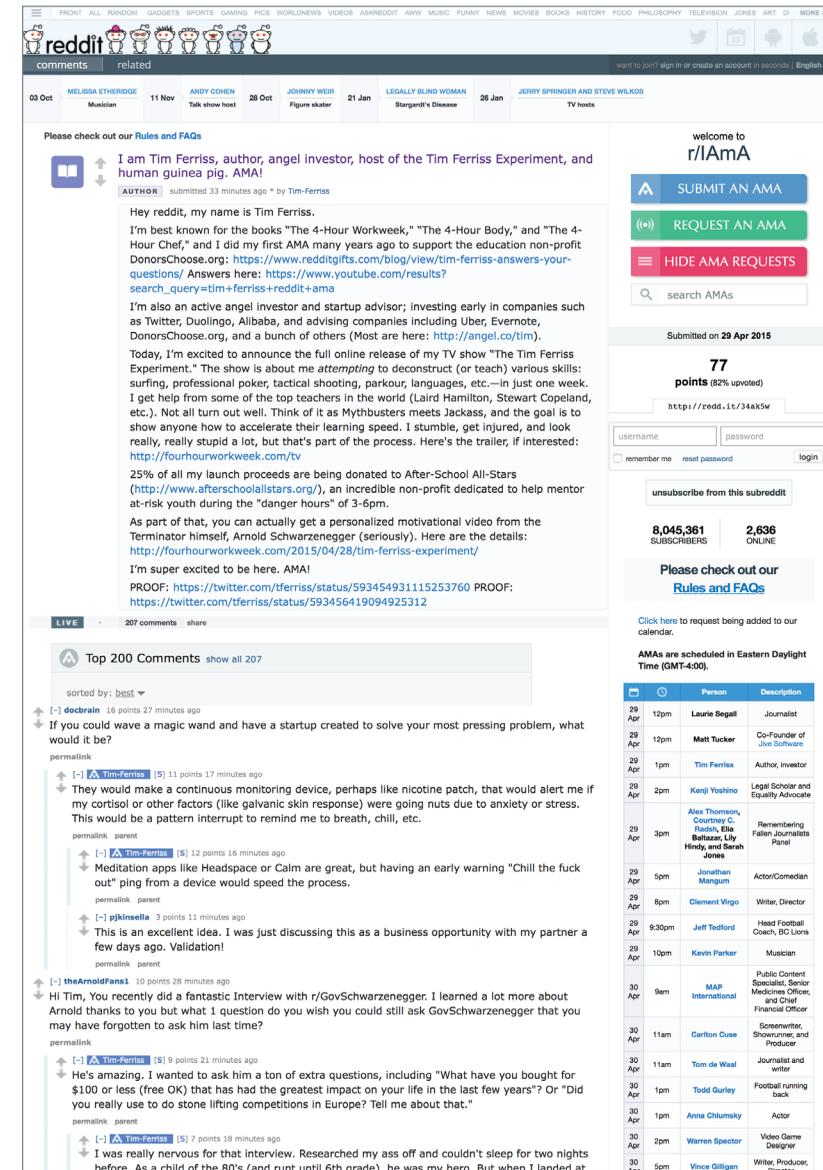


Fig.6: An example AMA thread, from author Tim Ferriss. (Ferriss, 2015)

The AMA interview format capitalizes on the voting functionality of Reddit's comment submission system to allow the community and the author of the AMA to determine the questions of most interest and value. AMA interviews can extend in duration from an hour to as long as a few days. Additionally, authors can also receive direct messages outside of the AMA, which can continue the dialogue well beyond the original post.

The AMA is of interest to our research for a number of reasons. This format is an example of the ways modern platforms have reshaped how we think of storytelling and of our personal experiences within the context of others. It underscores the unique nature and importance of the individual and their experience, while also illustrating a universal desire for this kind of knowledge, evidenced by the request/demand driven origin of the AMA interview (users asking for someone specifically to interview with them). AMA is a community of users with their own interests and experiences, who are fascinated with the interests and experiences of others.

Individuals participate in AMA interviews for a variety of reasons such as self-motivation; response to an "AMA Request" seeking someone with their experience; or by personal request by a friend or other user. Within AMAs, we identify three common user roles adopted by members of the community.

AMA authors, who make themselves available to the community for discussion. AMA commenters, who either ask questions of the author, or contribute other comments to the discussion. Finally, AMA requesters are individuals that make posts seeking out another person or type of individual to conduct an AMA.

While AMAs have grown in popularity in recent years, the format and presentation approach of the discussion can present challenges. A key issue is that the AMA interview relies on Reddit's existing comment system, which was not originally designed for interview style dialogue.

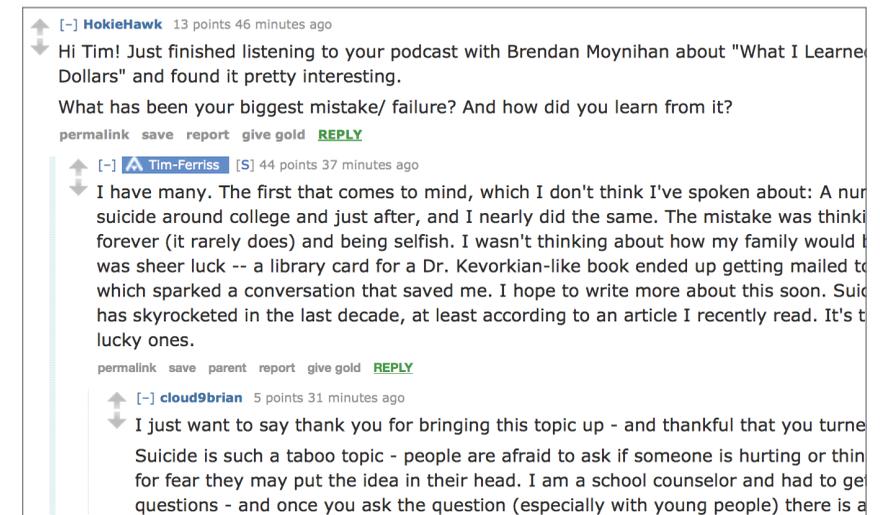


Fig.7: Example AMA comment thread.

