

# The LIA Summarization Systems at DUC 2007

*florian.boudin@univ-avignon.fr*

Laboratoire Informatique d'Avignon, France

**co-authors**: Frédéric Béchet, Marc El-Bèze, Benoit Favre,  
Laurent Gillard and Juan-Manuel Torres-Moreno

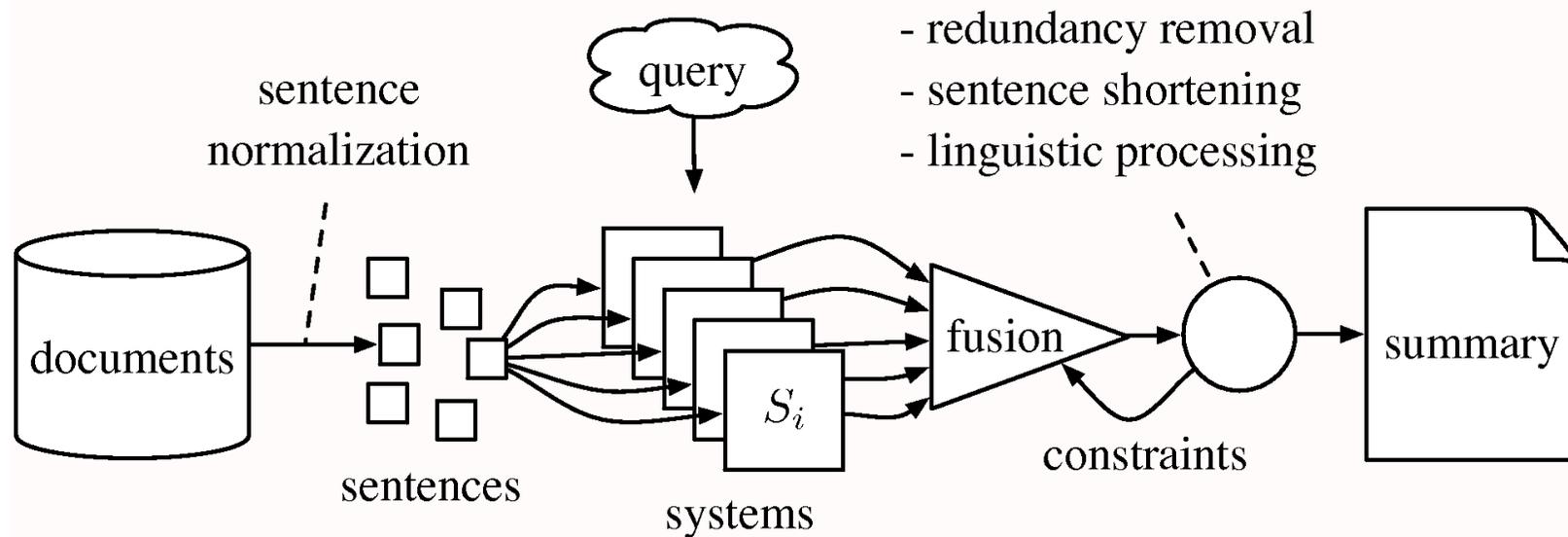
# Outline

- Main task
  - Using a fusion process ?
  - Results
  - Discussion
- Update task
  - Cosine maximization-minimization approach
  - Novelty boosting
  - Results
  - Discussion

# Main Task

# How is it working

- Use of several different summarizers as sentence selection components



# Using a fusion process ?

- Successful in other domains
  - Classification
  - Speaker Recognition
- Robustness
  - Small training dataset
- Reliability
  - Smoothing system performance variations

# More summarizers

- 5 systems in 2006, 7 systems in 2007
  - (S1) MMR+LSA (2006 & 2007)
  - (S2) Neo-Cortex (2006 & 2007)
  - (S3) n-term with variable length insertion (2006 & 2007)
  - (S4) LNU\*LTC (2007)
  - (S5) Okapi similarity (2007)
  - (S6) Prosit similarity (2007)
  - (S7) Compactness score (2006 & 2007)
  - (S8) Passage retrieval (2006)

# Fusion strategy

- Combining each system output
  - Ranked sentence lists
- Building a sentence graph
  - Sentences weighted according to their ranks and scores
- Output summary
  - The best path in the graph

