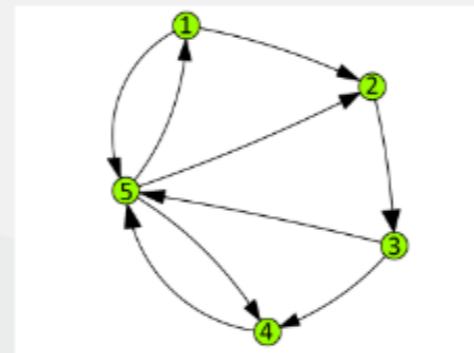


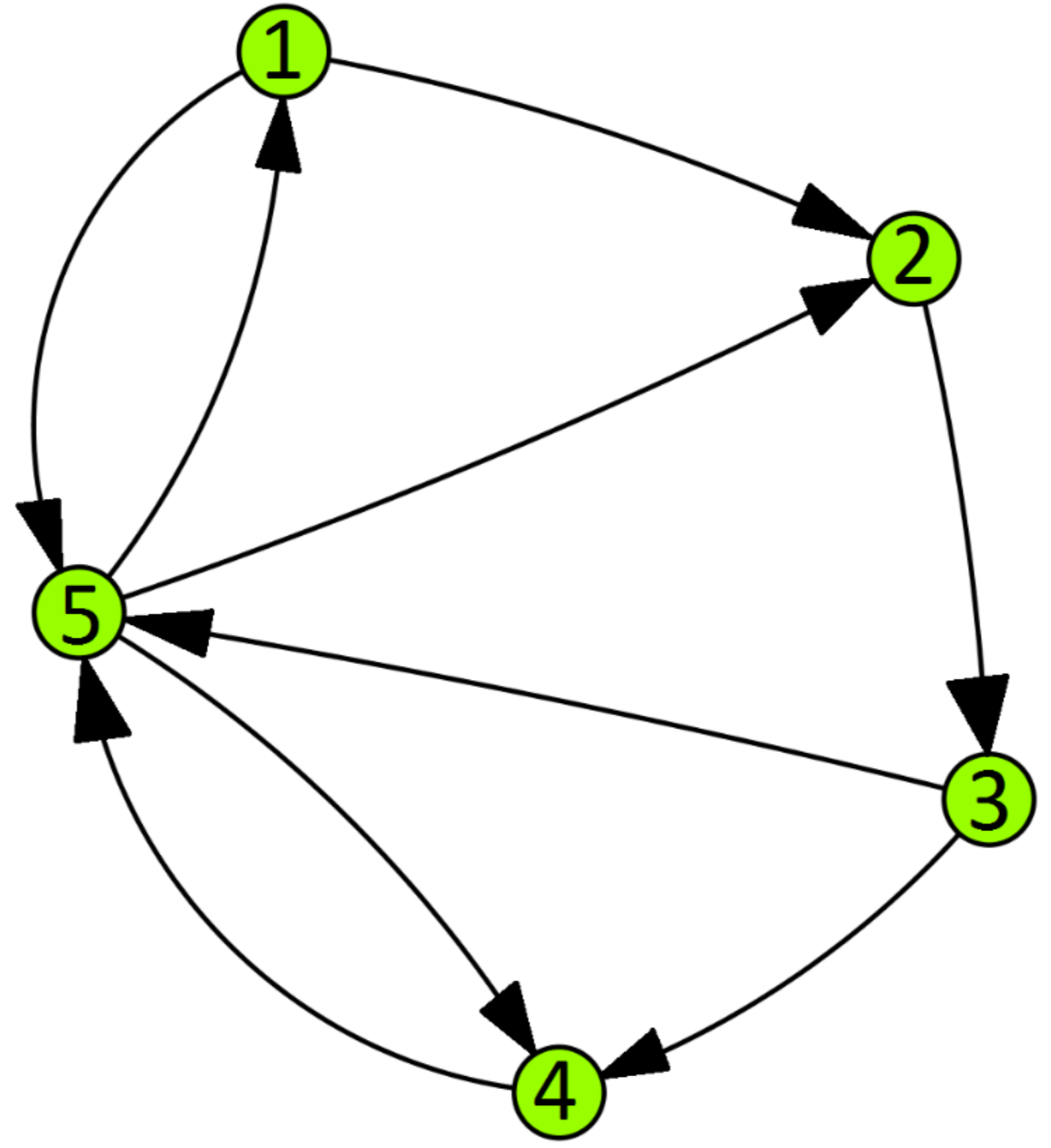
Mathematik hinter Suchmaschinen

Isabelle A., Elisabeth B., Klara B., Michael G.,
Marlene H., Anna M., Elisabeth P., Anna S.,
Lukas S., Serhat T., Julian W.

