

# An application-oriented terminology evaluation

the case of back-of-the-book indexes

**Touria Aït El Mekki**

LERIA

Université d'Angers

France

*touria at info.univ-angers.fr*

**Adeline Nazarenko**

LIPN - UMR 7030

Université Paris13 & CNRS

France

*adeline.nazarenko@lipn.univ-paris13.fr*

# Outline

## Introduction

1. The IndDoc indexing tool
2. The evaluation challenge
3. Our evaluation approach

## Conclusion

# BACK-OF-THE-BOOK INDEXES

## Exemple

### Index

<i>Acquisition</i>	7, 27, 78-81, 250
<i>AI</i> <i>see Artificial intelligence</i>	
<i>Artificial Intelligence</i>	43, 97, 134
<i>Knowledge</i>	<b>26-32</b> , 76-77, 89, 154
<i>acquisition (see also Acquisition)</i>	<b>211-215</b> , 228
<i>representation of</i>	25-29, 80, <b>132-136</b> , 250

### Structure nomenclature

List of relevant descriptors  
Semantic relations  
Cross-references

### References to textual segments

Segments of various sizes  
Underlined references

# EXISTING TOOLS

## Index management tools

L'expression de la *cause* va du lexique à la syntaxe. Dans certains cas, c'est en effet le lexique qui traduit la cause qu'il soit nominal (raison, motif, facteur...), adjectival (responsable de, fortuit, délibéré) ou verbal (provoquer, entraîner...) : nous avons donné dans le chapitre IV un aperçu de l'étendue et de la richesse du *lexique causal*. Dans d'autres cas, la *relation causale* est inscrite dans la *structure syntaxique* : c'est notamment le cas de la *subordination*, avec les *subordonnées circonstancielles*, des *participiales*, des *infinitives*, certains types de *complétives* et les *relatives explicatives*. Relevant à la fois du lexique et de la syntaxe, on trouve des *expressions lexicalisées* qui fonctionnent comme des *connecteurs* (en raison du fait que, par la faute de, étant entendu que...) mais aussi les constructions *causatives* (faire faire quelque chose à quelqu'un) qui explicitent le rôle du



*Causateur*, 6, 79-82, 7, 78,, 7, 80,  
81-82, 6, 79-82, 7, 78,, 7, 80, 81-82  
*cause*, 71r, 79i, 18,, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78,, 7, 80, 81-82  
*comme*, 44-46, 52, 59, 71, 79i, 18,, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78,, 7, 80, 81-82

*Consécutif*  
*Voir Conséquence*  
*Conséquence*, 9-15, 28, 57, 62  
*Inférence*, 11, 14, 21  
*Contrainte*  
*Voir Temps*  
*Contrefactif*  
*Voir Contrefactuel*



*Causateur*, 6, 79-82, 7, 78,, 7, 80,  
*cause*, 71r, 79i, 18,, 45, 58-59, 63, 64, 65, 67, 25,  
*comme*, 44-46, 52, 59, 71, 79i, 18,, 45, 58-59, 63, 64, 65,  
67, 25,  
28, 6, 79-82, 7, 78,, 7, 80, 81-82

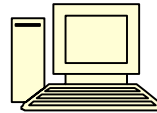
*Consécutif*  
*Voir Conséquence*  
*Conséquence*, 9-15, 28, 57, 62  
*Inférence*, 11, 14, 21  
*Contrainte*  
*Voir Temps*  
*Contrefactif*  
*Voir Contrefactuel*

### Indexer



- Entry markup
- Entry and subentry distinction

### (Word, Latex, Cindex)



- Index page design
- Occurrence lookup
- Alphabetical sorting

### Indexer



- Occurrence selection
- Occurrence clustering
- Index restructuring

➤ The form but not the content !

# EXISTING TOOLS

## Index acquisition tools

### Document structure (titles)

HTML Indexer, IXgen



*Causateur*, 6, 79-82, 7, 78, 7, 80,  
81-82, 6, 79-82, 7, 78, 7, 80, 81-82  
*cause*, 71, 79, 18, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78, 7, 80, 81-82  
*comme*, 44-46, 52, 59, 71, 79, 18, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78, 7, 80, 81-82  
*Conséquence*, 9-15, 28, 57, 62  
*Inférence*, 11, 14, 21

Few entries

References to too many occurrences

No structuration

### Linguistic analysis (NPs)

Syntactica, Indexing Online

*Causateur*, 6, 79-82, 7, 78, 7, 80,  
81-82, 6, 79-82, 7, 78, 7, 80, 81-82  
*cause*, 71, 79, 18, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78, 7, 80, 81-82  
*comme*, 44-46, 52, 59, 71, 79, 18, 45, 58-59, 63, 64, 65, 67, 25,  
28, 6, 79-82, 7, 78, 7, 80, 81-82  
*Conséquence*, 9-15, 28, 57, 62  
*Inférence*, 11, 14, 21