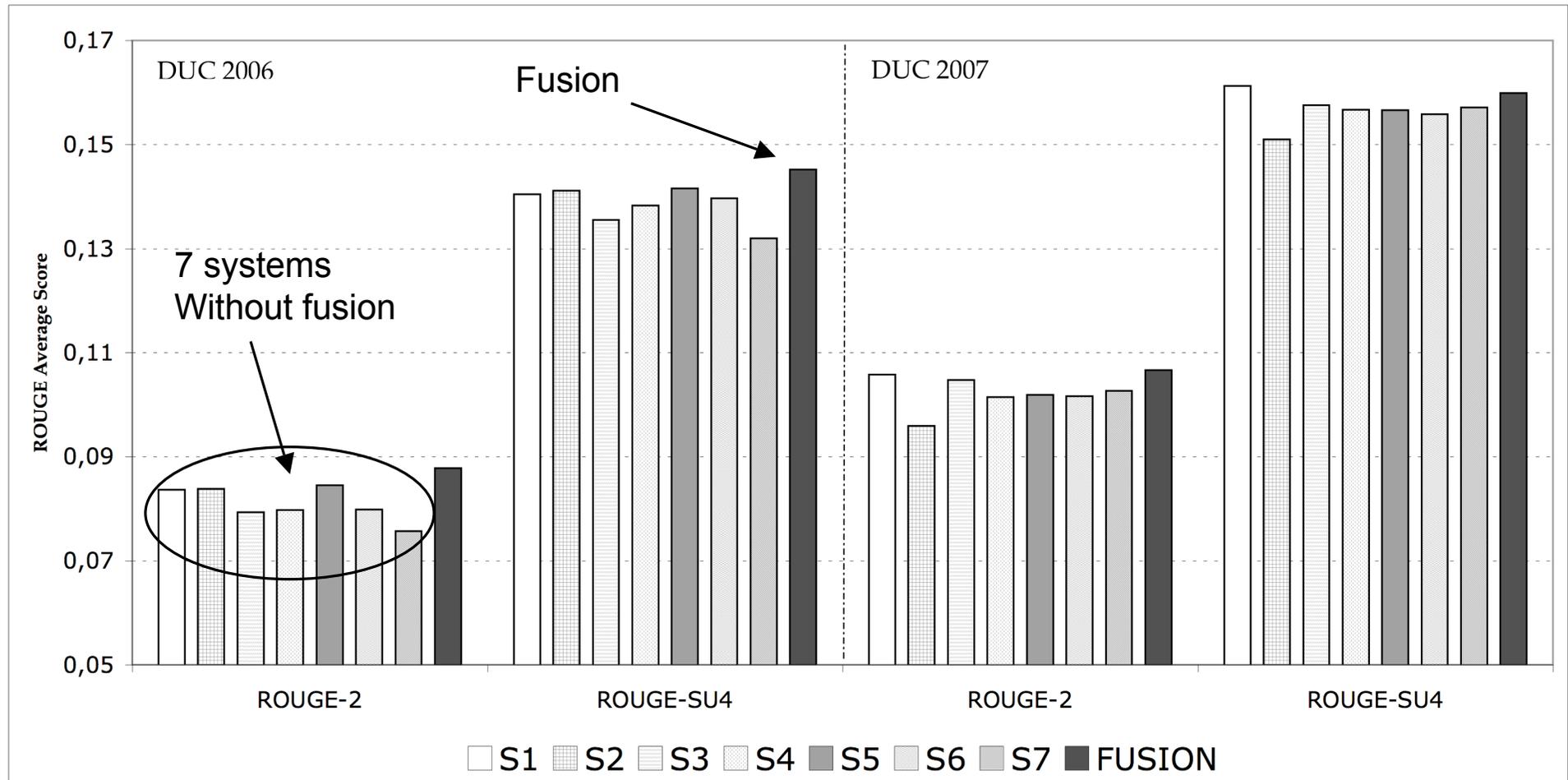
# Post-processing

- Person name rewriting
- Acronym rewriting
- Redundancy removal
  - word overlap
- Fusion, a second pass
  - New sentence lengths, redundancy and rewriting are backpropagated