The discussion hierarchy can contain multiple conversation chains (FIG.7) which puts cognitive demands on the

authors, commenters, and readers and ultimately can be particularly challenging to piece together after the interview concludes. For example, the order of the comments presented to the user is not chronological, but a function of both time and “votes” received during the discussion. This scattered presentation makes post-hoc viewing or analysis of the conversation difficult, which is further compounded by a lack of advanced search functionality or analytical tools.

AMA provides for us an ideal basis on which to apply our framework for identity within a digital social context. We have present the established community of users, actively participating and engaging with each other. We see a hugely diverse range of backgrounds and topics represented on the platform. We also have a system that has very low barriers to participation – contributing is as easy as making a free Reddit account.

Within the positive and unique aspects of AMA, we also see challenges within both the platform and the digital medium. Further, the natural diversity of the audience and subject matter make the study and design of any such platform a challenge from an audience and user targeting standpoint.

In considering the design of a platform on which conversation organically occurs, the designer must make deci-

sions on the functionality and presentation their system provides the user, to engage with the conversation.

With natural language data, especially conversation based (as opposed to long-form prose), the importance of context to each specific element of the dialogue is critically important. Online knowledge sharing platforms, for example, might default to basing context on a “Question & Answer” format, associating comments to their higher-order parent elements.

As dialogue grows over time, such simple models break down, lacking the context that deeper comments, spawning responses and questions of their own will need to provide the user an intelligible train of thought for following the conversation.

---

## Principles of Data Visualization

It is with these challenges in mind that we look to principles of data visualization and data understanding to evaluate an Online Knowledge Sharing Platform's user experience decisions. (Fry, 2010) proposes a framework for working with and, further, understanding data, which we propose to extend as a framework for how users explore and engage with data from a user experience standpoint:

1. Acquire – the matter of obtaining the data, whether from a file on a disk or from a source over a network.
2. Parse – providing some structure around what the data means, ordering it into categories.
3. Filter – removing all but the data of interest.
4. Mine – the application of methods from statistics or data mining, as a way to discern patterns or place the data in mathematical context.

5. Represent – determination of a simple representation, whether the data takes one of many shapes such as a bar graph, list, or tree.
6. Refine – improvements to the basic representation to make it clearer and more visually engaging.
7. Interact – the addition of methods for manipulating the data or controlling what features are visible.

(Fry, Computational Information Design, 2004)

When considering this framework in the context of designing Online Knowledge Sharing Platforms, we see the first two components – data acquisition and parsing – already occurring on the platform back-end. Users come to the network for the data they have come to rely and expect, and perhaps also take part in producing more of said content. Our platform of interest, reddit, provides this data to users through the webpage user experience, as well as to developers working on the platform through use of its public APIs.

The second two framework components, those of filtering and mining the data, are also features many users come to expect of their online platforms. Integrated search, and options to filter, at a high level have become table stakes for platforms of any scale. In our case, these

basic features were present on the native reddit experience.

From here we move into a space where platform user experience design heavily influences the usability and usefulness of the platform, in the form of the ability of the user to mine and represent the data. While much of the possible statistical and computational work can be dealt with on the backend system, the user must be able to engage and discern the data, to begin to develop their own opinions and perspectives of it.

In this interest, how the data is represented and offered to the user will heavily influence how they approach use of any such a system. This is one area we saw opportunity to explore design variations to the existing reddit experience, as through user research we found the flat forum design hindered any kind of deeper exploration of the data.

Finally, the ability to refine and interact with the data round out the essential elements to data exploration, as defined by Fry. With our purposes in mind, this labels the area of most interest and broader application of this thesis project.

---

## First Research Structure

In order to ground our understanding of AMA participants and their intents behind participation, we devised a research plan that would target users identified through means of the native Reddit platform.

The desired parties were identified as “Authors” who had written an AMA; “Requestors” who had made posts hoping to find an individual willing to author a particular AMA; and finally, “Participants” (or “Commenters”) who had engaged in an AMA in any capacity other than “Authors” or “Requestors”.

To find users fitting these roles, the Reddit API was used to pull account names behind each of, first: the authors of the one hundred top-voted AMA posts of the past month; secondly, the authors of the one hundred top-voted comments made in response to AMAs during the last month; and finally, the authors of the one hundred top-voted “AMA” request posts (labeled “Requests” in the API “Flair” subheading).

With these user names, three online surveys were created on the TypeForm service, all consistent in theme, but each containing specific language pertaining to their respective subgroup. The surveys were sent to the users directly through Reddit, as a private “direct message” wherein the research and researcher were identified as affiliated with Carnegie Mellon University, and the general context of the research was described.

Following an initial outreach to 150 participants, we received approximately 50 responses. A later message blast to the remaining 150 of the original 300 API-generated names resulted in a final set of 72 survey responses. Of these responses, thirty-six were “Commenters”, nineteen “Authors” responses, and seventeen “Requesters” responses.

The content of the survey aimed to understand four specific traits of each user, in regards to their experience with AMA: How did the user hear about AMA; In what manners has the user contributed to an AMA in the past; What motivated the user to participate; and, what would the user improve about the AMA experience.

Additionally, specific questions were added for each group: Authors were asked: how did they prepare for their AMA; Requesters were asked why they chose to ask for an AMA; and, in the event that the AMA Request was

fulfilled, how satisfied they were with the results; finally, Commenters were asked what specific kinds of AMAs they tend to have interest in and how they approach participation.

## First Research Findings



Fig.8: Example diagram of survey responses.

The answers collected were all open-format text input. With this set of data, answers were clustered by subgroup, and then further analyzed. Topic identification

from each response group was achieved through, first, basic noun phrase extraction, followed by hand coding of each noun phrase feature. With these coded noun phrases, the raw answers were again analyzed to identify the occurrences of each coded feature within the subgrouping's response set ("Authors", "Requesters", "Commenters"), of which the sums were tallied. This gave a clear picture of the number of responses, correlated to our coding scheme, submitted by each subgroup.

Finally, a simple visualization (FIG.8) of the forms of participation of each member of each subgroup was created to compare the diversity of engagement, and what shared roles may span our arbitrary subgroupings.