Too many entries

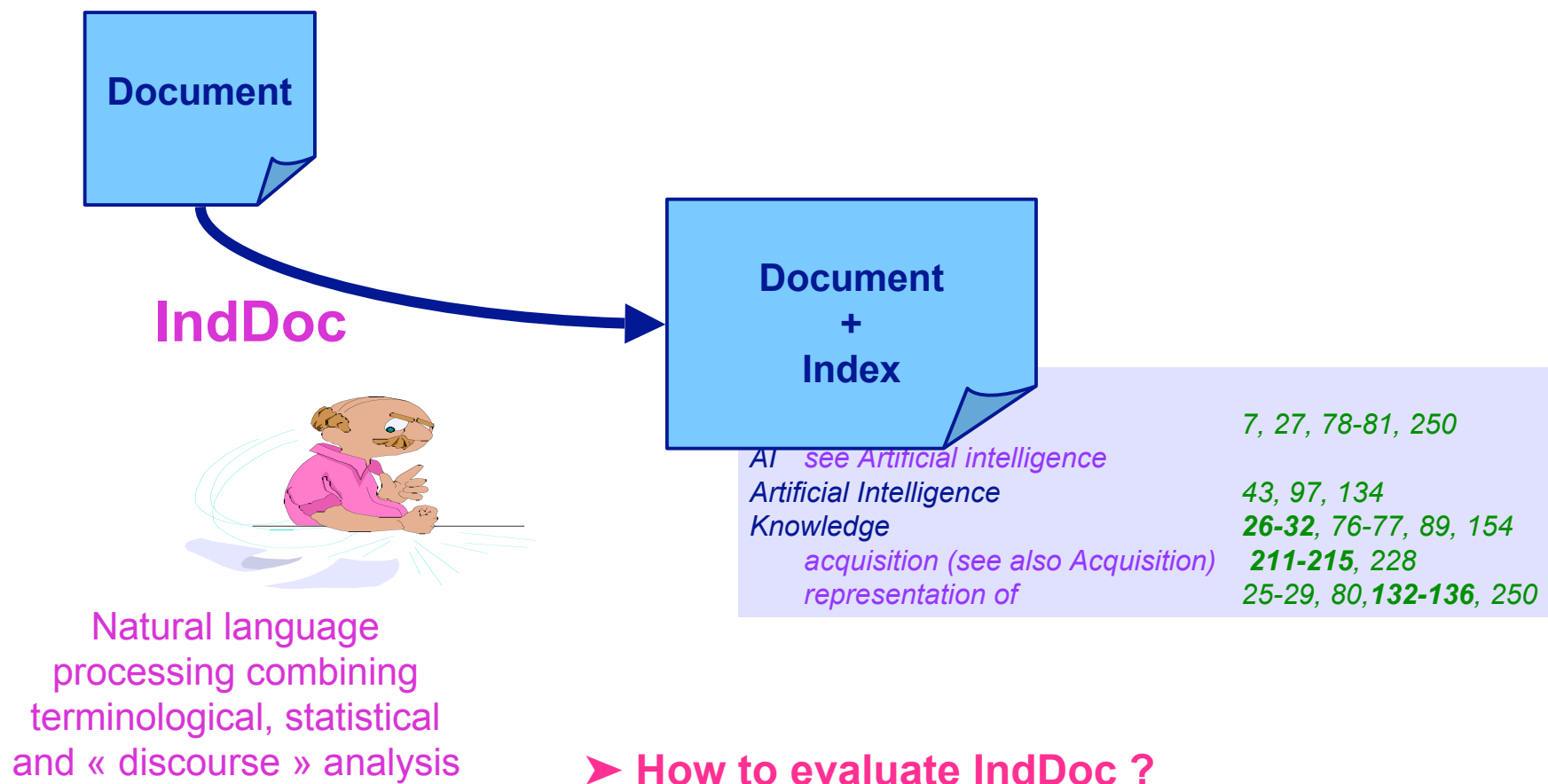
References to too many occurrences

No structuration

➤ How to select the relevant information ?