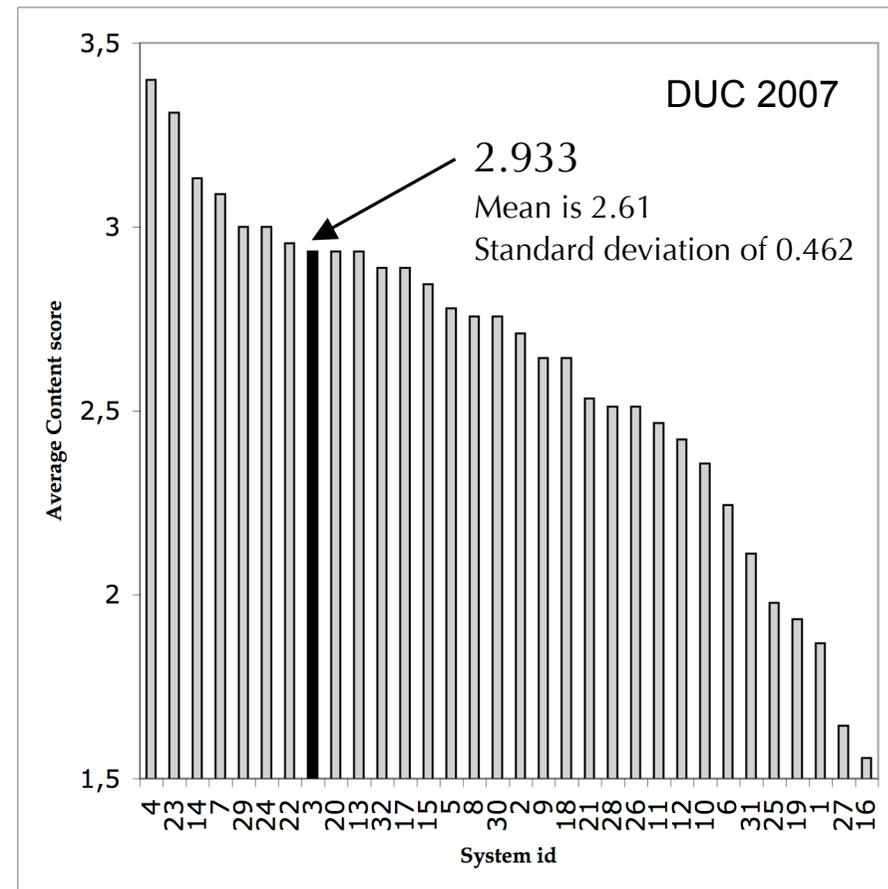
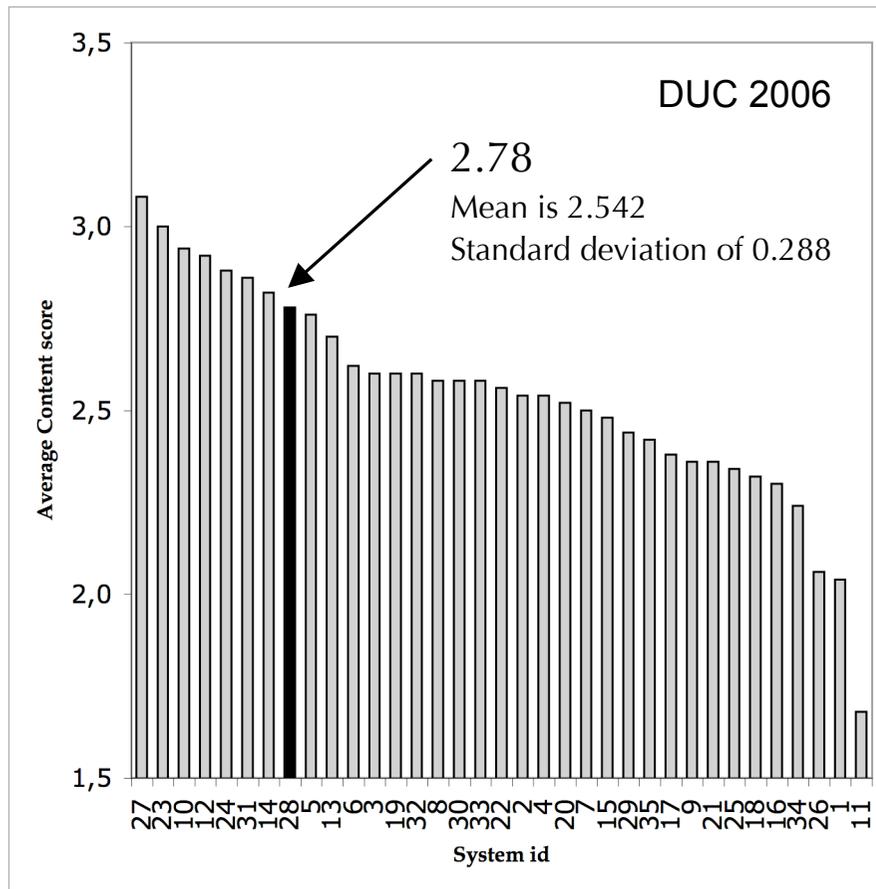
# Results