From this research, we found that the community of AMA participants come from a wide range of backgrounds and experience. We also found that Authors cited the fewest instances of involvement outside of their API-identified role as an Author. Our conclusion from this is that Authors are often time less involved, if indeed at all involved in the reddit community, but are more often outsiders who have been asked to take part.

We also found many users reported coming to AMA for the high quality content and discussion. One user cited enjoying AMA more than reading popular tabloid magazines such as *People*.

When asked how they might improve AMA, users most hoped to find ways of better following the conversation – e.g. seeing author-answered questions, downvoted questions, similar questions.

---

## Designing for an Improved AMA Experience

Drawing inspiration from work by Viegas, boyd (Viegas, boyd, Nguyen, Potter, & Donath, 2004), and Harris (Harris, 2011) the goal for this thesis project was to produce an interactive experience, deployable to the public in-the-wild, which would allow for the prototyping, testing, and exploration of ideas and concepts, validated by the very users of the Online Knowledge Sharing Platforms we are concerned with.

In order for such a platform to be built within the time frame of this one-year thesis project, development work began in parallel to early research at the onset of the project. Certain baseline features such as data capture, cleaning, analysis, and storage would play central roles in any design solution dealing with online social network content, meaning work could be began without waiting to conclude user research.

erate about what features to implement, and which to look for research to direct.

### The core features of the prototype version included:

- The Scraper was built to process the Reddit JSON API to archive information about every AMA thread submitted to the subreddit. Built atop the Reddit API, the scraper runs every fifteen minutes, writing an entry for each new AMA thread (and its metadata) into a series of CSV files serving as our static database.
- The search implementation is a basic linear search which queries the scraped CSVs – specifically the description “selftext” and title of each AMA post.
- Thread information visualization was implemented with Bokeh, a Python visualization library. Prototype visualizations included thread comment upvotes and most active thread commenters.
- Noun phrase extraction – labeled “Topics Discussed” – was achieved using the Python TextBlob library for noun phrase chunking. TextBlob is itself based on the Natural Language Toolkit for Python.

## First Version Application Structure

Early components such as a content scraper, search and categorization scripts, data visualization sketches, and server configuration were developed early on, as they were necessary components that would serve the prototype regardless of user research outcomes.

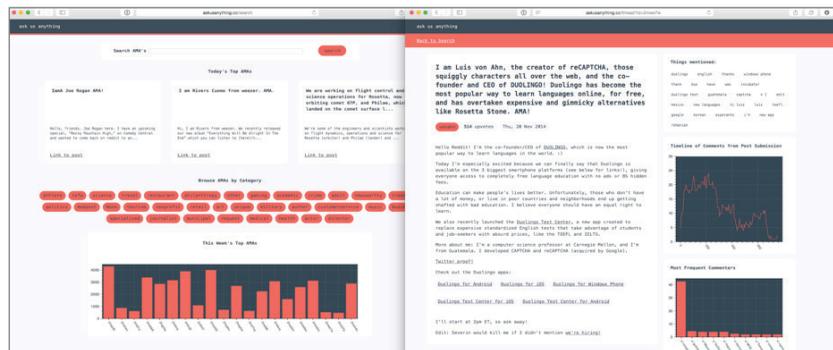


Fig.9: AskUsAnything prototype home page and thread.

The first version of Ask Us Anything was built with the intent of public deployment to AMA users and research participants. To meet this objective in the short timespan of one semester, it was necessary to be very delib-

restrictions through the PRAW library, which monitors request totals and cadence.

Analysis occurs upon the JSON being received via Reddit API/PRAW, and is passed directly to the HTML templates in Flask. These frameworks return renderable HTML from Python, which is passed as a variable within the Flask app and template.

This prototype version (FIG.9) was finalized and made accessible to the public at AskUsAnything.cc where it and all future versions of the project would be hosted.

---

## First User Testing & Findings

With our projects aim to more fully leverage online knowledge sharing platforms in helping shape and share personal stories and experiences, it was important to place the tool in the hands of users of such systems, so that they might consider and evaluate it with their own stories and experiences in mind.

To seed this prototype to a small set of users, we identified the Reddit Development subreddit group as a cohort of potentially interested, but previously unaffiliated users, who were of a small enough user base size so as not to overwhelm our server. A post was made on the RedditDev subreddit (Arrowood, 2015) introducing the project to the community, linking to the prototype, and asking for any feedback.

Early engagement with the prototype seemed very promising. From a very small group of users, we saw over 150 unique visits in the first two weeks, with an average time spent on the site of over four minutes. Feedback through the web form was nearly non-existent, but

some conversation was generated within the original Reddit post we made. One user expressed uncertainty as to the function or service provided by the tool, to which another user, unaffiliated with the project, came to its defense by highlighting what functionality he saw. This same user later followed up with suggestions to improve and build upon what he saw, speaking from experience he himself had building a web app that explores AMA data as well.

Feedback we received during these stages of research can be summarized as below.

#### **Functionality Requests for Authors:**

- The ability to post questions/responses
- The ability for an author to follow new chains of questions as they develop in real-time
- The ability to see priority of questions, aside from time-of-post ordering

#### **Functionality Requests of Commenters**

- The ability to see if the author is still active
- The ability to see when/where author responded
- See author activity aside from responses (upvotes, links shared, references made)

---

## **Second Version Improvements**

Based on the data we received during this phase, work began nearly immediately on what would become the second version of the project.

The first version of the project represented a survey implementation of features and components proposed to enable the aforementioned “archaeology” of the individual AMA. Through user testing we found many aspects of the prototype that did not translate to many users, and in general, that the proposition of an annotated or analytical tool for reading an AMA was a stretch for many users to see the value in.

There were aspects of the prototype that did appeal to many: the visual design pleased many of the users; the natural language processing demonstrated by the noun phrase topic extraction intrigued many users, and gave way to suggestions on how to further that functionality; finally, the highlighting of “Author” activity represented a step forward that many users mentioned hoping to see taken further – to see more of the Author’s activity.



to unpack a conversation at many levels to find interaction by the Author, while other patterns such as multi-author, or highly active or valuable outside contributions may lack the highlighting and emphasis they should warrant. The prototypes visualizations were looking primarily at activity and qualitative contribution metrics. What we lacked were insights into the overall flow of the conversation, and activity within that context. The new model for the thread visualization represents what we believed would be an improvement on the old visualizations, by providing an idea of not only the fingerprints of Authors and contributors on a particular thread, but also of the rhythm and cadence of the AMA. See side-by-side comparison (FIG.12):



Fig.12: Before and after of the thread data visualizations from first (left) and second (right) prototypes.

