

- ① CRAWL the web ...
- ② Efficiently index the documents (D)
- ③ search for phrase P and filter documents containing P , (D_P).
- ④ Rank order in terms of importance (D_P^{Rank}).

Among all the web documents that contain P , which ones are most reputed
... which ones are more visited by people.

Reputed pages are ones that lot of pages point to (e.g., nyt.com)
Not so reputed pages don't have many links pointing to them.

So let's model
the "who points
to whom" pattern.

