Pique Requirements

Maria Strasky, Sammie Jiang, Reid Oliveira, Erik Dolinsky

Product Vision Statement

Motivation and Opportunity

The Pique web application will be an automated engine for trending web content, displaying the top content across several platforms in one space. Pique will save users time and the users will know that they are able to keep up with the current trending content from platforms on which they may not be active (Twitter, Reddit, and Imgur) with minimal latency. This will be highly desirable as more social media platforms emerge and it becomes time consuming to keep up with them all individually.

Problem Statement

The problem of accessing popular and trending content across several sites being difficult and time consuming to access affects millennials and web-conscious professionals, the impact of which is wasted time and disinterest from seeing repeated content. A successful solution would be a web or mobile application that automatically aggregates popular and rapidly trending content and displays it in a collective space.

Product Position Statement

Designed for social media consumers who browse the internet for stimulating content, our system is software that pulls top and trending content from social media sites via their respective APIs and displays it on aggregate. Unlike Hootsuite or Flipboard, our product provides trending content from social media on aggregate; not just that which pertains to a particular brand (as Hootsuite does), and more than just trending news stories (as Flipboard does).

User Demographics

As Pique is intended to be quick and easy to use, users of Pique fall into just one category.

Users: Social media users who want to easily and quickly view the top trending public content each day across all major social media sites. They will be expecting a simple interface where you can immediately view the content, without having to navigate to different places.

Feature List

- Top and Trending public posts from Twitter, Imgur, and Reddit

    - Top – most likes, shares, and comments
    - Trending – quickly gaining likes, shares, and comments

- Ability to sort by hashtags; to search for content by a specific hashtag
- Ability to search for similar posts containing the same hashtag
- Filtering of repeated posts

Constraints

- Pique must use open-source libraries
- Pique must operate affordably
- Pique must be easily deployable and fully configurable; this will enable us to use both local machines and industry-standard virtual machines such as an Amazon Web Services EC2 instance to host the application throughout development.
- Pique must use publicly available API endpoints from a subset of major social media services, and is subject to the limitations set by each of these APIs (including limitations to content access, request rate limits, and others).

Scope and Limitations

- Features Included:

    - (see feature list)

- Features not Included:

    - Native Mobile Application (web application only)
    - User login accounts, which means no personalized content (everyone sees the same thing)
    - The "social" part of social media where you're sharing and interacting with other people (Pique is intended only for media consumption)

Assumptions and Dependencies

- We are able to poll each of the selected services for sufficient content and supporting data pertaining to said content so that we may formulate a scoring algorithm.
- Major browsers will remain compatible with the product (and vice versa) for the duration of its operational lifespan. These browsers include Google Chrome, Mozilla Firefox, and Apple Safari.
- Social media sources will remain responsive throughout the vast majority of the product's operational lifespan.

Use Cases

This section outlines the main use cases of Pique.

For all use cases that display stories on aggregate, individual pages of up to 50 stories will load. An additional page will be appended to the existing page(s) when the user reaches the bottom of the page. There is no option to filter by content source, although this could be implemented as an extension of the basic functionality in the future.

Each post will contain the post's text content if available, image content if available (or cover image if the post is a collection of images), timestamp of content collection, alias of content creator, source platform information, hashtags contained in the post, as well as the number of likes, comments, and shares (or equivalent metrics depending on the source).

Use Case 1

- **Identification:** Use case 1 – *Find Top Stories*
- **Primary Actor:** A web user who wants to view the top stories (stories with the most likes or shares) from social media sites.
- **Stakeholders and Interests:** This would impact all users of this product, as this is one the main functionality of Pique.
- **Preconditions:** None
- **Postconditions:** After this use case, the main page of the website will display the current top stories from social media sites.