- **Personalization of the AMA thread**

There is little to distinguish one AMA from another, visually, aside from varying lengths of description text and title texts, while the popular folksonomy of “Proof” is an often-included picture or message from the author, captured outside of the Reddit platform, provided to prove the identity of the author. These pictures are often charming, always unique snapshots of the author, typically with a short message to the community included. This provided us a relevant and charming artefact to use to differentiate those AMAs for which proof is included. See (FIG.13):

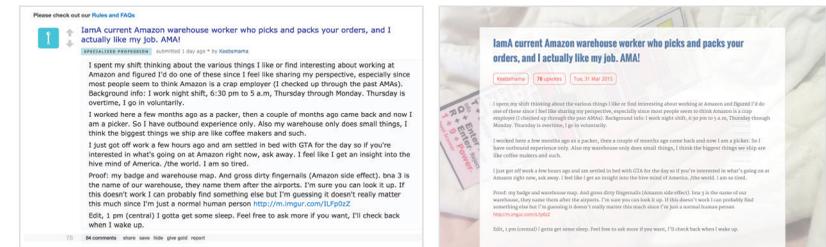


Fig.13: Before and after of the AMA thread’s description header, from native reddit (left) to the second prototype (right).

- **Topic and categorization of the AMA conversation.** The early implementation of noun phrase feature extraction was a performant and, when accurate, highly intuitive method of highlighting the common themes of the conversation. The problem was in the accuracy of such analysis on the often-times spotty corpus of

text within each thread. Spelling and grammatical errors can cause repetitive or similar but misspelled topics to rise to the top, while “meme” or popular references can lose meaning when taken out of context. To improve this experience, we made two improvements. The first was to design a more robust stop-words list, catered to Reddit and AMA, which would help us scrub from the results all of the features that were irrelevant or commonly confused our algorithm and users. The second was to use the categorization as a jumping point to explore the conversation. To this second point, topic linking was included to tie each noun phrase feature in the analysis section of the thread to anchor links wrapped around the points in the conversation where each occurs.

## Second Version Application Structure

This second version of Ask Us Anything was a considerable step forward from the more basic prototype we first shared with users. The improvements to the visualizations, the thread hierarchy representation, and the improved natural language processing represented challenges to implement for a novice programmer, as well as to deploy in a fashion that would be widely viewable and performant to end users on our modestly sized hosting server. The final structure of the application, including libraries and APIs utilized, is represented below (FIG.14).

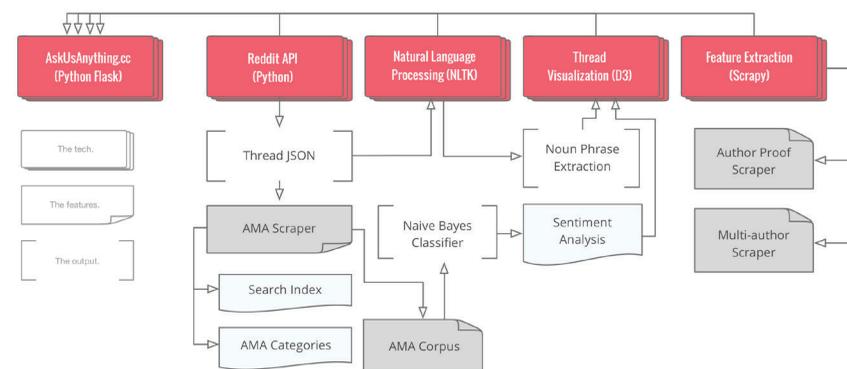


Fig.14: Outline of the application’s structure.

---

## Second User Testing

The new version of AskUsAnything was deployed in late March, 2015. To share this with a broader audience, in the hope of finding more perspectives from which to judge and evaluate the system, we reached out to the moderators of the AMA subreddit on Reddit.com. We sought their guidance on how to share the news of AskUsAnything with the community, and offered to do an AMA as Designers/Researchers exploring the value of AMA and its implications for the design of future knowledge sharing platforms. We were denied, and told that the education/thesis based work did not fall in line with their guidelines for the Author/AMA Topic for the subreddit.

We instead resorted to our prior methodology of using the Reddit API to find relevant, highly qualified users to test our designs with. Another API scrape was conducted to find the Authors of the top 100 AMAs from the month of March. We crafted a short survey, pertaining to the new user experience, and messaged each directly with a link to their AMA thread on our site, and a survey URL.

We had approximately 10% turnout on surveys. This number is hard to determine, as the methodology of finding the top 100 authors leaves you susceptible to pulling in user names of accounts that were setup for the express purpose of authoring AMA. For example, the polarizing digital security and privacy activist Edward Snowden conducted an AMA using an account setup solely to do his AMA. These accounts are often times created with the help of Reddit's AMA staff, and are abandoned a short time after the AMA concludes.

These two challenges represent a notable reality and design consideration for digital knowledge sharing communities: First, over specificity of theme or context can make the cross-expertise or cross-topical collaboration difficult if not forbidden (as represented by our interest in designing something for a community, and sharing it with that community, only to be refused an audience due to the context of our project falling outside the communities inherent purpose); Secondly, the ephemerality or anonymity of online identity can run counter to the nature and purpose of some systems for higher-order analysis and engagement.

The result of our qualitative research were very polarizing, as many users found tremendous enjoyment and value to the experience, while others felt it lacked some of the nuance and detail they came to value in the native

Reddit experience. This insight is valuable to understand for any community one may design for, as the expected heuristics of any platform or prior-existing practice truly shapes and defines the nature of the experience. As users became used to reading AMA threads, inclusive of the secondary and tertiary discussion threads that would occur in response to comments made within the AMA, any endeavor to reframe or exclude parts of that experience can confuse and frustrate many users.

---

## Second User Testing Findings

Having thoroughly researched multiple classes of user types of our network of interest, reddit “Ask Me Anything”, we found the composition of the community to be unique and amorphous. Users engage in this kind of online knowledge sharing with diverse intents and perspectives in mind.

