

Medicine Digital Libraries: How To Cure Information Overload by Modern Methods of Artificial Intelligence

Zdenko Staníček Filip Procházka
Jiří Šmerda Vladimír Dosoudil Radka Findeisová

Masaryk University, Institute of Computer Science

Masaryk University, Faculty of Informatics,
Knowledge and Information Robots Laboratory

21th November 2007



Outline

- 1 Introduction
 - Digital Libraries Definition
- 2 Motivation
 - First Motivating Use Case
 - Information Retrieval Process
 - Another Motivating Use Case
 - Solution Today
- 3 Our approach
 - Our Goal
 - Present Work
 - Future Work
- 4 Summary
- 5 Discussion



Definition According Digital Library Federation

Definition (Digital Libraries)

„Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.“

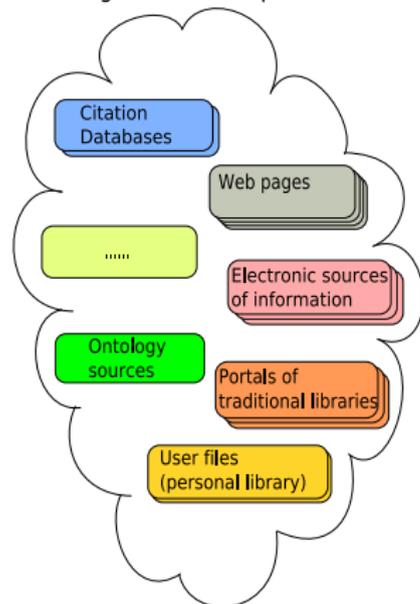


Renowned Authors and High Impact Journals

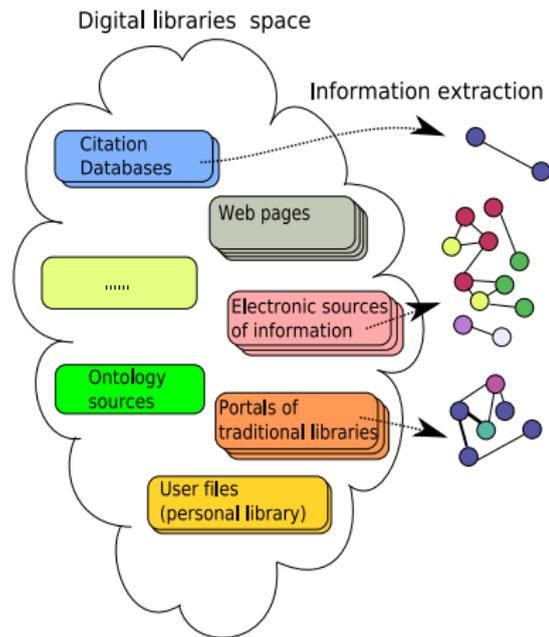
- I'm a surgeon and I do research in heart transplants.
- The problem is that there are many digital libraries and many papers concerning the heart transplants and I have no time to explore all of them.
- I wonder what renowned authors have published about heart transplants in high impact journals, I'm especially interested in relevant papers, both electronic and printed versions.

