

Use Case 4.2: Leverage the deeper graph to surface more relevant works

Example Story: As a researcher, I would like to see resources in response to a search where the relevance ranking of the results reflects the "importance" of the works, based on how they have been used or selected by others, so that I can find important resources that might otherwise be "hidden" in a large set of results.

In this use case the importance calculated will reflect importance in the scholarly world and will be different from those in commodity systems, as well as including items that would not appear in commodity systems (e.g. manuscripts). A benchmark will be doing better than Amazon with richer results.

Out of scope: n/a

Potential Demonstrations

- A. Do a "page-rank" style algorithm (centrality, minimum path, etc) across the full linked data graph, assigning appropriate weights to certain kinds of annotations and relationships and reflecting those weights in the relevance ranking of search results for a set of common queries.
- B. Possibly allow (for demo purposes) the user to see the comparative results s/he would have gotten from Amazon. (Added by DavidW, not discussed by group)
- C. Boost the ranking of any resource that has external relationship links by a computation over those relationships.

Data Sources

- Annotations on resources
- Scrape references from LibGuides
- Any relationships between resources or links to them in the broader linked-data web
- Usage data of any sort (StackScore or similar would be a start)

Ontology Requirements

- None new beyond ability to include annotations, references, relationships and usage data as described data sources

Engineering Work

- Understanding of works and granularity at which links between items should be understood to understand importance
- Deal with questions of how to work in a rather data sparse environment -- how to merge information from page-rank like ordering with some other order that doesn't rely on a dense graph
 - Will need to do iterative/experimental work on what will influence rank. Possibilities include:
 - e.g., a Libguide that cites a work -- ideally with an OCLC number or DOI
 - could scrape links into the catalog
 - usage data
 - external ranking
- Might use various axes of similarity: Example -- Griffin Weber's analysis that is performed nightly and stored in the triple store in a different namespace -- geographic proximity, association with MeSH terms calculated from occurrence in publications

Who will do what?

- See [Use Case 4.1: Identifying related works](#)