Beispiel:

- Websites und Links
- Wichtigkeit:
Anzahl der eingehenden und ausgehenden Links



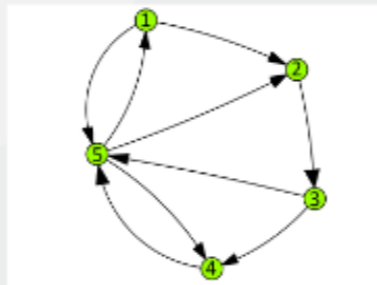


Überblick

- Verbindung zwischen Webseiten
- Wichtigkeit der Seiten
- Test in der Praxis

Random Walk

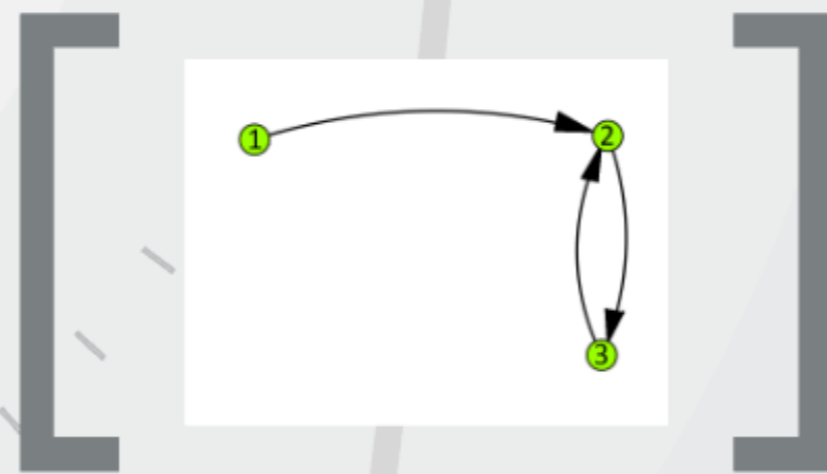
- Methode zur Bestimmung der Wichtigkeit
- = zufälliges Gehen durch das Netz
- häufiges Verlinkt -> wichtig