- **Main Success Scenario:**

  1. The user accesses the Pique main page for the first time in a session, or accesses "Top" from the menu bar at any point during a session
  2. Previews of top stories are displayed on the web page

- **Extensions and Alternative Flows:**

  - The user Reads a Story (extension)
  - The user Finds Related Stories (extension)

- **Open Issues:** None

Use Case 2

- **Identification:** Use Case 2 – *Find Trending Stories*
- **Primary Actor:** A web user who wants to view trending social media stories (stories that are gaining likes or shares quickly).
- **Stakeholders and Interests:** This would impact all users of this product, as this is one of the core functionalities of Pique.
- **Preconditions:** None
- **Postconditions:** After this use case, the main page of the website will show trending stories instead of the previously displayed stories.

- **Main Success Scenario:**

  1. The user accesses "Trending" from the menu bar
  2. Current trending stories populate in place of the previously displayed content

- **Extensions and Alternative Flows:**

  - The user Reads a Story (extension)
  - The user Finds Related Stories (extension)

- **Open Issues:** None

Use Case 3

- **Identification:** Use case 3 – *Find Stories by Hashtag*
- **Primary Actor:** A user who wants to view all stories that contain a specific hashtag.
- **Stakeholders and Interests:** Although not all users may use this functionality, it will remain available for those who are interested.
- **Preconditions:** The Pique backend has gathered and stored posts containing the desired hashtag.
- **Postconditions:** After this use case, the Pique main page will show stories containing the desired hashtag in place of the previously displayed content.

- **Main Success Scenario:**

  1. The user searches for a hashtag in the menu bar
  2. Stories containing the desired hashtag populate in place of previously displayed content

- **Extensions and Alternative Flows:**

  - The user Reads a Story (extension)
  - The user Finds Related Stories (extension)
  - No posts for that hashtag exist; an error message is shown (alternate flow)

- **Open Issues:** None

Use Case 4

- **Identification:** Use case 4 – *Find Stories with Top Hashtag*
- **Primary Actor:** A user who is interest in a particular trending topic, but may not know what that topic is.
- **Stakeholders and Interests:** Although not all users may use this functionality, it will remain available for those who are interested.
- **Preconditions:** Pique has evaluated its most popular hashtags at that point in time.
- **Postconditions:** After this use case, the Pique main page will show stories containing the desired hashtag in place of the previously displayed content.

- **Main Success Scenario:**

  1. The user targets a top hashtag that is displayed in the menu bar
  2. Stories containing the top hashtag populate in place of the previously displayed content

- **Extensions and Alternative Flows:**

  - The user Reads a Story (extension)
  - The user Finds Related Stories (extension)

- **Open Issues:** None

Use Case 5

- **Identification:** Use case 5 – *Read a Story*
- **Primary Actor:** A web user who wants to read a story previewed on the web page.
- **Stakeholders and Interests:** This would impact all Pique users, as this is one of the main functionalities of Pique.

- **Preconditions:**
  Any one of:

    - Use case 1 (Find Top Stories)
    - Use case 2 (Find Trending Stories)
    - Use case 3 (Find Stories by Hashtag) (Main Success Scenario)
    - Use case 4 (Find Stories with Top Hashtag)

- **Postconditions:** After this use case, the main page of the website will have been re-arranged so that the selected story is expanded.

- **Main Success Scenario:**

    1. The user opens a story that they want to read in full
    2. The story expands on the page and all other stories below it are moved down to accommodate this

- **Extensions and Alternative Flows:**

    - The user Finds Trending Stories (extension)
    - The user Finds Related Stories (extension)

- **Open Issues:** None

Use Case 6

- **Identification:** Use Case 6 – *Find Related Stories*
- **Primary Actor:** A web user who wants to view all stories that have a specific hashtag, having viewed a story that contains the specific hashtag.
- **Stakeholders and Interests:** This would impact all Pique users. It's possible that not all users will use this functionality, but likely most of them will.
- **Preconditions:** Use case 5 (Read a Story)
- **Postconditions:** After this use case, the page will display stories that contain the targeted hashtag.

- **Main Success Scenario:**

    1. The user targets one of the hashtags contained in a story
    2. The web page displays all stories with that hashtag in place of the previously displayed content.