# ASSISTING THE INDEXING WORK

## The IndDoc system



# Outline

Introduction

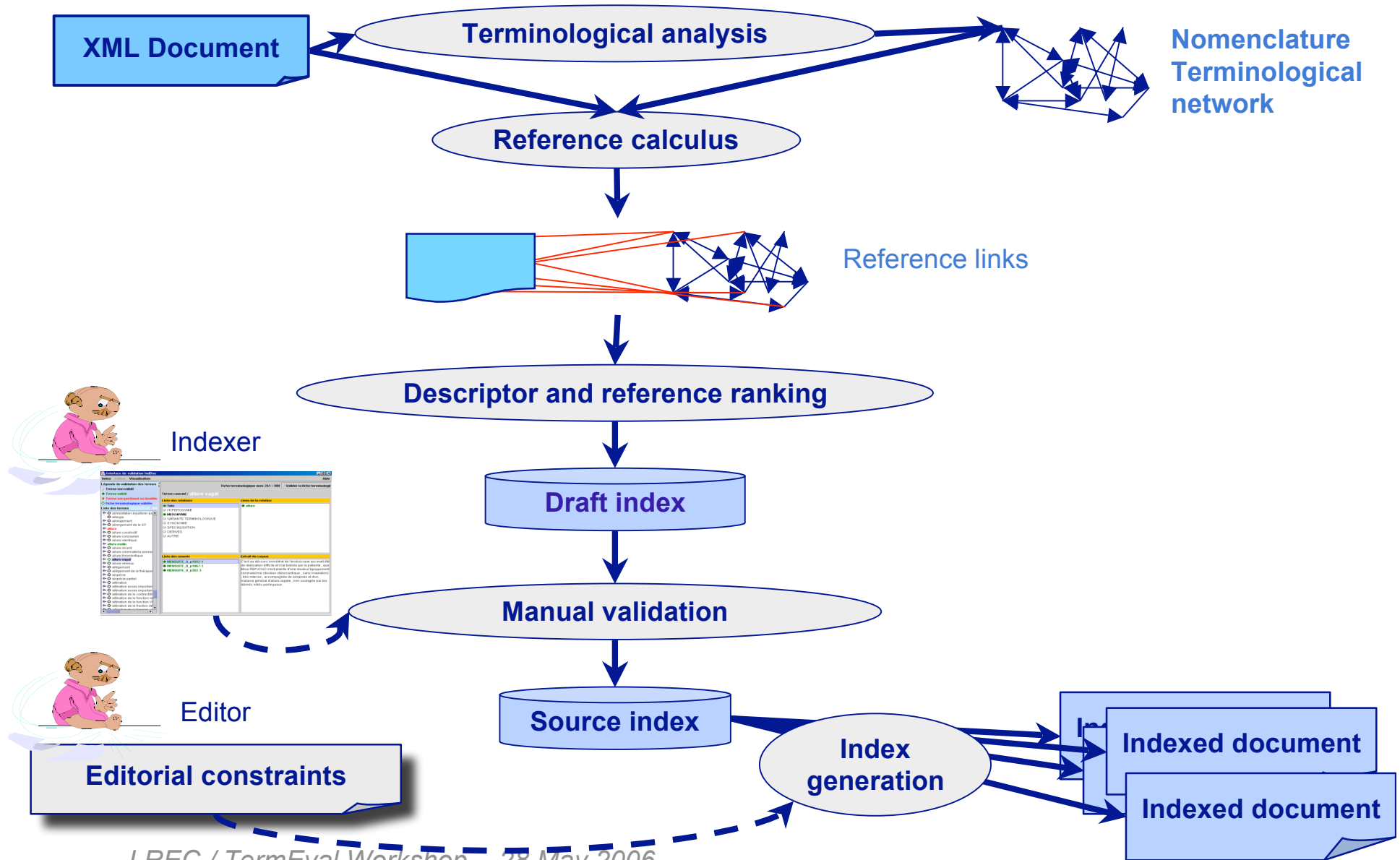
**1. The IndDoc indexing tool**

2. The evaluation challenge

3. Our evaluation approach

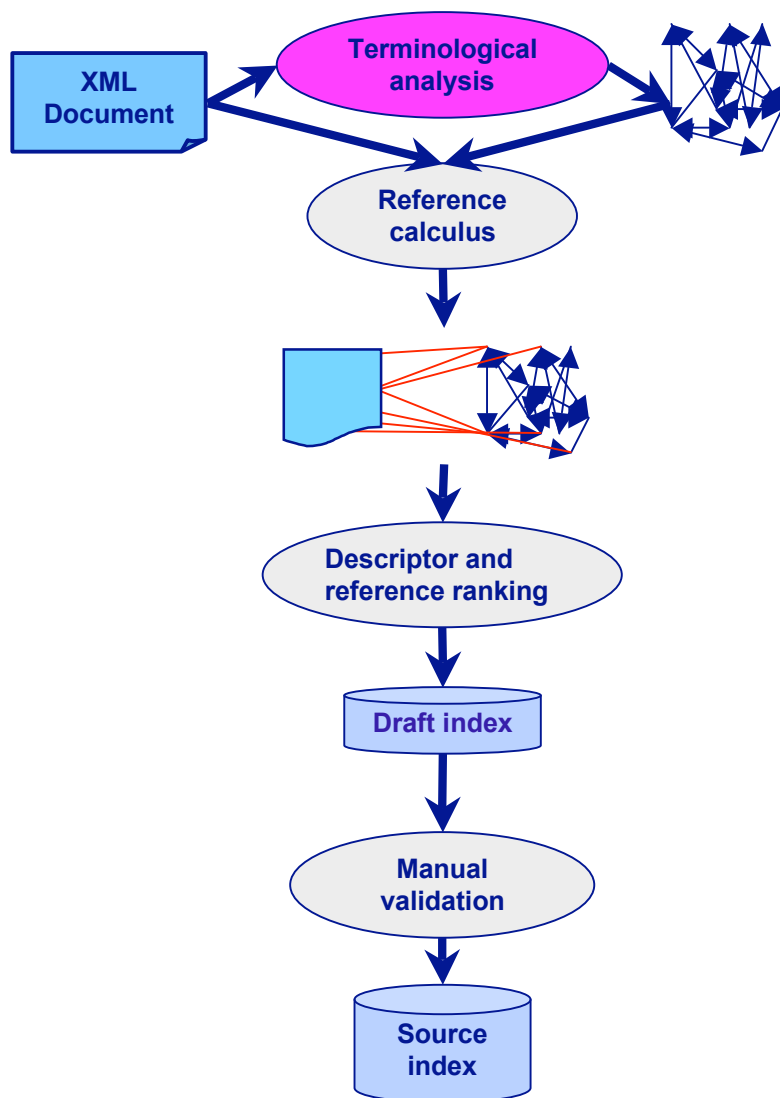
Conclusion

# INDDOC ARCHITECTURE





# TERMINOLOGICAL ANALYSIS

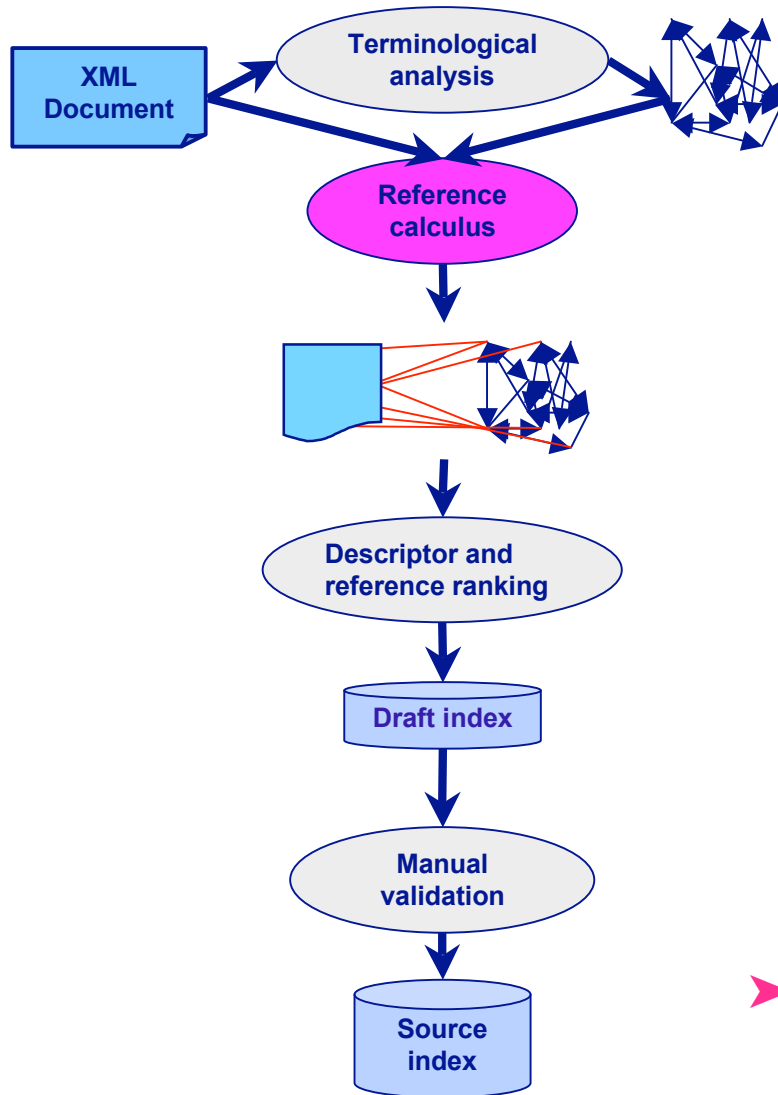


- Term extraction
- Morphological analysis (variant forms)
- Identification of semantic links
  - Syntactic internal structure
  - Pattern-based recognition
  - Dictionary-based synonymy detection