Comparison between 2006 and 2007

# Automatic evaluation



# Manual evaluation (1)



# Manual evaluation (2)

Measure	2006	2007
Mean linguistic quality	3.57	3.42
Grammaticality	4.08	4.11
Non-Redundancy	3.84	3.62
Reference Clarity	3.42	3.36
Focus	3.74	3.56
Structure and Coherency	2.76	2.47

- Linguistic quality scores of our submission in 2006 and 2007
- Unchanged linguistic processing module
- Small difference between the two evaluations

# Fusion - Conclusions

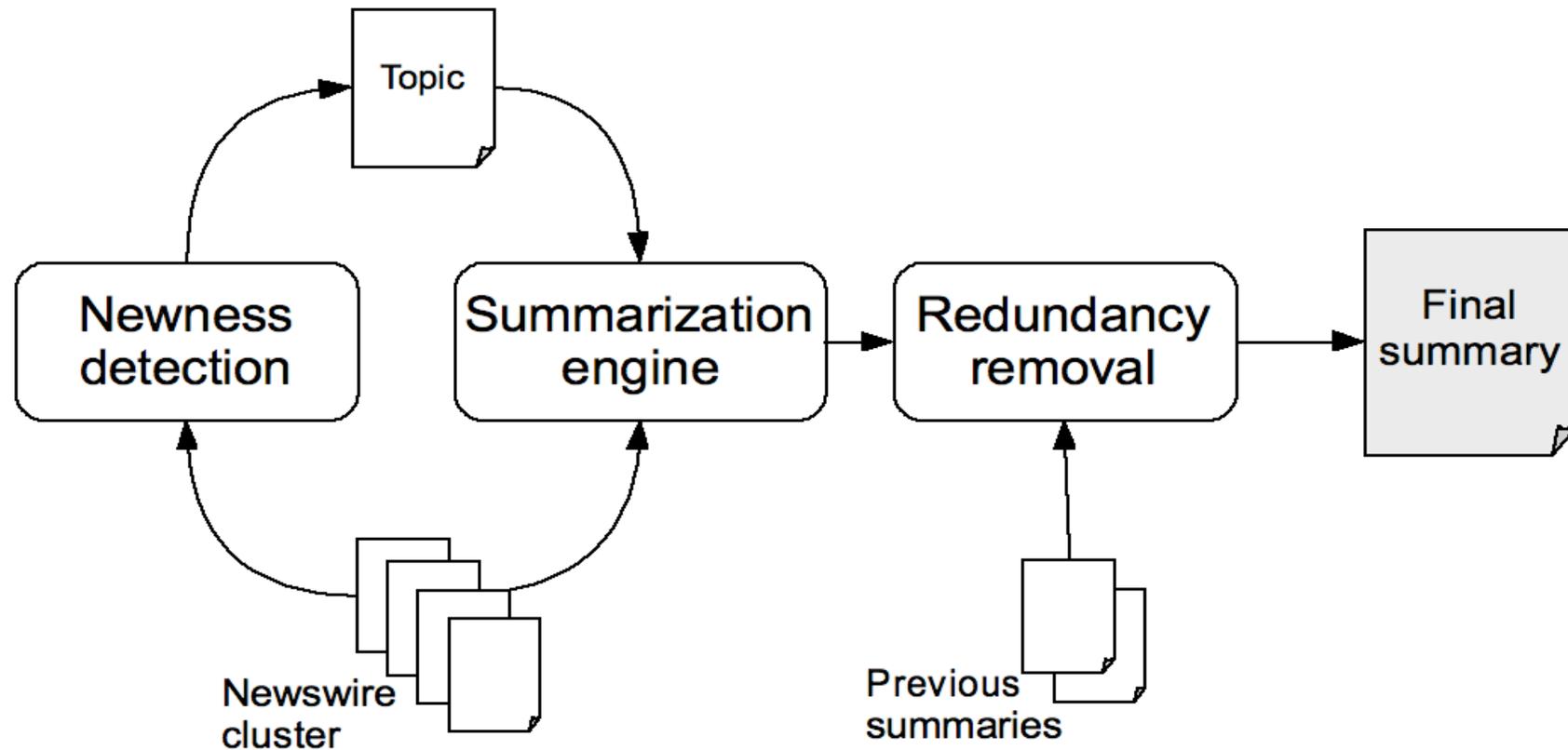
- Outperforms the best system
- Prevent overfitting
- Toolkits available (we use the AT&T FSM toolkit)
- Flexible
- Parameter tuning using a development corpus

# Update Task

# Principle

- Based on a very simple user-focused Multi-Document Summarizer (MDS)
  - Similarity with topic
- Added features:
  - Cross summaries redundancy removal
    - Cosine maximization-minimization
  - Novelty boosting
    - Topic enrichment

