



When users do not know what they are looking for...
It is more challenging searching a full text.

Do you prefer?



Natural Language Processing for Information Retrieval

David D. Lewis & Karen Sparck Jones

Presentation by
Ismail Uyanik
Deniz Kerimoglu





Document Retrieval

DR involves retrieving texts for humans to read.

Retrieval depends on indexing.

Indexing poses constraints:

- Variability
- Underspecification
- Reduction





File View Control Help

Hmmm.. Let's find out
something about
cheap production
methods for simple
prefabricated housing.





IR is a hard task!



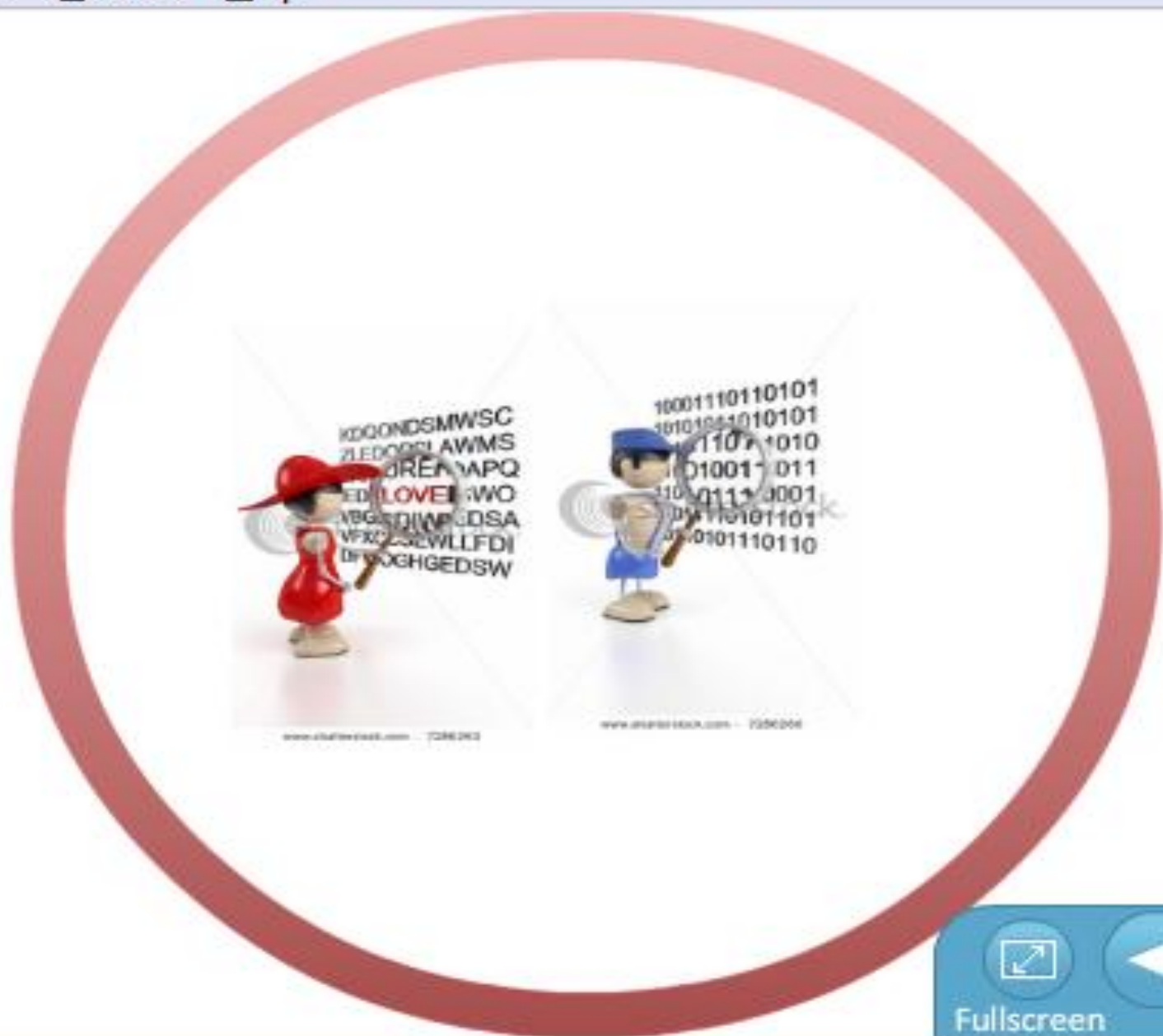
INFORMATION RETRIEVAL

Harder when users do not know what they are really looking for...

... and more challenging when searching a full text directly.



File View Control Help





Which one do you prefer?

SEARCH MODE ON
PUBLICATION-YEAR **GT** 2000 **AND**
PUBLISHER **EQ** MCGRAWHILL
SEARCH MODE OFF

I want to read all books published by
McGrawhill in 2000 and afterwards,
please.



END USERS PREFER THE LATTER.





Proposed Research Agenda

Future DR systems should include simple flexible natural-language indexing forms.

Such approaches allow and encourage users to concentrate on request development.





And one more thing...

It is essential to test, test, and test again

Conclusion

Although DR makes use of controlled indexing languages, indexing increasingly involves terms drawn from the natural language.