- **Extensions and Alternative Flows:**

    - The user Finds Trending Stories (extension)
    - The user Reads a Story (extension)

- **Open Issues:** None

Non-Functional Requirements

Performance Requirements

- Pique must provide users with content access at a lower content consumption latency (that is, the time between content creation on the source site to consumption by the Pique user) than accessing multiple social media sites for the same top or trending content.
- Pique must be significantly easier to use than browsing multiple social media sites either simultaneously or successively.
- Database Access Object must be able to handle a load of up to 100,000 requests per second, in order to handle both a large number of users and simultaneous I/O associated with bulk data processing by the sorting node.
- The Pique landing page must respond and load within 0.5 seconds of a request, barring latency due to legacy client hardware. This stems from the requirement that Pique must be fast and easy to use, in order to convince users to choose Pique over browsing multiple social media platforms.
- The Pique data cache must not exceed a reasonable storage limit.

Security Requirements

- Pique must be registered as an application with each of its source social media platforms.
- Content creators hold all rights to the data that is collected by and displayed on Pique. However, they will have agreed to the collection of this data by third-party applications as part of the terms of use of each of the respective source applications.

Software Quality Attributes

- Pique will need to be adaptable, especially when considering adding further social media sources, or removal of existing social media sources from Pique's lineup without major development effort.
- The Pique landing page will remain available 99% of the time during operation.
- Pique source code will remain fully tested and maintainable with every release, ensured via peer code and test suite review.
- Pique will remain intrinsically portable on the JVM.

Requirements Changes and Rationale

- *Thu. Nov. 3*: Reddit upgraded from a potential source to a confirmed source. At this stage, we were debating whether we should include Facebook as a source (see below), but also understood that we should keep a diverse set of sources. According to our findings, Reddit would be a good fit for the intended user base, and not too difficult to implement.
- *Tue. Nov. 8*: Use case 4 *Find Stories with Top Hashtag* upgraded from a potential feature to deliverable. Our modular and flexible design of the data tier allowed for relatively simple implementation of this feature, and we feel that it would be a particularly valuable feature for a subset of users.

- *Mon. Nov. 14*: Facebook removed as a content source. This was done for the following reasons:

  - The Facebook Graph API (v1.0) originally had the ability to obtain individual posts easily, but this has since been removed. It would now be significantly more difficult to pull individual posts from user and group feeds. We have discussed how this may be attained, but this task would require setting up special architecture to monitor specific groups and popular users on Facebook, and grow this viewing network with a machine learning algorithm. Another possible solution would be to provide user-specific content via user login, but this again would require further system architecture, and would deviate from the purpose of Pique as a means of consuming top, publicly available data. Any potential solutions were deemed too great a development effort for the product release timeline.
  - The proportion of viral original content (in more general terms, the content that Pique aims to show its users) posted on Facebook is much lower than that which is posted on other platforms, and a good portion of that information is first posted on other sites focused on public content (such as Reddit, Imgur, or Twitter) and then later propagates to private social media (See Lifespan and propagation of information in On-line Social Networks: a Case Study (https://arxiv.org/pdf/1403.1486.pdf)). Thus, including Facebook would work against Pique's promise of minimizing latency between content creation and consumption, as a good portion of content gathered from Facebook would be outdated, and likely already gathered from another source.
  - There is increasing concern about 'echo chambers' in social media, most of which is directed at Facebook, whose private platform and graph-centric approach to social media propagates the creation of these echo chambers. See Echo Chambers on Facebook (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2795110) for more information, and *Research at Facebook*'s own Exposure to Diverse Information at Facebook (https://research.facebook.com/blog/exposure-to-diverse-information-on-facebook/) for a contrasting viewpoint. This would work against the goal of having Pique serve a broad set of top and trending data.

- *Tue. Nov. 15*: Imgur upgraded from a potential source to a confirmed source. After the removal of Facebook, we realized that we needed to further diversify our sources, and Imgur was at the top of the list, both in terms of user base and ease of implementation.