- No validation

*book / books*  
*anti-obesity prescription / prescription for obesity*  
*clinic / hospital*  
*hospital / children hospital*

➤ **A network of terms**

# REFERENCE CALCULUS

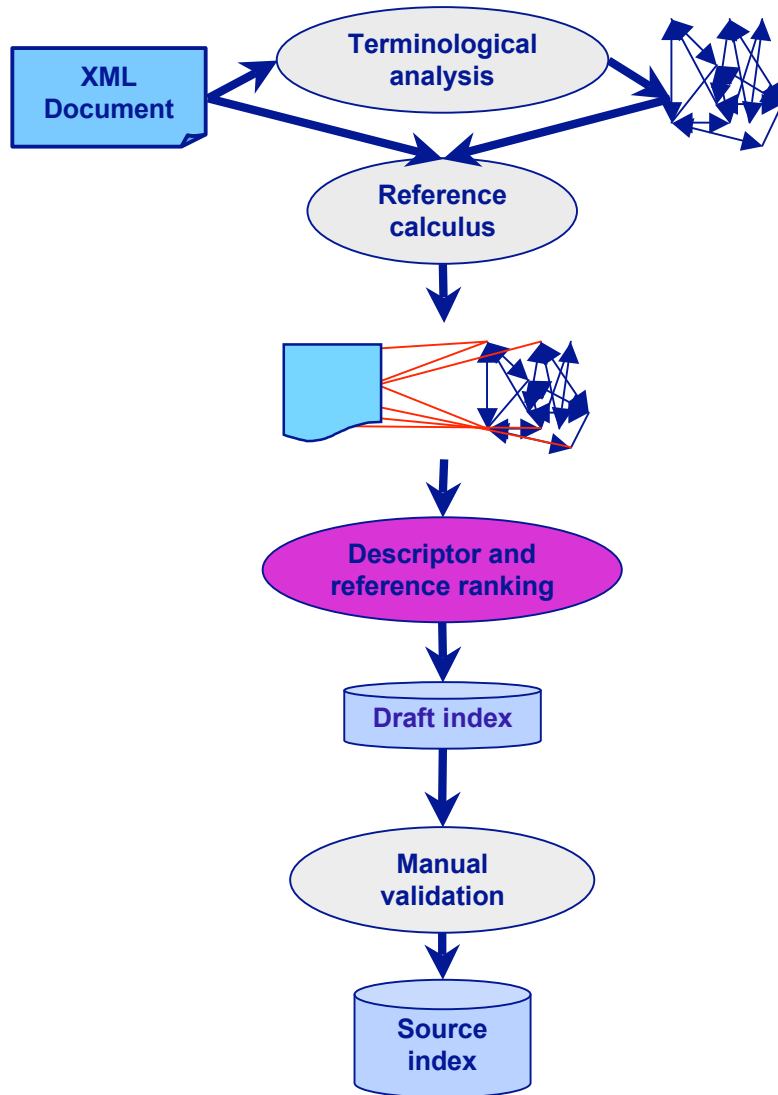


## For each descriptor

- Identification of descriptor occurrences
- Reference segment sizing (segmentation algorithm)
  - Statistics
  - Document structure
  - Typography
  - Terminological cohesion
  - Semantic density (# of relations)
  - Dictionary of linear integration linguistic markers

