

Analytics: Design Questions

General questions

Q1: Clarity on relevant metrics

- Are these accurate definitions?
 - View = user lands on a show page for a collection, work, file
 - Download = user downloads object: file
 - Either through “Download image” action on a work or file show page, or the “Download” action associated with a file. Is there anything else that can be downloaded?
 - Access = is this simply a synonym for “view” or is it a unique metric, such as “views + downloads,” or something else?
 - Making it a unique metric is likely problematic

Q2: Aggregated metrics

- R10, R11, maybe others seem to suggest that views of a work include views of a child work or fileset of that work. For my benefit at least, I think we need to talk through aggregated metrics more thoroughly.
 - For example, a user could 1) view a (parent) work show page, 2) click on a child work of that work to view its show page, 3) click on a file to view the fileset show page.
 - How many views of the parent show page are there? 1 or 3?

Specific requirement questions

The “Rx” in the headings below refers to the specific row of each requirement in the [requirements spreadsheet](#).

R2: Dashboard - Graphics for works by type, etc

- Distinction between Work Type and Resource Type needs to be clarified, but assuming these are two different things that are both of interest to repo admins, this is enough to start design.

R3: Dashboard - Table of number of Works by Collection

- Should the focus of the table really be “number of works”? If a repo admin wants to know how many works various collections have, I’d suggest that is information better found in a report.
 - A table of **largest collections**, showing the five or so collections with the most works, would be appropriate for the dashboard, if that information would be useful for repo admins.
- Alternatively (instead of or in addition to the “largest collections” idea above), a focus for a dashboard widget related to collections could simply be letting the repo admin know

which collections have seen the **most recent activity**. Such a table could easily include a column with the number of works in the displayed collections, but the purpose for the table on the dashboard would be to inform about recent activity, not counts of works.

R4: Dashboard - Graphic widgets allow for modifiable date ranges

- Since the goal of a dashboard is to show current and recent activity, I assume that date ranges should only be modifiable in terms of how far back the date window goes (e.g., default is 30 days but could be modified to be 60 days, 7 days, etc.).
 - If a repo admin wants to see the data for a dashboard widget for some specific date range, a report would be more appropriate.
- Enabling unique date ranges for each dashboard widget is going to make the UI cluttered and less usable. Another approach might be to provide a single control that adjusts the date range for all dashboard widgets. If it is easy to change (e.g., a dropdown menu with [1 day | 7 days | 30 days | 90 days] options), the repo admin could pretty quickly change it to view specific widgets with the desired range.
 - In other words, this would be a UI control, not a configuration setting.

R5: Report - number of downloads and pageviews by collection

- View for a collection = when the collection show page is viewed?
- Download for a collection = ?
 - You can download files that are in works that are in a collection, but you can't download a collection, can you?
 - So maybe this means, total downloads of all files in all works in this collection?
 - If so, does the work have to be accessed via the collection to count?
 - A file could exist in more than one work (I think?), and a work could exist in more than one collection (depending on collection configuration) so if we don't consider the context in which the work is accessed, just counting file downloads and assigning that to the downloads for a given collection could be misleading.

R6: Report - number of times collections were accessed by month

- What does "accessed" mean here?
 - Is it views? So a count of the total views of a given collection's show page?
 - If so, how is it different from R5? Or is it not different, just that this requirement says also show the counts aggregated by month?

R7: Report - use a calendar 'widget' to generate statistics (month, group of months, by fiscal year, etc.)

- I interpret this as simply meaning, for the reports described by R5, R8, R9, R10, R11, R12, and R13, the user should be able to control the time period for the which the statistics are generated.

R8: Report - number of times my collections were accessed by month

- The requirement title is a duplicate of R6. Is it a duplicate or something different?
- Based on the additional description here, I'm more confused about what "accessed" means, because the description here says "accessed = pageviews of works in the collection."
 - So rather than how many times the collection show page was viewed, it is the sum total of the times the show page of a work in this collection was viewed?
 - If so, this seems problematic, given a work can exist in multiple collections (based on configuration choices).
 - Or maybe the idea is "accessed = total number of times a show page for a work in this collection is viewed via this collection"?

R9: Report - number of unique downloads for objects in any given collection

- "objects" = files? Nothing else can be downloaded, correct?

R10: Report - number of times the page with my work has been viewed (pageviews)

- Views of a work show page, correct?
- Although it's not stated here, I assume the user would want to see a table of all of their works, with the view counts.

R11: Report - Timeseries graphic showing number of times a page with my work has been viewed over time

- This requirement is just supplementing R10 with a timeseries graphic for a selected work, correct?

R12: Report - Number of unique visitors of my work

- Distinct from, but would be presented similar to R10

R13: Report - number of downloads of a file in a work

- "a file in a work" - should this be "the total downloads of all files" in a work?
 - Or do you really want the report to show individual file downloads per work, in which case we'd need to nest each file under each work? Doable, I'm just trying to clarify whether the interest is in "downloads per file within the context of the work", or just downloads per file. If it's the latter, it might make more sense to have a separate "files" report.

R14: Item Statistics - At the work-level, statistics should be generated for pageviews and downloads that capture all activity on the work

- Clear, other than my concern about aggregating views that I described in Q2 earlier.

R15: Item Statistics - table and/or chart of monthly pageviews and downloads for the item

- How many months back do we try to show?
 - Not relevant until a repo has been in operation for more than a year so we could ignore for now, but in the long-term this might need to be something like a) the months of the current year, plus yearly totals for past years, or b) the months of the current year by default, with a dropdown menu to replace the year being viewed with a previous year.

R16: Item Statistics - On the showpage, indicate pageviews and downloads

- How is this different from R14?

R17: Item Statistics - Analytics at the item level are captured in such a way that allows for aggregation at all higher levels

- I guess this is not a UI concern, but relates to the aggregation issue.