# How is it working



# A simple user-oriented MDS

- Documents segmented in sentences
- Sentences  $\vec{s}$  filtered and stemmed
- Each sentence is scored in relation to the topic  $\vec{t}$ 
  - *cosinus* angle written  $\eta(\vec{s}, \vec{t})$
  - *tf.idf* weights
- Drawbacks
  - Summaries do not inform the reader of new facts
    - Cross summaries redundancy removal techniques
    - Novelty boosting

# Two-step cosine maximization-minimization (1)

- Improved sentence scoring method
  - Cross summaries redundancy removal

$$R(\vec{s}, \vec{t}, \Pi) = \frac{\eta(\vec{s}, \vec{t})}{\phi(\vec{s}, \Pi) + 1}$$

← sentence|topic

→ Sentence|early summaries

$$\max R(s) \implies \begin{cases} \max \eta(\bullet) \\ \min \phi(\bullet) \end{cases}$$

# Two-step cosine maximization-minimization (2)

- Limits
  - All sentences are scored in relation to the same topic
    - Selected sentences are syntactically related
  - Force irrelevant sentences to enter the summary
- ↘ Propose a novelty boosting technique

# Novelty boosting

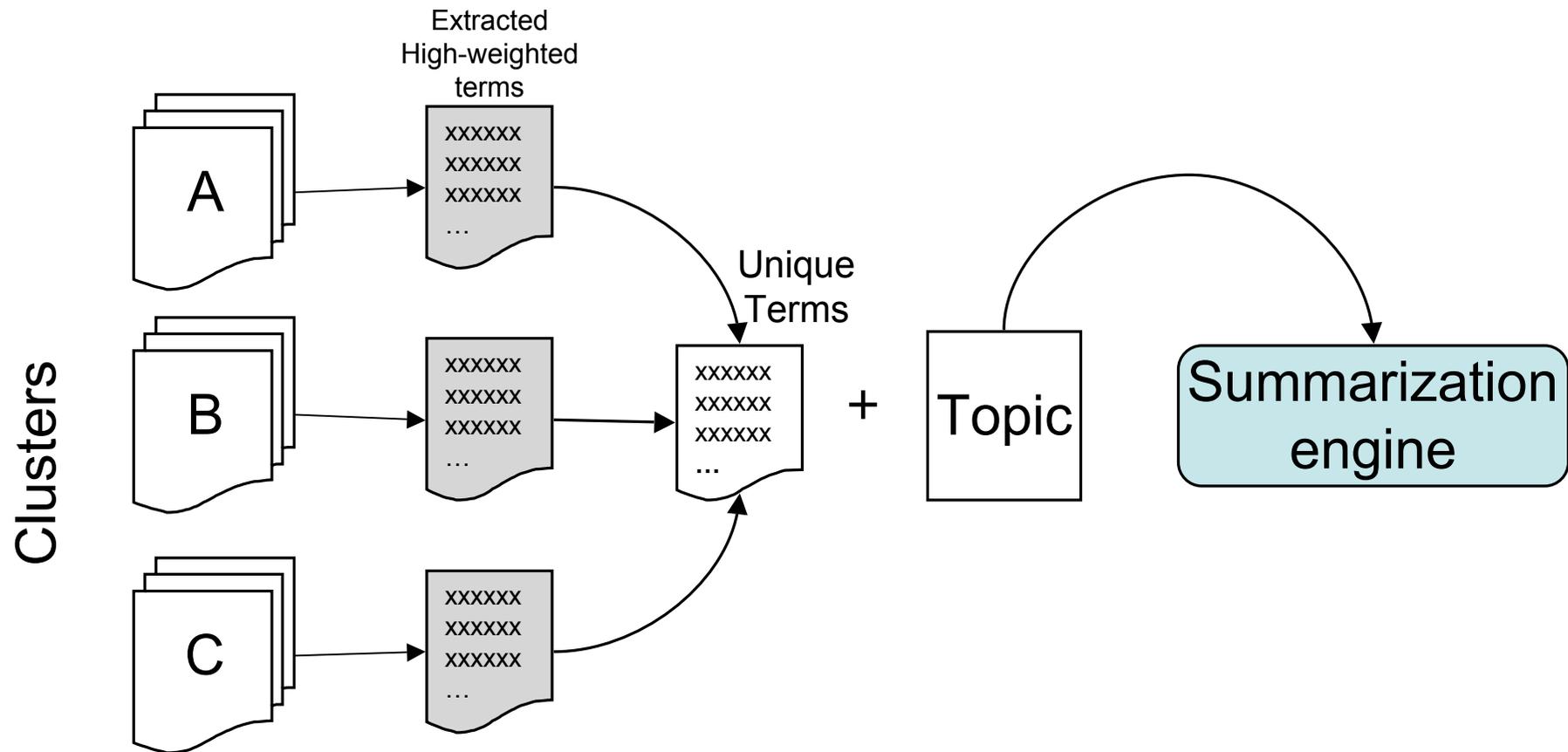
- Point summary to the major cluster novelty
  - Novelty in comparison to early clusters
  - Extraction of high weighted term lists
- Topic enrichment using the unique terms