Authors may have felt a strong self-driven desire to contribute, while others may have been pressured into contributing by a friend or colleague.

Commenting participants themselves may choose to engage out of non-consequential curiosity (“I find the concept of communicating with celebrities entertaining.” – “Commenter” research participant) or out of interests stemming from shared or similar experience (“Through personal connection with people I know it motivates me to develop more knowledge in fields I like. – “Commenter” research participant).

An interesting subgroup of Online Knowledge Sharing

communities are the users who request information and data from other users. These users, labeled “Requesters” in our research, serve an important role, particularly in Reddit AMA, as a large number of user authors are compelled into action through the requests of others. Through our API targeting, we found 100 AMA Requesters and asked them a series of questions, including “What motivated you to request an AMA?”. Examples of the range of answers we received are below.

Some requesters seek perspectives of people who have done work they respect or are fans of:

“I was working about a very well known book (The Things They Carried) and while doing my research I found that the author ceaselessly gives interviews and the like. And since Americans respect their veterans a lot I thought it would be a great AMA if the author, Tim O’Briend could go to reddit.”

Other requesters look for entertainment:

“The fact bands where also participating so I wanted to request my favorite band ‘Trivium’.”

While some requesters rely on this kind of dialogue to supplant older models of media and publishing:

“A look into who someone is via their own words on random subjects is much more enlightening than people or teen beat.”

Given this diversity of intentions and motivations for contributing, and factoring the inherent cultural and individual diversity of users represented by any social network, the value of extensive user research cannot be understated for any endeavor such as our own.

---

## Implications for Online Knowledge Sharing Platforms

The work of this thesis project spans multiple disciplines and research interests. As a design research project, the focus of the project was on realizing the full potential for online knowledge sharing platforms in helping shape, develop, and share personal stories and experiences.

The value of such work is only itself fully realized if it enables the work of future designers and researchers of similar systems. Further, any such design innovation or intervention should only be made if the designer can be certain, through careful testing and research, that the foundation of the community – the active and contributing member base – find no new challenges or alienation within the new system. To this end, here outlines a number of principles used in setting up research with this projects community of users.

### 1. Understanding the community

All knowledge sharing platforms rely on their collective members and users to produce, participate in,

and pass on the knowledge content from which they derive their name. This community will be representative of a wider base of individuals than just those who hold credentialed membership (user accounts, logins, invites, etc.).

Roles such as authors, commenters, and requesters serve as more publicly facing members of the community, while moderators, supporters, and non-participating followers round out a broader spectrum of users, each with specific interests and use-cases.

### 2. Understanding the history of the platform

The growth and development of such a community is a process that take quite a while. Knowing the genesis of such platforms, how the community evolved in support of the context, and how the community differentiates itself from other similar communities will serve to guide you in basing judgments and decision on which areas of opportunity to investigate.

### 3. Understanding the wider value

Online networks often develop heuristics, vernaculars, and communication models which support and inform the message and content produced therein. In the case of online knowledge sharing platforms, such models can serve as barriers to those outside the network and community.

Given the nature and spirit of the sharing taking place within such communities, it is important to find the value inherent within, and to further identify the outside or broader community that might benefit or be able to contribute, which has gone un- or under-addressed.

These principles propose no answers, nor do they offer any specific questions. Instead, they position the designer and researcher to make decisions and investigate opportunities with an honest and ambitious eye towards improving the community for-which they are designing. With these principles established, one can look towards more universal principles of design for guidance on how to approach improving the experience of the platform.

---

## Implications for Broader Online Social Networks

This thesis project has aimed to explore the implications for various design approaches with regard to online knowledge sharing platforms. In doing so, it has drawn inspiration and guidance from the work of many past designers and researchers. In order to better understand the implications for this work beyond the specific context and community for which it is operating within, recall our basing of design decisions and evaluation of user experience according to the framework (Fry, 2004) proposes for data understanding:

1. Acquire – the matter of obtaining the data, whether from a file on a disk or from a source over a network.
2. Parse – providing some structure around what the data means, ordering it into categories.
3. Filter – removing all but the data of interest.
4. Mine – the application of methods from statistics or data mining, as a way to discern patterns or place the

data in mathematical context.

5. Represent – determination of a simple representation, whether the data takes one of many shapes such as a bar graph, list, or tree.
6. Refine – improvements to the basic representation to make it clearer and more visually engaging.
7. Interact – the addition of methods for manipulating the data or controlling what features are visible.

(Fry, Computational Information Design, 2004)

Current online social networks hold incomprehensible amounts of information about their users, much of a distinctly private nature. As users come to rely on such networks to provide timely, accurate, and relevant access to this data, the aspects of data understanding that we focused out design attention towards – refinement and interaction – represent an immense area of opportunity.

While this project explored propositions for how to improve a native experience through these principles, many other opportunities still exist to be explored. In looking to define the data-driven user experiences of social networks.

---

## Future Work

Choosing one existing platform to iterate and prototype on, we naturally open discussion of related future work on other, similar platforms and communities. The broad appeal and rising popularity of Online Knowledge Sharing Platforms ensures similar work will continue, and the use and value of such systems will grow over time.

Within AMA, we saw a format that relied heavily on identity and a sense of authority or credibility of authors. Future work might explore the role anonymity or differing forms of merit/credibility might influence user and design decisions. The role of the “throw-away” account, or the shroud of anonymity as an enabler of honesty, transparency, or unpopular truth/opinion is an intriguing question of today’s modern hyper-public networks.

The long-form style of many AMA discussions differ highly from some Online Knowledge Sharing Platform’s which focus on brief, quick answers to specific questions. Work similar to ours on such systems (e.g., Quora, Yahoo! Answers, r/AskReddit) would find huge audiences

of users on platforms with many of the same, if not more opportunities.

Larger scale, community oriented networks of Online Knowledge Sharing, such as Wikipedia, rely on collaboration and iteration to enable accuracy and quality. Question & Answer style communities such as StackOverflow and Quora have come to rely on similar models. These networks present great opportunity for similar work on enabling more contextually annotated, interactive user experiences that could challenge the voting-based discussion thread style employed commonly.

