


Search and Find in Scientific Literature

Richard Charkin
Macmillan Publishers

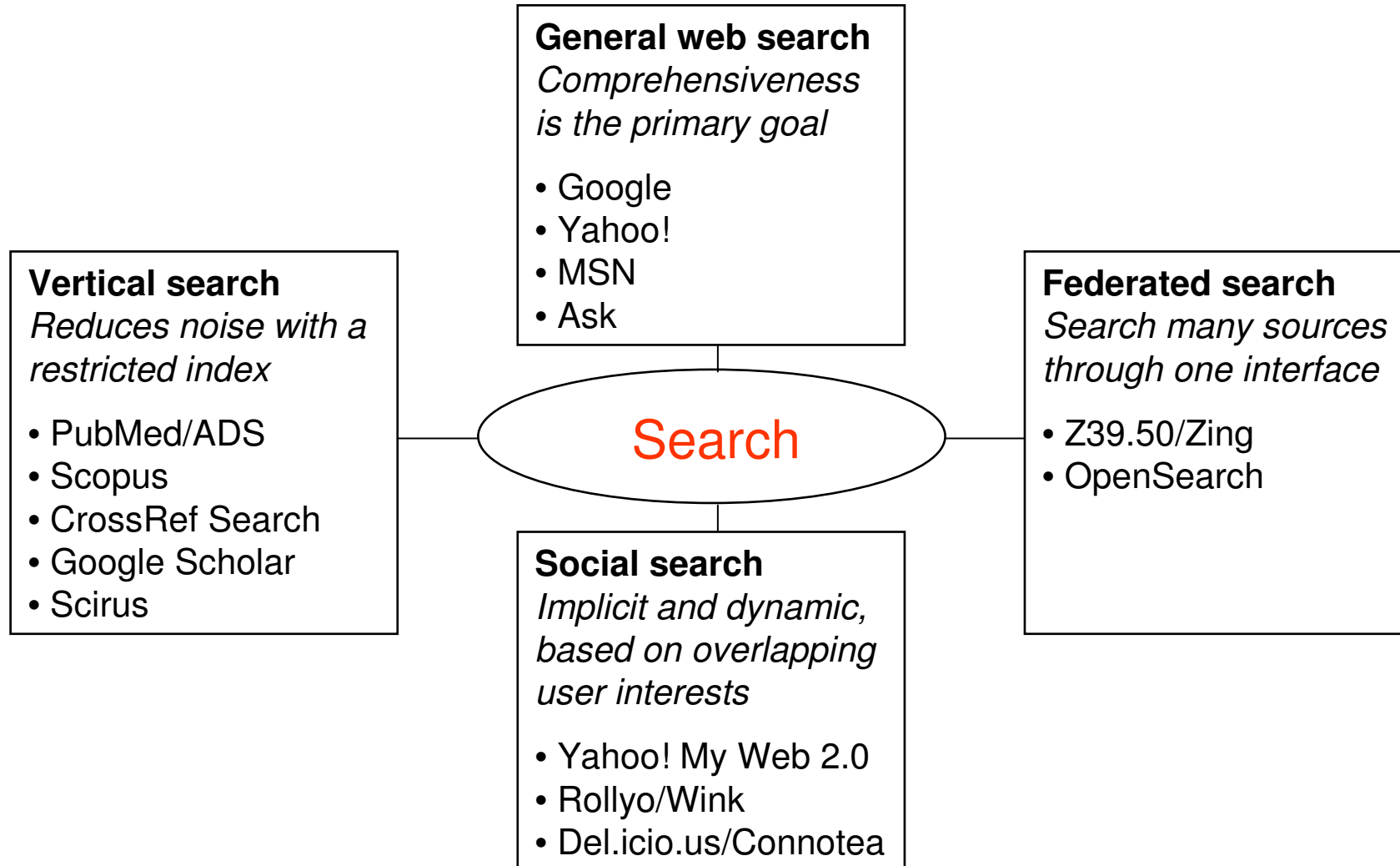
Web search is now fundamental to science

- Ten years ago scientists accessed literature in libraries, now they use web browsers
- Well over 60% of external referrals to Nature.com come from search engines
- Search is the most heavily used feature on Nature.com



The ability of scientists to search literature and data sets has already revolutionised the research process

Search is multifaceted



Current challenges in scientific search

- **Inconsistent naming:** One gene can have 20+ names, some of them also common words
- **Fuzzy definition:** Even core scientific concepts like 'gene' can be ill-defined
- **Inconsistent Formats:** Commonly used data standards are mostly non-existent
- **Image and video Data:** This kind of information is intrinsically difficult to index and search



Some partial solutions

- Standardised naming schemes and ontologies (e.g., IUPHAR, GO)
- Tagging and 'folksonomies' (bottom-up metadata and ontologies)
- Steadily improving text-mining algorithms Metadata made more interoperable with Semantic Web technologies like RDF

nature.com
the world's best science on your desktop

Thank you