➤ A list of textual segments associated with each descriptor

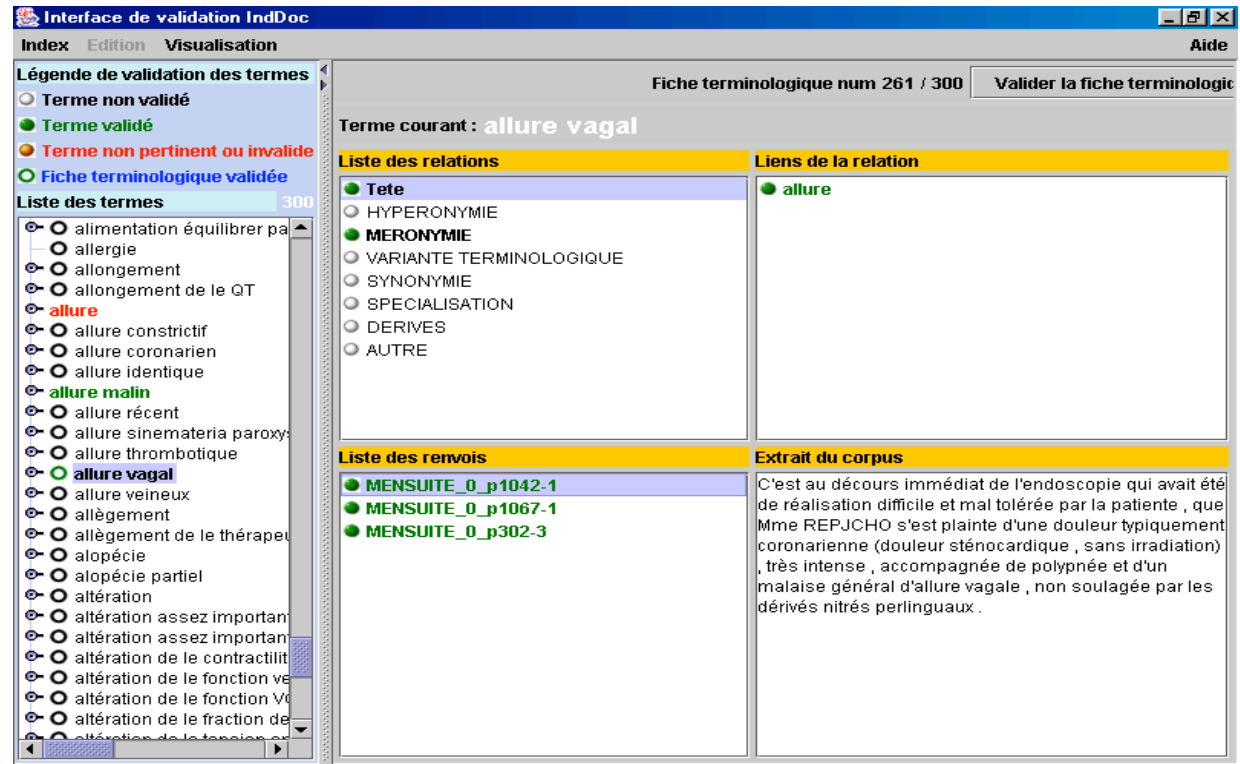
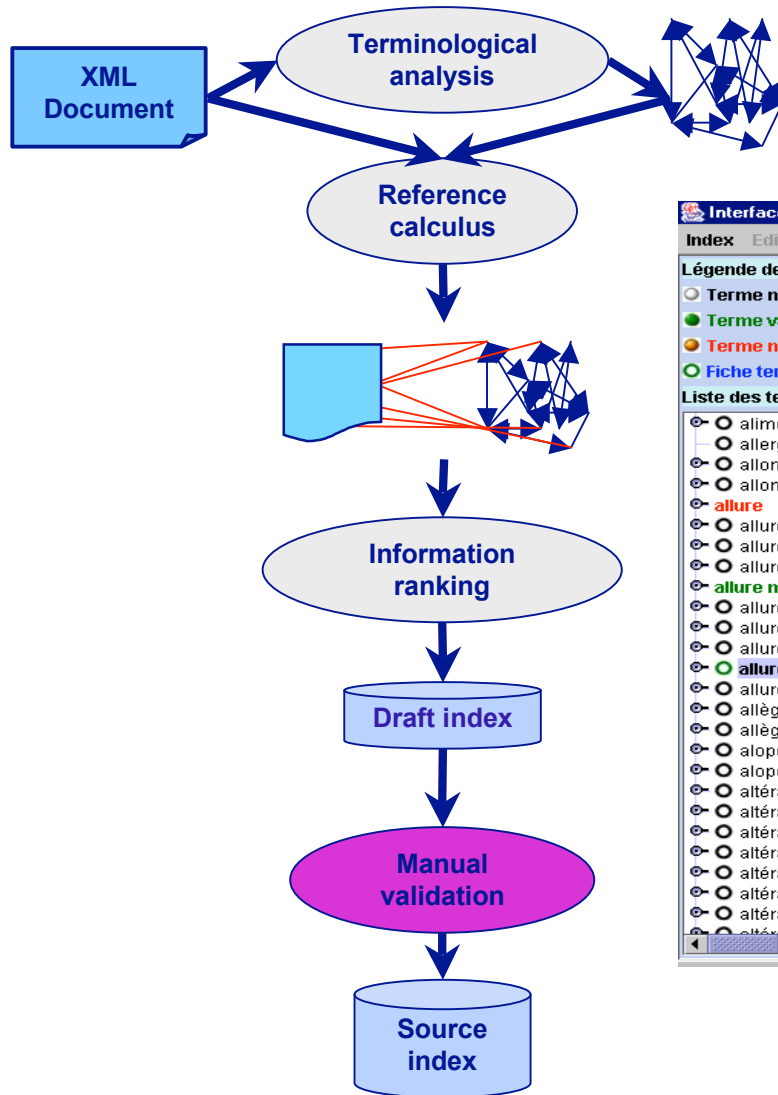
# DESCRIPTOR AND REFERENCE RANKING



## Relevance weights calculus based on

- Term frequency and distribution
  - Size and distribution of reference segments
  - Emphasis markers
    - Document structure
    - Typography
    - Linguistic markers (dictionary)
- **Ordered list of descriptors**
- **Ordered list of reference segments for each descriptor**