Information Retrieval Process Illustration

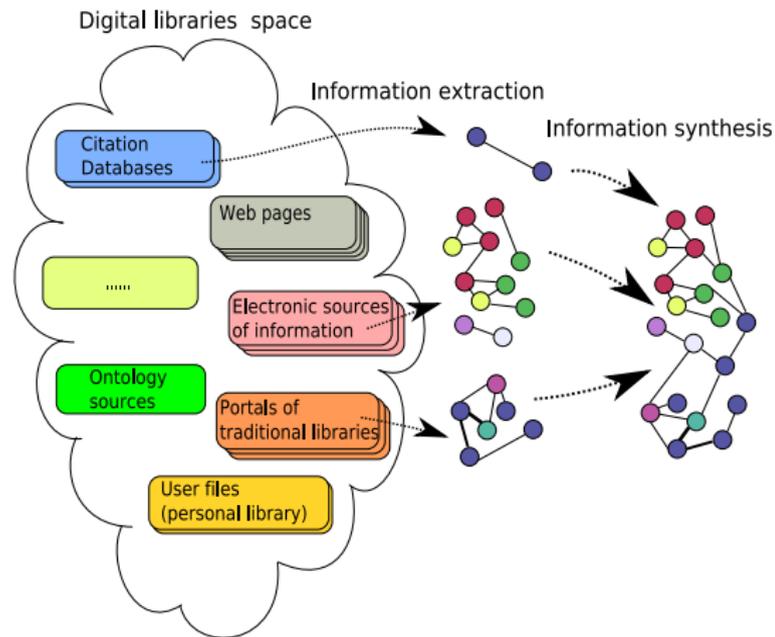
Digital libraries space



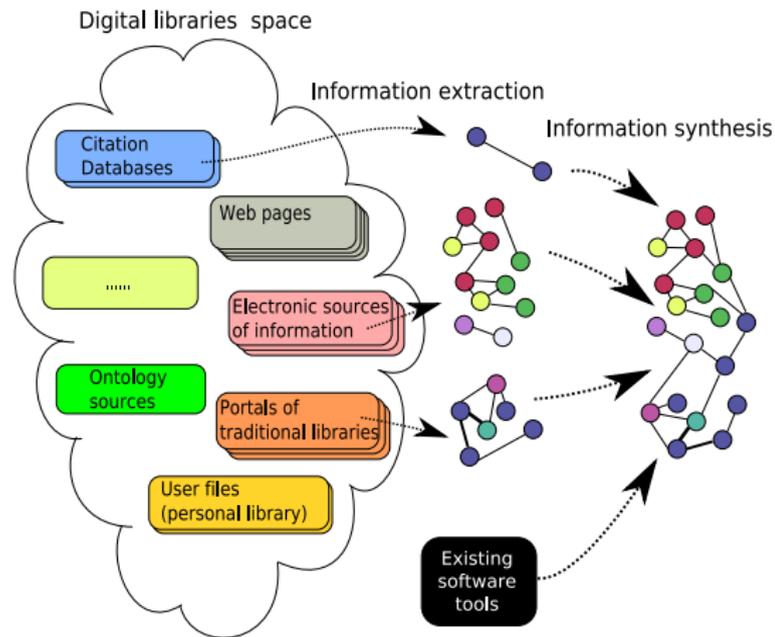
Information Retrieval Process Illustration



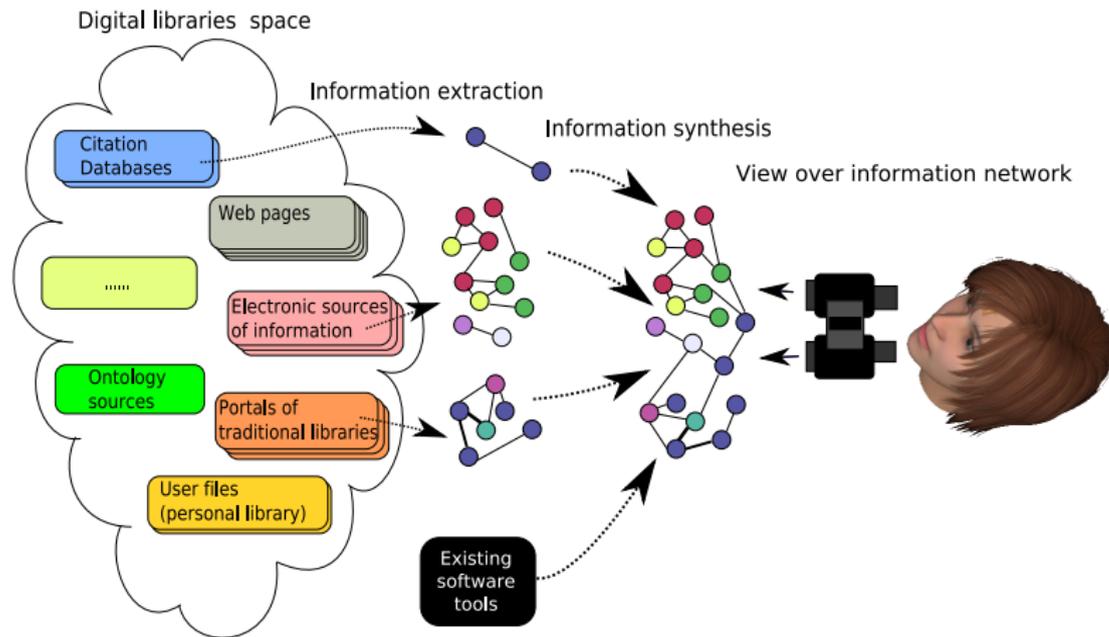
Information Retrieval Process Illustration



Information Retrieval Process Illustration



Information Retrieval Process Illustration



Accessible Paper

- I've got a reference to a paper located in a collection which is not bought by my institution.
- I need to find the paper in accessible collections.



Solution Today

We are not able to satisfactorily solve the mentioned use cases today.

- 1 *Ease of Access*: data are available, but not easily available
- 2 *Information Capability*: information capability of a union of all data sources is more than a union of information capabilities of each data source
 - $IC(DS_1 \cup DS_2) \supseteq IC(DS_1) \cup IC(DS_2)$

Our Goal

Our goal is to develop a scope into the digital library space which will support the information retrieval process by

- information extraction from heterogenous data sources
- intelligent information synthesis
- work with uncertain and context dependent information
- use of pattern matching
- advanced visualization methods

All of them will help users to focus on important things and filter the unimportant.



Present Work

- we focus on metadata
 - bibliographic records
- we use the Knowledge and Information Robots technology
- the application is under development at Institute of Computer Science, MU
- we are able to measure data sources utilization
- we would like to involve various users in a process of testing the application - alfa testers are welcome



Future Work

- personal libraries
- user adaptation
- digital libraries monitoring
- work with ontologies
- work with fulltexts
- include other data sources



Summary

- we have introduced the motivating use cases and the information retrieval process
- we have presented our approach with Knowledge and Information Robots technology and pointed out major issues



Discussion

Thank you for your attention!

Discussion:

- Have you encountered similar problems in your work which could be solved by our approach?
- Do you have ideas of other use cases?
- Do you consider our approach to be meaningful?
- Where do you see the major difficulties?

Contact: Jiří Šmerda - smerda@ics.muni.cz

