Hyrax Analytics Use Cases

Use Case 1: Administrative Dashboard

On the Administrative Dashboard, administrators should be able to get a general overview of the contents of the repository and any activity that has taken place recently. Much of the infrastructure for this is already in place in Hyrax, but squaring the requirements below with what is currently implemented is critical.

Requirements

- I can see graphics for the number of:
    - Collections (donut chart of published vs. unpublished, time chart for collections over time)
    - Works by Work Type (donut chart?)
    - Works by Resource Type (donut chart?)
    - Works by Visibility Type (donut chart?)
    - Files
    - Repository Growth (new items in the repository over time)
    - User activity (bar chart over time)
    - Visitors (boxes with numbers)
- I can see the number of collections and information about each collection including:
    - Number of works in the collection
    - Number of files in a the collection
    - Number of users interacting with the collection
- I can see graphics for the total number of queries run on the repository over time
- View graphics of visibility, in progress (ie in workflow), other such administrative content on a dashboard
- All widgets on this page (charts) should allow for setting a new date range for the widget

Questions

Use Case 2&3: Role-Based Reporting Dashboards

The reporting dashboard should be able to report usage statistics (views, downloads) in a granularity (admin set, time range, or a facet) specified by the user. The dashboard is role-based, allowing the user to query and report on information based on their role in the repository.

Requirements

As an administrator

- I can report out
    - the number of downloads based on filetype (mime type) or resource type.
    - the number of times my *collections* were accessed by month.
    - the number of unique downloads for objects in any given *collection*.
    - the total number of queries conducted for my *collections.*
- I can
    - use a calendar 'widget' to generate statistics (month, group of months, by fiscal year, etc.).
    - view the number of times my *collections* were accessed by month.
    - view the number of unique downloads for objects in any given *collection*.
- I can execute a search/browse or navigate to a collection (user or administrative) and, based on the results, download a file of pageviews and downloads for those items.
- I can execute a search/browse or navigate to a collection (user or administrative) and, based on the results, view a table and/or figure of pageviews and downloads for these items.
- I can save a search/report as part of my user profile.
- I can see the Top 100 most viewed works (show as table with thumbnails & number of views).
- I can see, on my user profile, pageviews, downloads, and any other available statistics about My Works.
- I can see Keyword and Phrase searches (e.g. fish (453) or stuart collection (36)).

As a Depositor of one or more works in the repository, I need access to statistics about the discovery and use of my works. Statistics of interest about each work include:

- Total number of times the page with my work has been visited.
- Number of times a page with my work has been visited over time.
- Number of unique visitors of my work.
- Number of times a page with my work has been visited by user.
- List of referring URLs.
- Geo-location of visitors (based on IP address or for privacy a partial IP address).
- Total number of downloads of a file in a work.
- Time series of downloads.

Questions

- Gabriela added Keyword and Phrase searches. Is that useful to anyone else?

Use Case 4: Item Level Statistics

At the work-level, statistics should be generated for pageviews and downloads that capture all activity on the work.

Requirements

- I can see a table and/or chart of monthly pageviews and downloads for the item.
- I can download the daily (or some other more granular timestep) pageviews and downloads for the item.
- As available, I can view other statistical information for the item (e.g. citations, tweets, mendeley saves, shares).
- Location of downloads (map, table of places/countries)
- On the showpage, indication pageviews and downloads

Questions

- Aggregation of statistics at the parent work?

Use Case 5: Analytics Management

downloads and pageviews - need to make sure this stuff is getting counted. also: counting other things like citations, etc.

Requirements

- Access analytics via a REST endpoint - json output of most viewed author, most downloaded, and other such information (to contribute to a larger scale reporting system - e.g. Mexico National, ARL, etc.)
- Integration of legacy statistics

Questions

- What is a download?
- Clarity on what makes a download event happen?
    - A view should not = a download

Use Case 6: Setup and Configuration of Analytics

The repository manager should be able to setup Google Analytics account and control its tracking scope in the dashboard.

Requirements

- enter the Google Analytics ID associated with my Google Analytics account so that repository usage activities are recorded in Google Analytics (Gary notes this will be added to general settings in Hyku) for my single tenant repository instance.
- enter the OAuth information associated with my Google Analytics account so that repository usage activities are displayed in the UI at single tenant repository instance,
- flip statistics and related UI should be so that a repository administrator can turn them off as needed.
- configure a page or pages on which analytics will not be collected. This is particularly useful when using external monitoring software that requests a page from the repository and compares it to a stored value to confirm the stack is working as expected.
- in hosted repository in a multi-tenant context, only access the usage data about the content in my repository, and not the usage data of other tenants of the service.
- Piwik integration should mirror Google Analytics integration yes!
- add Generating a widget or gem of external analytics tool useful (e.g. Chartkick, D3)

Questions

Use Case 7: Analytics Documentation and Best Practices

New repository administrators/creators could benefit from some documentation of how analytics work in Hyrax and what best practices the Samvera community has to offer on analytics platforms and sources to use and caveats.

Requirements

- Clarity of what aggregated statistics mean and how they are generated
- Clarity on what a pageview is and what a download is and how they are generated
- Documentation is needed for all analytics pages and features

Questions

Use Case 8: Analytics Reports delivered by email

Depositors – especially publishing faculty, researchers, students – would like to receive messages containing information and analytics about how their content is being access and used.

Requirements

- Send any user an email with consolidated analytics about their deposits ("your article was accessed 21 times this month, your data was downloaded 5 times this week")
- Send the depositor an email when a particular milestone has been reached ("congratulations! your dissertation has been viewed 100 times!"
- Provide the depositor the ability to configure their analytics notifications (probably in user profile)
    - turn on/off
    - frequency of notification (weekly, monthly, quarterly, etc.)

Questions

Issues/Uncategorized:

- I can embed the x most viewed and downloaded items in this collection on the collection landing page (push to/coordinate with Display Sets WG?)
- Where downloads of statistics data are downloadable, downloads are available on a monthly or daily basis.
- Export figures and tables from Hyrax (as pdf, csv)
- File download statistics should include "direct-link" visits like the bookmarked uri for a pdf that won't open a html page (it will impact Google Analytics particularly)