---

## Discussion

This project has brought together concepts of identity, storytelling, social network design, data visualization, and user experience design & research that represents what we believe is unique, novel, and valuable work for a design-based thesis project.

User research shows that networks such as those we considered provide immense value to individuals, and that work to improve these networks can find support and enthusiasm.

Such work does risk endangering the unique proposition and heuristics users have come to rely on when using such systems. Through research-based understanding of the user community, and the value the network provides to users, as well as designing with data understanding and “accessibility” principles in mind, we see great potential to make these networks even more valuable sources of knowledge, inspiration, and meaning to users.

---

## Bibliography

Arrowood, C. (2015, 2 15). Building a tool for Reddit AMA - Take a peek! Retrieved 4 28, 2015, from reddit.com: [http://www.reddit.com/r/redditdev/comments/2v5s0c/building\\_a\\_tool\\_for\\_reddit\\_ama\\_take\\_a\\_peek/](http://www.reddit.com/r/redditdev/comments/2v5s0c/building_a_tool_for_reddit_ama_take_a_peek/)

Boe, B. (n.d.). PRAW: Python Reddit API Wrapper. Retrieved 4 29, 2015, from <http://praw.readthedocs.org/>

Bostock, M. (2011, 2 18). D3: Data-Driven Document. Retrieved 4 29, 2015, from Introduction: <http://d3js.org/#introduction>

boyd, d. (2014). *It's Complicated: The Online Life of Teens*. New Haven, CT: Yale University Press.

boyd, d. (2008). *Why Youth Heart Social Network Sites: The Role of Networked Publics in Teenage Social Life*. Cambridge, MA: The MIT Press.

Campbell, J. (1949). *A Hero with a Thousand Faces*. New York, NY: Pantheon Books.

Donath, J. (2014, 04 25). We Need Online Alter Egos Now More Than Ever. Retrieved 04 28, 2015, from Wired.com: <http://www.wired.com/2014/04/why-we-need-online-alter-egos-now-more-than-ever/>

Eco, U. (2009). *The Infinity of Lists*. New York, NY: Rizzoli.

Feinberg, M. (2013). Comparative Appraisal of Expressive Artifacts. CHI 2013 Extended Abstracts (pp. 1999-2008). New York, NY.: ACM Press.

Ferriss, T. (2015, 4 29). I am Tim Ferriss, author, angel investor, host of the Tim Ferriss Experiment, and human guinea pig. AMA! Retrieved from reddit.com: [http://www.reddit.com/r/IAmA/comments/34ak5w/i\\_am\\_tim\\_ferriss\\_author\\_angel\\_investor\\_host\\_of/](http://www.reddit.com/r/IAmA/comments/34ak5w/i_am_tim_ferriss_author_angel_investor_host_of/)

Forte, A., Andalibi, N., Park, T., & Willever-Farr, H. (2014). Designing Information Savvy Societies: An Introduction to Assessability. SIGCHI Conference on Human Factors in Computing Systems (pp. 2471-2480). New York, NY.: ACM Press.

Fry, B. (2004). *Computational Information Design*. Thesis, Massachusetts Institute of Technology, School of Architecture and Planning, Cambridge.

Fry, B., & Reas, C. (2011, 3 1). Processing.org. Retrieved

4 29, 2015, from Overview: <https://www.processing.org/overview/>

Harris, J. (2011, 12 08). Cowbird. Retrieved 4 28, 2015, from Cowbird: <http://cowbird.com/about/>

Heer, J., & boyd, d. (2005). Vizter: Visualizing Online Social Networks. 2005 IEEE Symposium on Information Visualization (p. 5). Washington, DC.: IEEE Computer Society.

Kamvar, S. D., & Harris, J. (2011). We Feel Fine and Searching the Emotional Web. Fourth ACM International Conference on Web Search and Data Mining (pp. 117-126). New York: ACM Press.

Kittur, A., & Kraut, R. E. (2008). Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. 2008 ACM conference on Computer supported cooperative work (pp. 37-46). New York, NY.: ACM Press.

Kittur, A., Suh, B., & Chi, E. H. (2008). Can You Ever Trust a Wiki? Impacting Perceived Trustworthiness in Wikipedia. ACM Conference on Computer Supported Cooperative Work (pp. 477-470). New York, NY.: ACM Press.

reddit Inc. (n.d.). Retrieved 4 29, 2015, from reddit API: <http://www.reddit.com/dev/api>

Turkle, S. (2012). Alone Together: Why We Expect More from Technology and Less from Each Other. New York, NY: Basic Books.

Turkle, S. (1995). Life on the Screen: Identity in the Age of the Internet. New York, NY: Simon & Schuster.

Turkle, S. (1984). The Second Self: Computers and the Human Spirit. Cambridge, MA: The MIT Press.

Viegas, F., boyd, d., Nguyen, D. H., Potter, J., & Donath, J. (2004). Digital Artifacts for Remembering and Storytelling: PostHistory and Social Network Fragments. 37th Hawaii International Conference on System Sciences. New York: ACM Press.

Wang, G., Gill, K., Mohanlal, M., Zheng, H., & Zhao, B. Y. (2013). Wisdom in the Social Crowd: an Analysis of Quora. 22nd International Conference on World Wide Web Pages (pp. 1341-1352). Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Weninger, T., Zhu, X. A., & Han, J. (2013). An Exploration of Discussion Threads in Social News Sites: A Case Study of the Reddit Community. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 579-583). New York, NY.: ACM Press.

# Fall Thesis Poster

# ASK US ANYTHING.cc

## STATEMENT OF OPPORTUNITY

Online knowledge sharing platforms and communities have become a popular and reliable source of information on the social web. As these platforms become more prevalent, the design of such systems to enable not only participation and quality [5], but also higher-level analysis and comprehension [3] - "information assessability" [1] - is of critical importance.

## BACKGROUND & PROJECT DESCRIPTION

Traditionally, online knowledge sharing platforms take the form of either encyclopedic data [3], or conversational question and answer threads [5]. New models of knowledge sharing have developed over time [3], the such model is that of the "AMA" or "Ask Me Anything".

Ask Me Anything is a crowd-sourced interview made famous on the social website Reddit [4]. In an AMA, an individual (typically of an interesting background or qualification) makes themselves available to the community for an interview, through comment submission, using the websites traditional discussion thread.