Enrichment  
Bag of words

$$B_X^{boost} = B_X \setminus \bigcup_{i=1}^{n_p} B_i$$

Early clusters's  
Bag of words

# Example (Novelty boosting for cluster C summary)



# Summary construction (1)

- Arranging the most high scored sentences
- No special order within the summary
- Limit of 100 words
  - ↳ high probability of truncated last sentence
- Propose a better last sentence selection method

# Summary construction (2)

Last sentence selection method :

- If remaining word number  $> 5$ 
  - After-last preferred if
    - Length 1/3 shorter
    - Score greater than a threshold
      - » threshold obtained empirically
  - Otherwise truncate sentence
- Else produce non-optimal sized summary

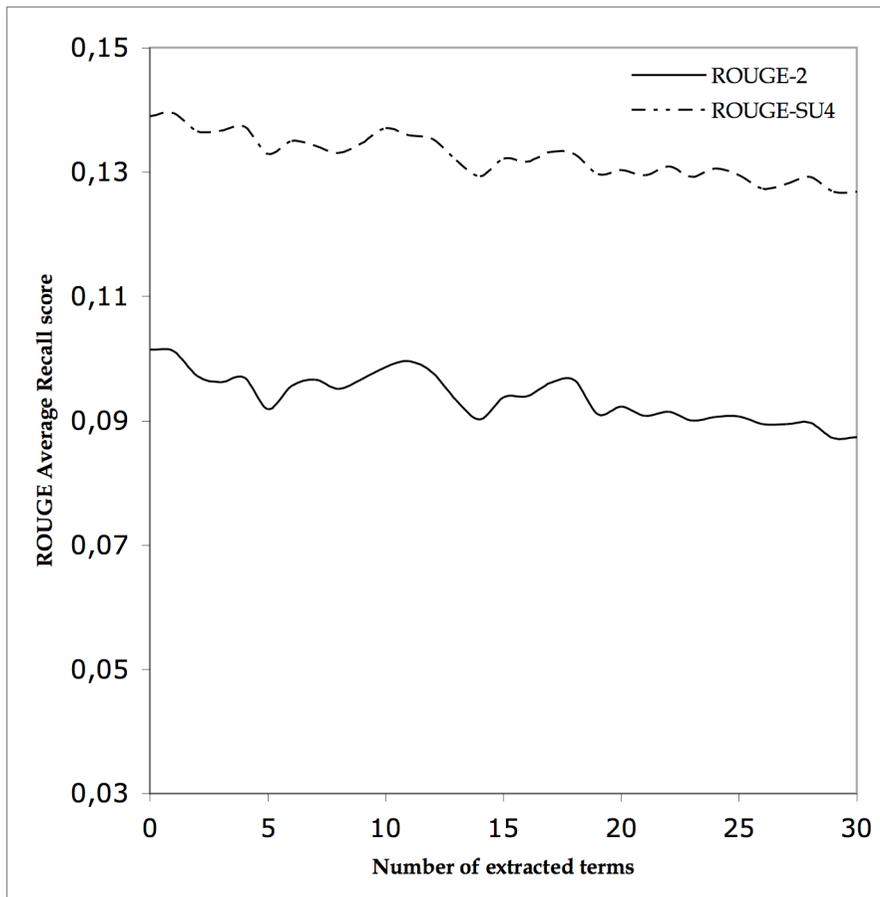
# Post-processing (1)

- Within summary redundancy removal
  - Cosine similarity with threshold
  - Threshold obtained empirically (~ 0.4)
- Sentence Rewriting techniques
  - Person name rewriting
    - *Vice President Al Gore ...*
    - *... Al Gore ...*

# Post-processing (2)

- Sentence Rewriting techniques
  - Acronym rewriting
    - *Massachusetts Institute of Technology ...*
    - *... MIT ...*
  - Link words removal, say clauses removal
    - ~~*Moreover, the president is ...*~~
    - ~~*... said the judge.*~~
  - Cleanup punctuation