# VALIDATION INTERFACE



➤ Validated index

# Outline

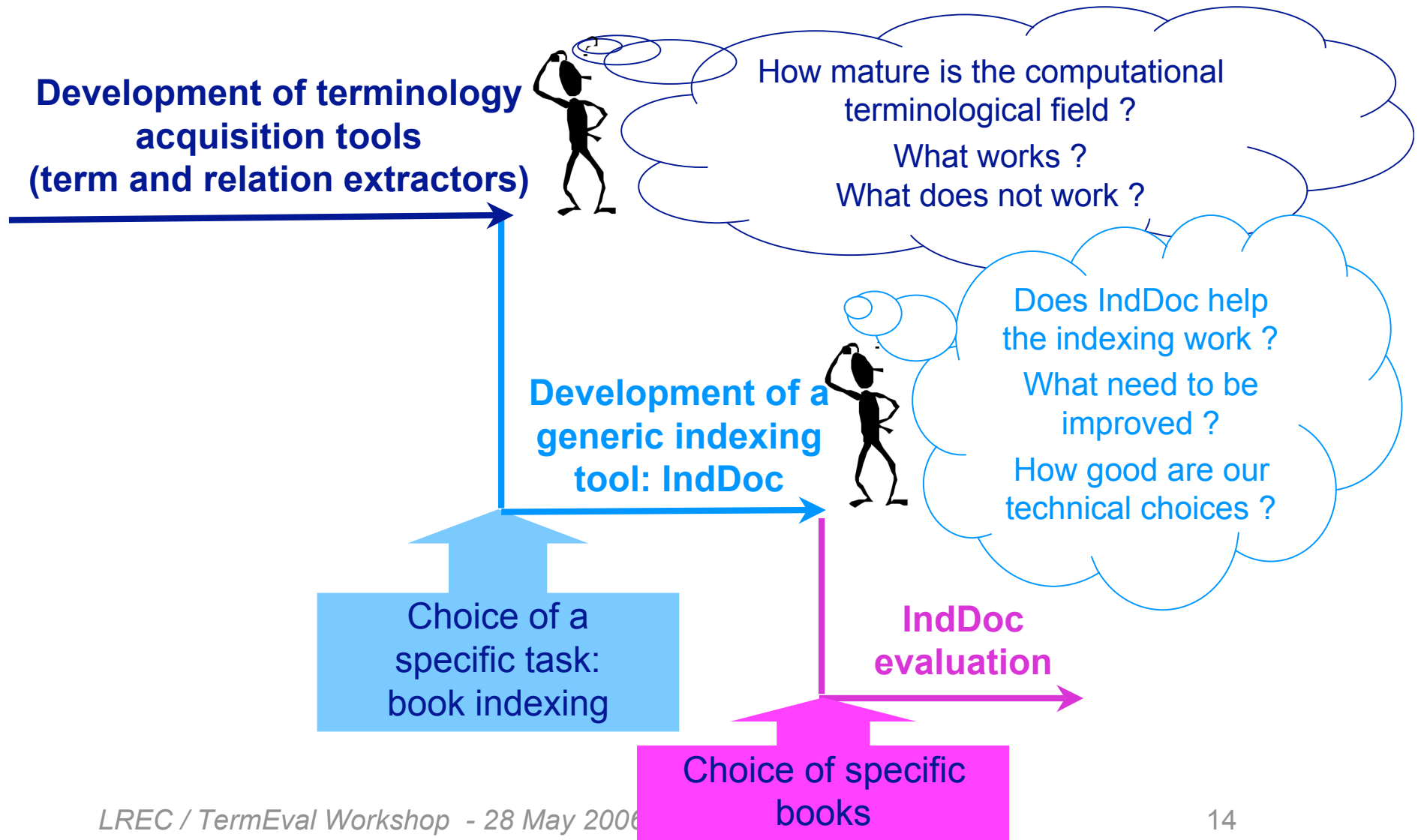
## Introduction

1. The IndDoc indexing tool
- 2. The evaluation challenge**
3. Our evaluation approach

## Conclusion

# EVALUATION

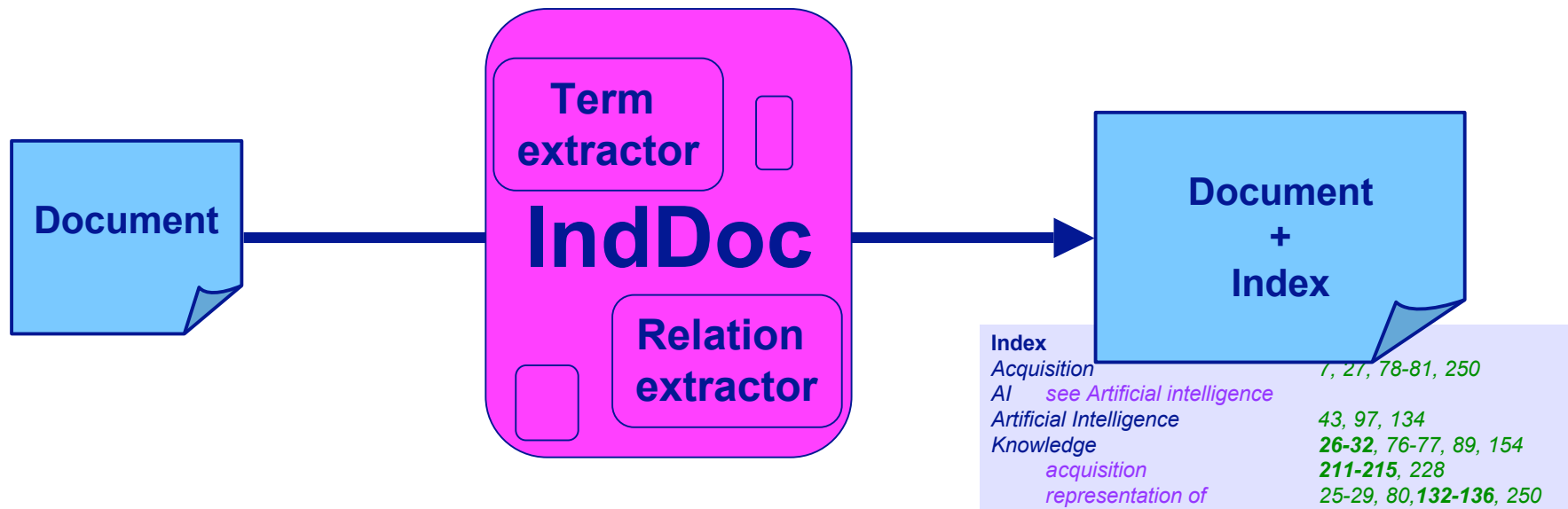
## Why?