This thesis project, entitled "Ask Us Anything", is the design and development of a platform aiming to enable the aforementioned analysis and "information assessability" of these AMA interviews. The platform further enables user-study and research into the applications and usefulness of such a tool provided user-member types who participate in "AMA".

## USER RESEARCH

In order to further understand the type of user who participates in an AMA, a survey was sent to 300 Reddit AMA users. Using the Reddit API, three groups of 100 users each (see aside) were identified by merit of **Authors** of the top 100 AMA submissions from the past month, **Commenters** who wrote the top 100 AMA comments from the past month, and **Requesters** of the top 100 AMA Requests of the last month. We received a total of 76 responses.

The findings visualized below represent preliminary patterns identified in the survey analysis. Included within are patterns of participation, coded means of discovering AMA, and coded reasons for contributing.



## PROTOTYPE

### RESEARCHING THROUGH DESIGN

Ask us Anything (<http://AskUsAnything.cc>) is our prototype web application, built to analyze live AMA threads. The prototype features natural language analysis, metadata analysis, as well as categorical and historical analysis.

Our intent is to deploy this prototype to the Reddit community, in the second semester, and with further improvements. We see immense value in allowing the analytic tool we build to be used by ourselves, as researchers, as well as by the community user-base of study.

The prototype was built in Python Flask, utilizing Boken for data visualization, and NLTK for natural language analysis.



Homepage View

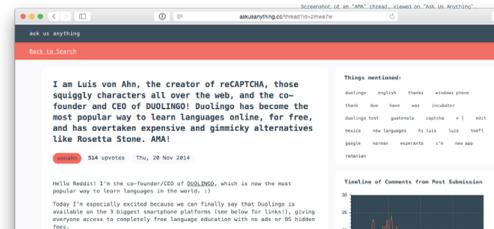
The homepage presents a categorical, as well as time-based representation of AMA threads - those from the past and those currently occurring.

AMA Thread View

The thread view presents the user with author highlighting, topical non-phase clustering, comment activity history, author frequency, and options to sort.

## LITERATURE REVIEW

[1] Foray, A., Ambaldi, M., Park, T., Willinger-Ferr, M. Designing Information Query Interfaces for Information Assessability.  
 [2] Foster, A., Bai, B., Dai, E. Can You Ever Trust a Wiki? Detecting Perceived Trustworthiness on Wikispaces.  
 [3] Wang, G., Gill, K., Hensel, M., Zhang, H., Zhou, B. Wisdom in the Social Crowd: an Analysis of Quora.  
 [4] Wieringer, T., An, A., Han, J. An Exploration of Discussion Threads in Social News Sites: A Case Study of the Reddit Community.  
 [5] Kotler, A., Kozl, K. Harnessing the Wisdom of Crowds in Wikipedia: Quora Through Crowdsourcing.  
 [6] Shalun, S. The Second Self: Computers and the Human Spirit.  
 [7] Miller, T. A Book of One's Own.  
 [8] Kasey, M. On Twitter.  
 [9] Campbell, J. A View with a Thousand Faces.



## PROJECT TIMELINE

- AUGUST
  - Literature Review
- SEPTEMBER
  - Concept Explorations, Community Identified
- OCTOBER
  - Research Plan & IRB Approval, Begin Prototype Development
- NOVEMBER
  - Survey & Analysis
- DECEMBER
  - First Prototype Complete, Identify User-Testing Participants, Writing CHI WIP
- SEMESTER BREAK
- JANUARY
  - IRB Revision, Begin User-Testing, CHI WIP Submit
- FEBRUARY
  - User-Testing Concludes, Prototype Revisions
- MARCH
  - User-Testing Synthesis, On-going Improvements
- APRIL
  - Soft-Launch of App
- MAY
  - Thesis Book, Spring Poster Show

# Spring Thesis Poster

## ask us anything for reddit ama

Investigating the role online social networks play in the development and sharing of personal stories and experience.

### The Platform & Community

Online knowledge sharing platforms support the generation, distribution, and discussion of information benefiting the crowd. Reddit is a news sharing website allowing users to vote on popular news and shared items, featuring only the highest rated content.

Within Reddit, a subcommunity called "Ask Me Anything" (AMA) hosts open-format, public interviews with diverse constituents including Barack Obama, Edward Snowden, patients with ALS and Cancer, and even a Vacuum Repair Man.

More so than just knowledge, people come to AMA for perspective. These are not quick tips and answers, but rather whole stories that users seek to hear and share.

"... mine was about making strides to move on from a poor beginning to adulthood, and as part of my counseling and rehab processes, I decided to use an AMA as a way to begin talking about it."

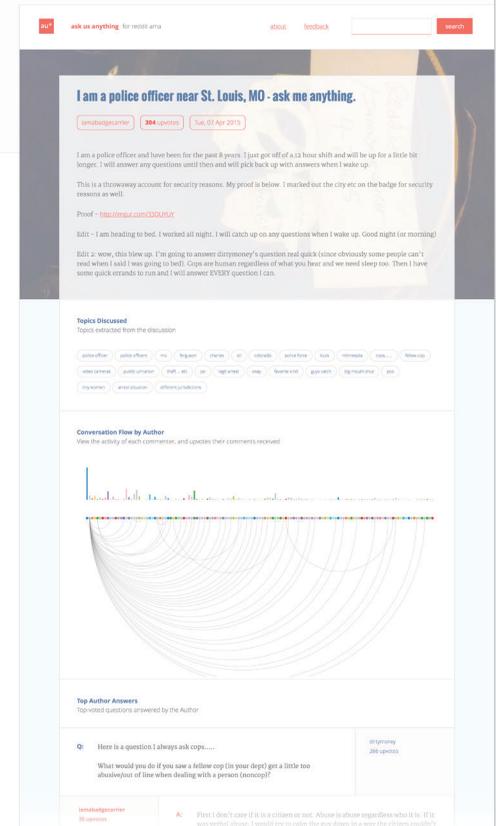
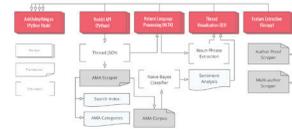
- A recent AMA Author on "Why did you lose an AMA?"