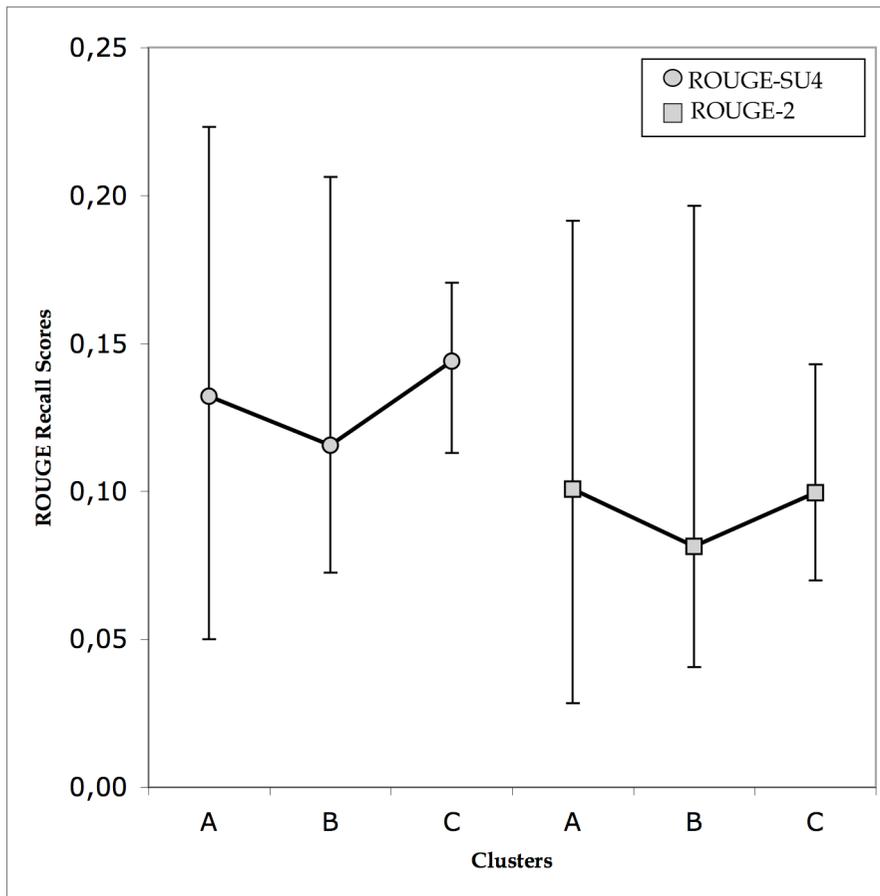
# Experiments (1)



Automatic evaluations (ROUGE-2 and SU4) in relation to the number of extracted terms