# EVALUATION

## What ?

1. Evaluation of the quality of the results: resulting indexes
2. Indirect evaluation of the quality of the system that produced the results: IndDoc
3. Indirect evaluation of the quality of the system modules that compose the system: terminological tools



# THE EVALUATION PROBLEMS

## **Problem 1: There is no index reference**

- Professional indexers produce recommendations rather than models
- There is no real indexing tradition in France ( $\neq$  USA, GB, Canada)
- Indexing is subjective

## **Problem 2: There is comparable tool**

- IndDoc and existing tools rely on very different indexing strategies
- Indexing tools are designed for English

## **Problem 3: Our indexing method is cooperative**

- What is the role of the system?
- What is the role of the indexer?

## **Problem 4: It is too early to perform large-scale user tests**

- User tests are difficult to set up and expensive
- User tests require to have a complete prototype
- Developers need to have evaluation feedback as soon as possible



# Outline

## Introduction

### 1. The IndDoc indexing tool

### 2. The evaluation challenge

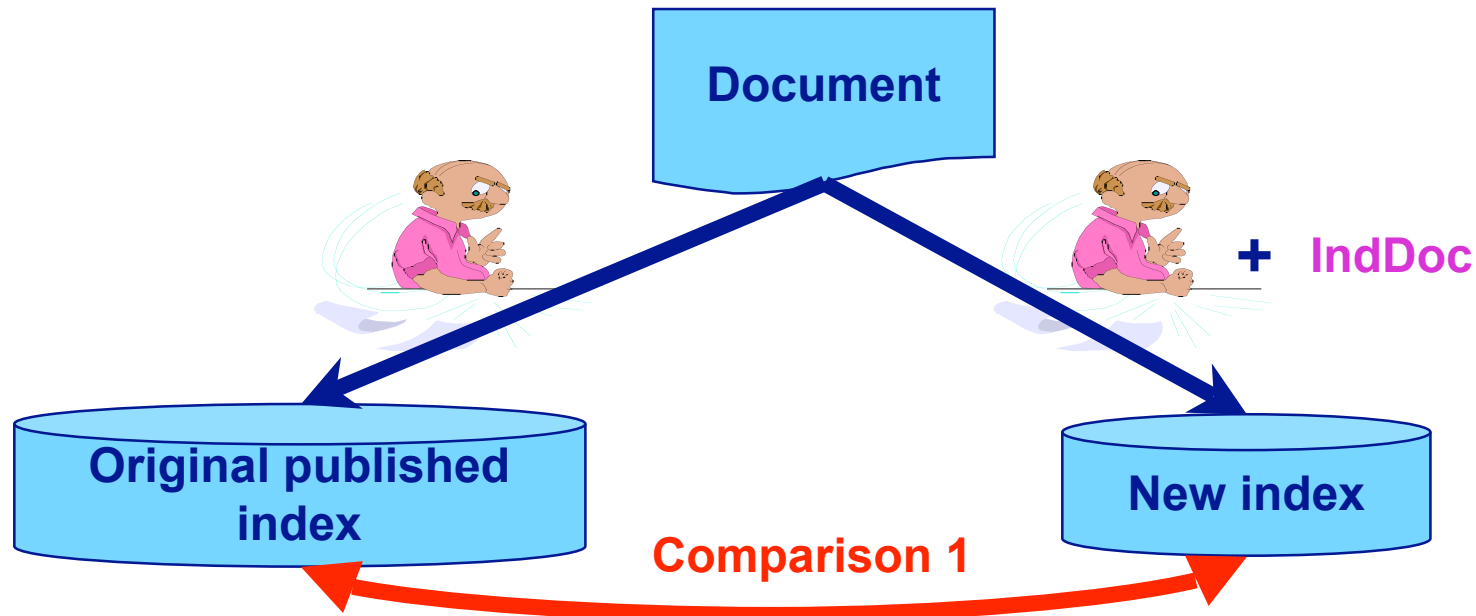
### 3. Our evaluation approach

- Comparison of a manual index and an IndDoc index
- Comparison of IndDoc and existing tools
- Comparison of the indexes before and after validation

## Conclusion

# FIRST COMPARISON

## Manual vs. IndDoc indexing



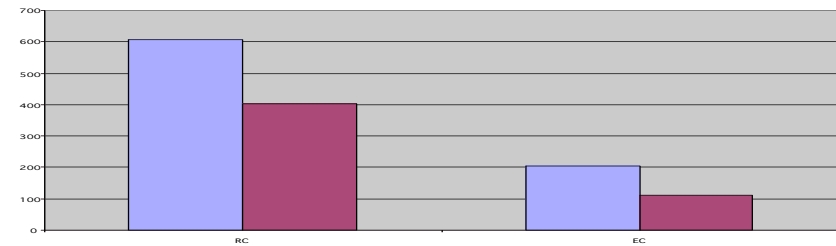
### Experiments on two monographs already published with indexes

- The manual indexes are provided by the book authors
- The books (without index) are analysed by IndDoc
- An distinct indexer validates the draft indexes
- The resulting validated indexes are compared with the published ones

# FIRST COMPARISON Results

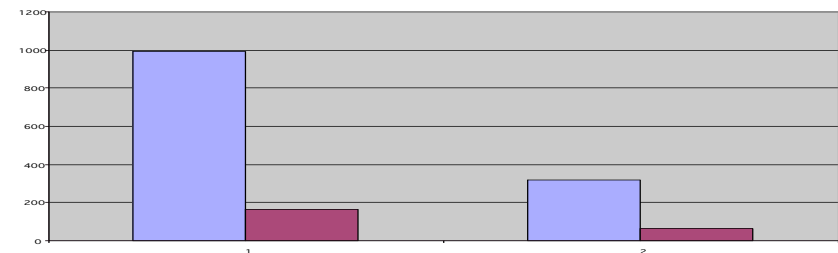
## Descriptors

- More descriptors: +50% and +80%
- No missing author descriptor (silence)