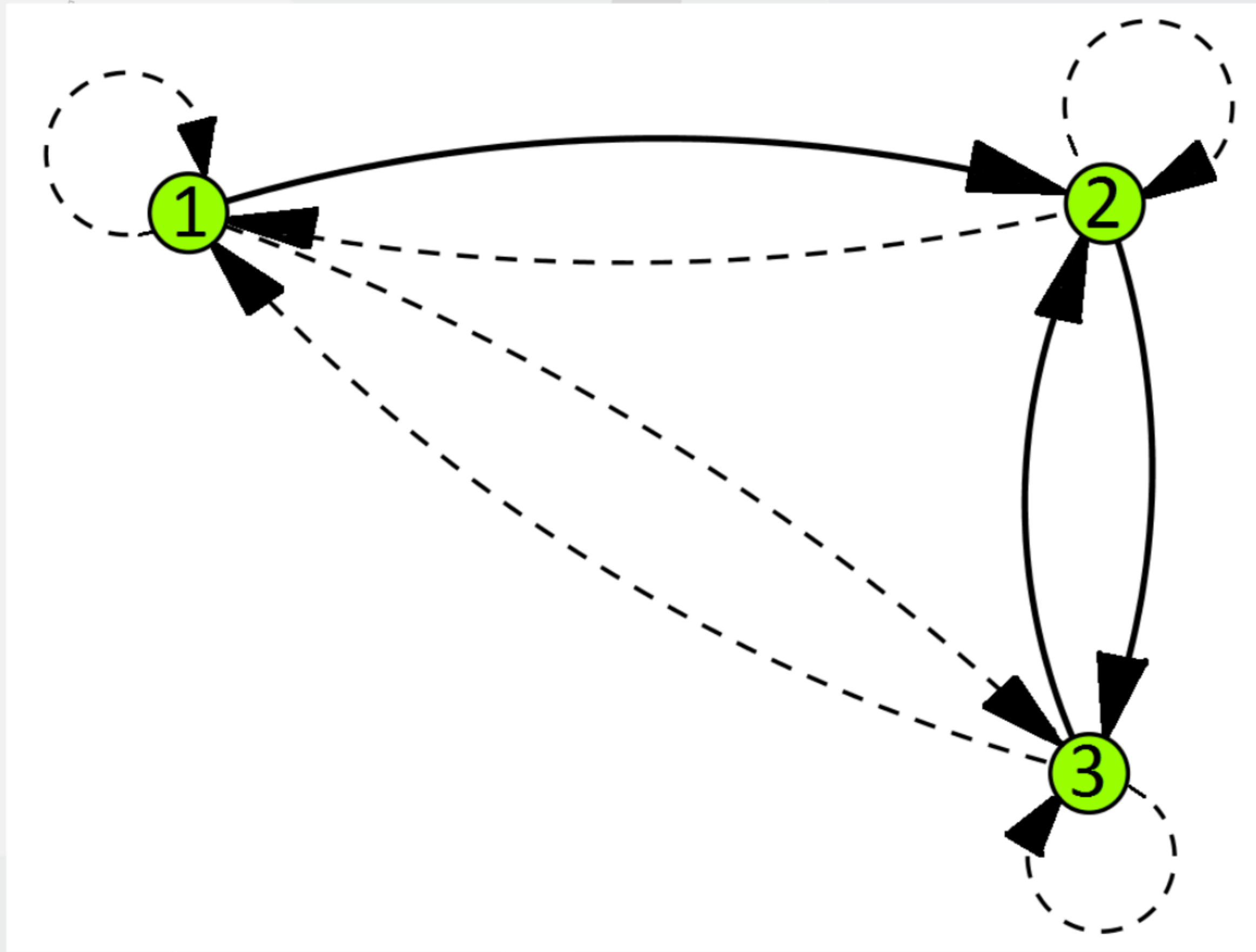
Teleport-Variable

zufälliges Gehen:

- Sackgassen?
- Kreise?