### Challenge

AMAs are fast moving conversations that can be difficult to review after the fact. They are presented using the general reddit interface which can be non-intuitive to novices. As the conversation grows in length, it also increases in complexity, as many of the highest quality responses are in-fact responses to other responses. This nested nature makes the skimming and review of an AMA a challenge, and masks the benefit and information from many users.

### Response

Build a working, publicly accessible prototype for AMA users to review, explore and consider AMA conversations. Evaluate with top author, commentator and requestor participants in the AMA community. The prototype's architecture is illustrated here:



### The New User Experience

#### Features

**Question & answer highlighting**  
Reformat the discussion to provide an easier to skim summary of author answers, with ability to expand full conversation threads.

**Topic extraction**  
Show the most frequently mentioned topics, linking to the points in the conversation where each is mentioned.

**Conversation mapping**  
Show contributions and length of involvement in discussion for each user; upvotes, and sentiment analysis for each comment.

#### User Feedback

"If it hasn't been done already, after the AMA ends, all the questions and their answers should be provided in an interview form..."  
Most common request among 25 AMA Participants

"There should be some system of automatically up-voting posts where people comment so they don't get lost in the vast chaos that is Reddit..."  
Most common request among 22 AMA Authors

"Consolidation of similar questions and increased visibility for replies written by the [Original Poster]?"  
Common request among 35 AMA Participants

#### Feature Examples



### The Final Stages

With the final AskUsAnything experience nearing completion, I am soliciting more feedback from users. With nearly 1,000 monthly unique visitors and high engagement, this feedback is of great interest. Over 100 users have already been surveyed, and final rounds of user testing are concluding as of the time of this poster's creation.

### Discussion of Findings

Reddit users who have tested AskUsAnything have had generally positive things to say. Many users have challenged to dig further into the data and content, showcasing deeper insight. Some users want to go beyond the singular AMA, in order to find related and similar AMAs. The value in AMA, and platforms like it, are in the knowledge and experience of many.

Designing for AMA has proved to be a challenge. As systems on the internet give form to new methods of communication, designing outside of the bounds of the formative community or platform risks making the message and mastery of users. The experience I have had here will hopefully give valuable guidance and principles for designers of future OKRs.

### AskUsAnything.cc

By Chris Arrowood  
Advised by Aisling Kelliher  
Master of Design, Interaction Design  
Carnegie Mellon University  
School of Design

---

## Appendix

### PROJECT TIMELINE

- **AUGUST**  
Literature Review
- **SEPTEMBER**  
Concept Explorations,  
Community Identified
- **OCTOBER**  
Research Plan  
& IRB Approval,  
Begin Prototype  
Development
- **NOVEMBER**  
Survey & Analysis
- **DECEMBER**  
First Prototype Complete,  
Identify User Testing  
Participants,  
Writing CHI WIP
- SEMESTER BREAK
- **JANUARY**  
IRB Revision,  
Begin User Testing,  
CHI WIP Submit
- **FEBRUARY**  
User Testing Concludes,  
Prototype Revisions
- **MARCH**  
User Testing Synthesis,  
On-going Improvements
- **APRIL**  
Soft-launch of App
- **MAY**  
Thesis Book,  
Spring Poster Show

---

## Project Timeline

---

## First Survey Questions

### Authors

- **How did you discover “AMA”?**  
*A bit of background on when you first heard of, read, or were asked to participate in an “ask me anything” thread.*
- **In what ways have you participated in “AMA” discussions in the past?**  
*Please feel free to check all that apply.*
  1. Author of AMA by request
  2. Author of AMA by volunteering
  3. Requester of an AMA
  4. Commenter participant in AMA
  5. Non-contributing browser of AMA
- **What motivated you to do an “AMA”?**  
*Were you internally (self) or externally motivated to do an “AMA”? Can you describe that motivation?*
- **Can you describe your experience authoring and conducting your “AMA”?**

*How you prepared before? Where were you during the “AMA”? What happened at the conclusion of the “AMA”?*

- **How you prepared before? Where were you during the “AMA”? What happened at the conclusion of the “AMA”?**

*This is optional, but if anything comes to mind...*

### Participants

- **How did you discover “AMA”?**  
*A bit of background on when you first heard of, read, or were asked to participate in an “ask me anything” thread.*
- **In what ways have you participated in “AMA” discussions in the past?**  
*Please feel free to check all that apply.*
  1. Author of AMA by request
  2. Author of AMA by volunteering
  3. Requester of an AMA
  4. Commenter participant in AMA
  5. Non-contributing browser of AMA
- **What motivates you to participate in “AMA”?**
- **How satisfied are you with the information you get through “AMA’s”?**  
*0-5*

- **How would you improve the “AMA” experience for participants like yourself?**

*This is optional, but if anything comes to mind...*

- **How would you improve the “AMA” experience for participants like yourself?**

*This is optional, but if anything comes to mind...*

## Requesters

- **How did you discover “AMA”?**

*A bit of background on when you first heard of, read, or were asked to participate in an “ask me anything” thread.*

- **In what ways have you participated in “AMA” discussions in the past?**

*Please feel free to check all that apply.*

1. Author of AMA by request
2. Author of AMA by volunteering
3. Requester of an AMA
4. Commenter participant in AMA
5. Non-contributing browser of AMA

- **What motivated you to request an “AMA”?**

- **Was your “AMA” request fulfilled?**

- **If your “AMA” request was fulfilled, were you satisfied with the information you discovered through “AMA”?**

0-5

---

## Second Survey Questions

### Authors

- Were you able to view your recent AMA thread on AskUsAnything.cc? (You received a link to your thread via Reddit PM)
- Have you re-read your AMA on Reddit, since it completed?
- If you answered 'Yes' to the previous question, could you perhaps tell us why you re-read your AMA?
- Did you find the reading experience on AskUsAnything informative?
- Perhaps you could tell us why you gave the answer you did to the above question?
- Do you believe the AskUsAnything thread accurately captured your AMA, as you remembered it?
- Perhaps you could elaborate on your rating, provided above, of the "accuracy" of the AMA?
- How does the presentation of AMA information in

AskUsAnything compare to the Reddit experience?

- Perhaps you could expand on your rating to the above question on how the information presentation compares to Reddit?
- Can you think of any ways to improve AskUsAnything?