## Relations

- More relations: +166% and +300%
- Authors acknowledged that relations are useful although they did not introduce a lot of them manually



# FIRST COMPARISON

## Results (con't)

### Reference ranking

Detailed comparison of the reference lists for a sample of 30 descriptors having numerous references in the published index

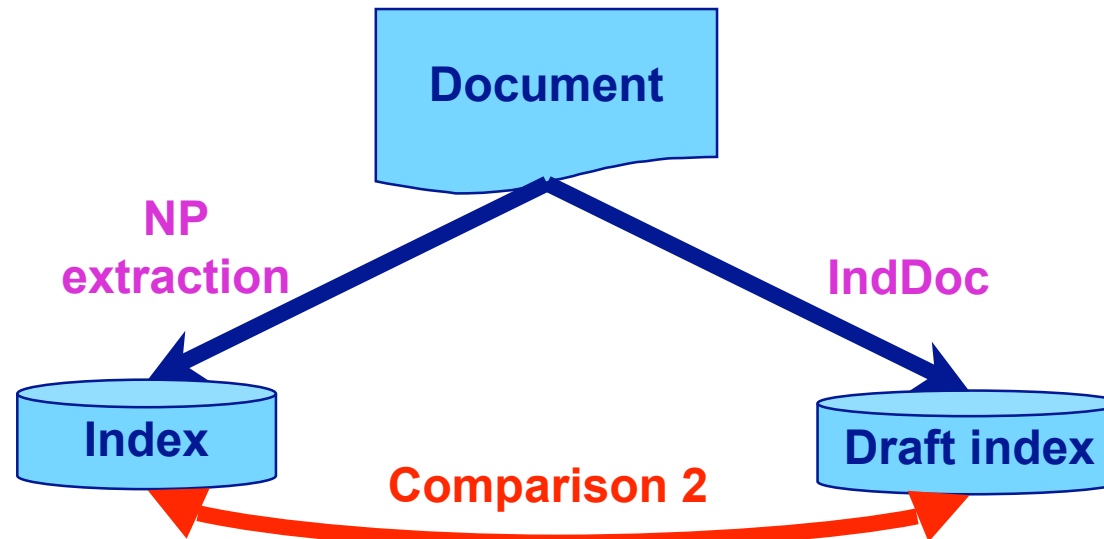
Correct ranking : 77%		Incorrect ranking: 23%	
Same ranking	Similar top ranking	Incorrect ranking	Undefined cases
17	6	4	3

### Indexing time

- No precise indication from the authors
- Estimation: validation is approximatively as long as manual indexing

# SECOND COMPARISON

## Basic vs. rich textual analysis

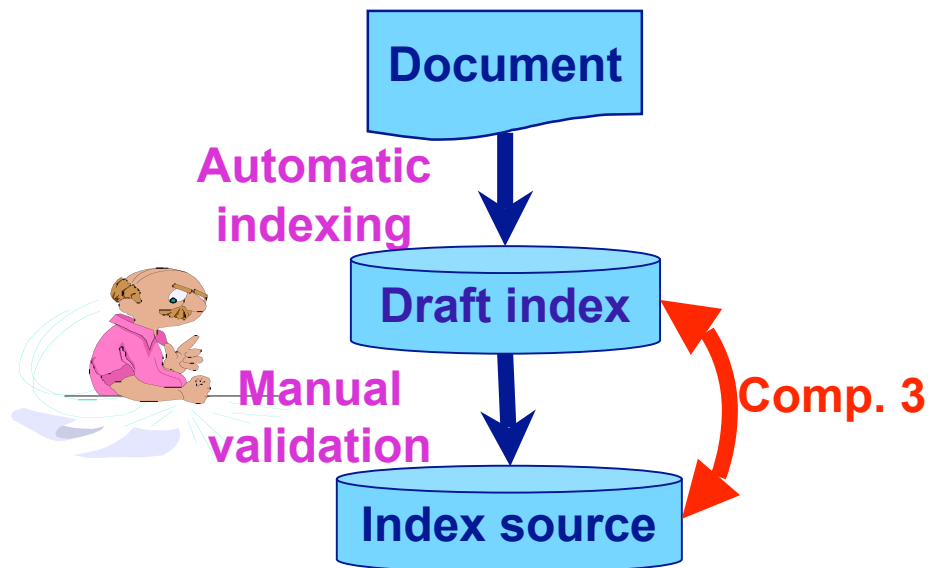


### Experiments on one monograph

- As no indexing tool exist for French, we simulate one by extracting the nominal phrases from the text manual indexes are provided by the book authors
- The books are also analysed by IndDoc (without validation)
- The resulting indexes are compared: IndDoc indexes are smaller

# THIRD COMPARISON

## Towards to a cooperative process



Books	LI	KR	KA
Desc. selection Precision	33%	44%	
Desc. selection Ranked precision	83%	77%	71%
Rel. selection Precision	65%	71%	80%

# AN OVERALL PICTURE

**Indexing using IndDoc takes as much time as indexing manually  
BUT**

- 1. The resulting index is much richer**
- 2. The index elements are ranked**  
(one can automatically produce an index of any given size)
- 3. The draft index is good enough to be validated**

**➤ The IndDoc approach is good !**

# CONCLUSION

## **A pragmatic approach**

- Evaluate as much as possible
- Make use of existing materials
- Exploit any experiment to evaluate (how partial and specific they be)

**Such a pragmatic evaluation validates the overall approach  
BUT it gives only fragmentary results. It is not sufficient !**

## **To go further, we still have to perform...**

- Large-scale evaluation
- Detailed subcomponent evaluation
  - The component intrinsic quality
  - The impact of the component on the global system
- User-based evaluation



# INDEX

## An hybrid network