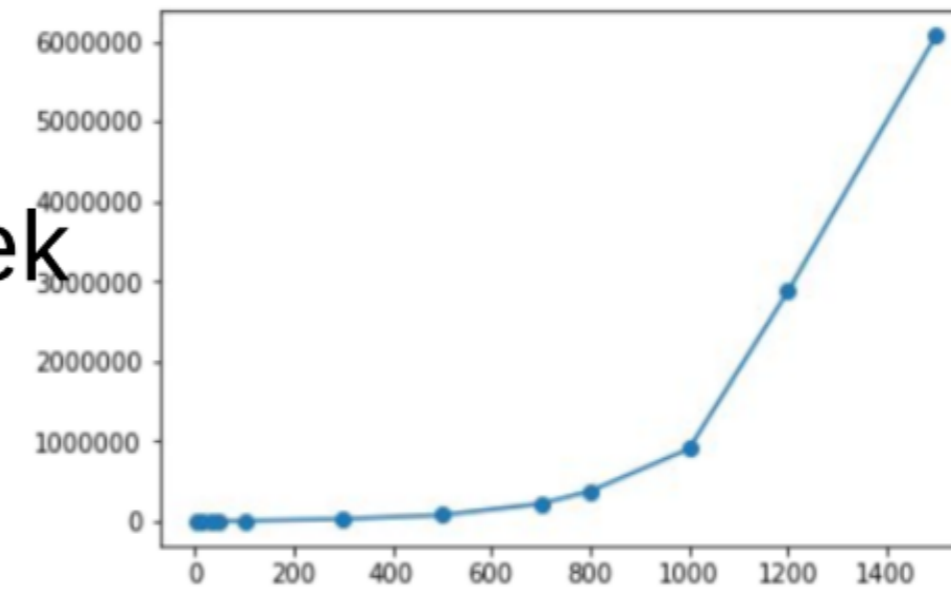


Rechenzeit

•Google kennt 130 Billionen Seiten

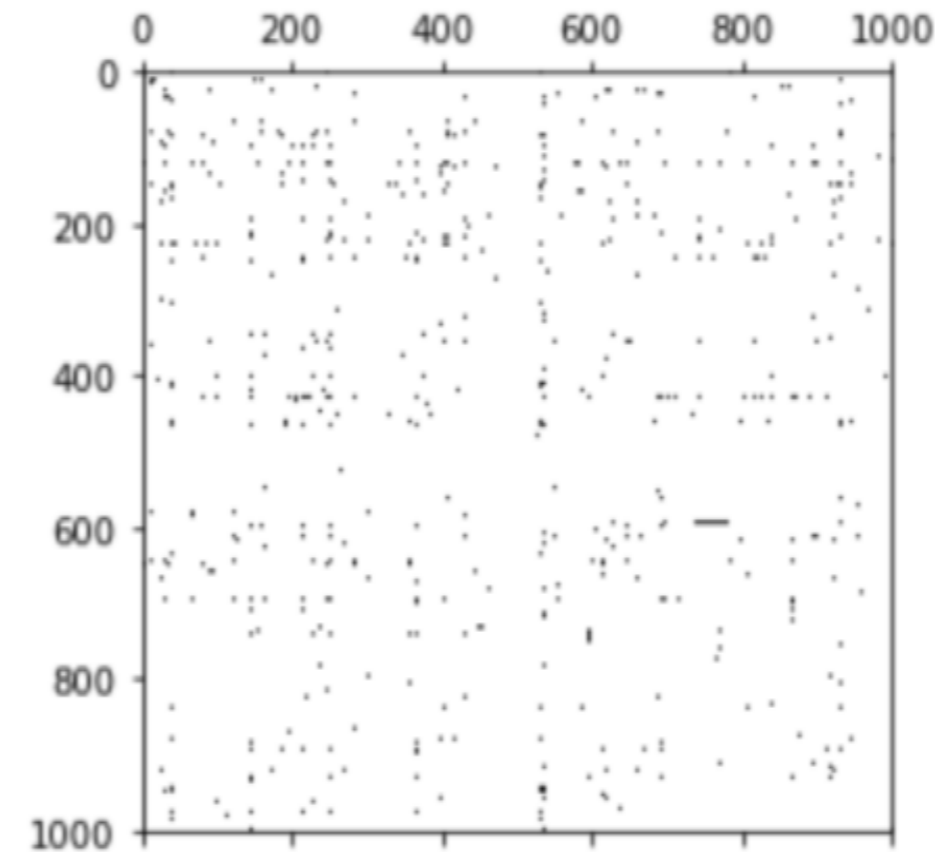
•Naiv ganze Matrix berechnen:

- $130 \text{ Bio} * 130 \text{ Bio}$
- $1.000 * 1.000$ braucht 1 sek
- Quadratische Laufzeit
- ca. 300 Milliarden Jahre
- Universum 4,7 Milliarden alt



Trick

- Jede Seite wenige Links
- Matrix Großteils 0
- Diese nicht berechnen
- Jetzt lineare Laufzeit
- 130 Billionen Seiten
→ 1 Tag
- Und nicht 300 Milliarden Jahre



Simple English Wikipedia

- 150.000 Seiten
- 2.200.000 Links
- Alle heruntergeladen (8 MB)
- Berechnet in 1s
- Dank Optimierung

Resultate

- Auf 100 wichtigsten Seiten
 - Nr. 1: USA
 - 38 geografische Artikel
 - Davon 21 Staaten
 - Sonst sehr allgemeines
 - Nr. 16: Tier
 - Nr. 17: Biologie
 - Nr. 18: Mensch
- 22.500 Seiten ohne Verlinkung

