

# Week 10: Social Search

---

## Assignment 1: Tagging

*a) Describe how social media tags are similar/different to anchor text (e.g. with respect to creation, application to search, etc.).*

Similarities:

- combine some words with a resource
- queries might use same words when searching for this resource
- can be noisy

Differences:

- tags are cheaper,
- tags have less quality,
- tags have high ambiguity.

*b) Which types of tags are most valuable for search, which most useless?*

Most valuable:

- Content-based
- Context-based

Most useless;

- Subjective
- Organizational

## Assignment 2: Communities

*a) Explain **briefly** how the HITS algorithm can be used to find a community of bloggers about "tennis"!*

1. Calculate authority and hub scores for all websites
2. Find a blog with a high authority, which posts about tennis
3. Take all hubs with high authority, which link to this blog.
4. Take all blogs with high authority, which are linked to by these hubs.

*b) How would you do it using PageRank?*

1. Do a topic-sensitive PageRank about tennis,
  2. Take all documents with a PageRank greater than a certain threshold.
- (does not distinguish authority and hub scores)

### Assignment 3: Profiles

*New users of a system usually don't have a history and therefore empty profiles. How can this so-called **cold-start problem** be solved?*

*a) Explain briefly **one** possible solution to deal with new users in a filtering setting.*

Find users with similar provided user information and apply: average of their constraints, most used queries, or most occurred relevant/non-relevant queries.