- Novelty boosting introduces « noise »
- Enhances the readability

# Experiments (2)

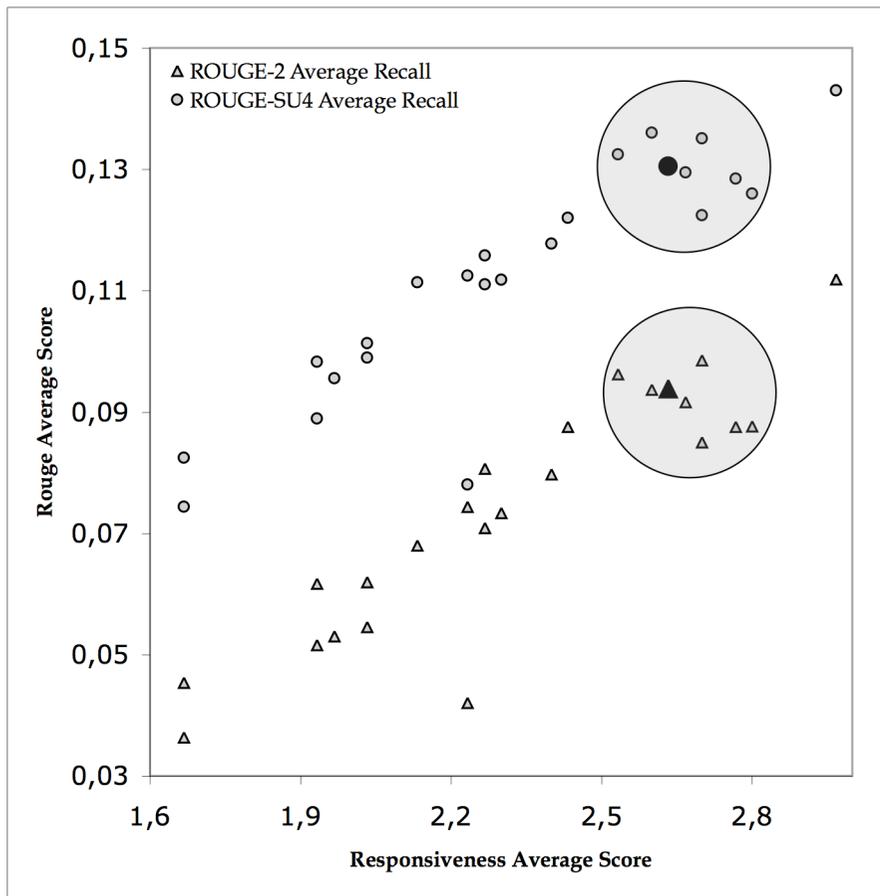


Automatic evaluations (ROUGE-2 and SU4) for each cluster of documents (A~10, B~8 and C~7 articles)

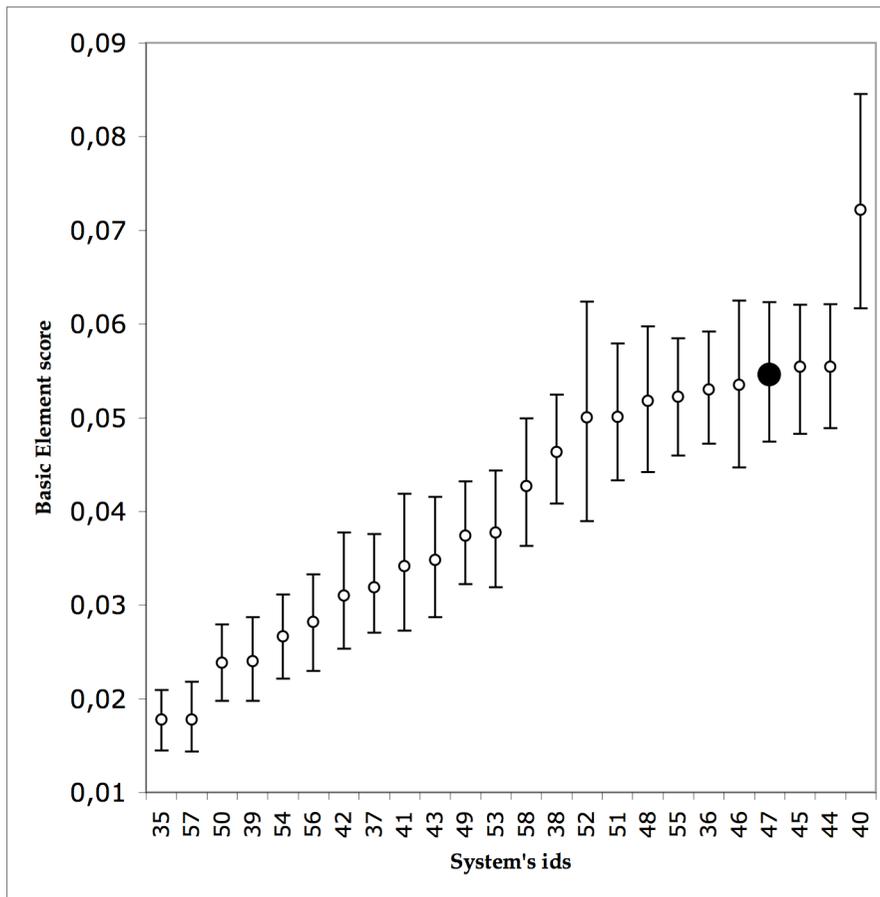
- Enhances system stability and reliability
- Non-optimal enrichment
  - Slight decrease with cluster B

# Results at DUC 2007

# Results (1)



# Results (2)



Automatic evaluations (Basic Elements) for each system at DUC 2007

- BE score 0.0546
  - Mean = 0.0409
  - Standard deviation = 0.0139

# Conclusion

- Very simple approach
- Summary quality enhanced across time
- Novelty boosting
  - Helps preventing within redundancy
  - Introduces “noise”
- Language Independent

# What's next ?

- Enhance cross summaries redundancy removal process
  - Change granularity
    - Considering previous sentences instead of summaries
- Dynamic novelty boosting
- Improve sentence rewriting techniques

# Thank You !

Florian.boudin@univ-avignon.fr

**co-authors**: Frédéric Béchet, Marc El-Bèze, Benoit Favre,  
Laurent Gillard and Juan-Manuel Torres-Moreno

This work was partially supported by the *Laboratoire de chimie organique de synthèse*,  
FUNDP (*Facultés Universitaires Notre-Dame de la Paix*), Namur, Belgium