Multimodal information approaches for the Wikipedia collection at ImageCLEF 2011

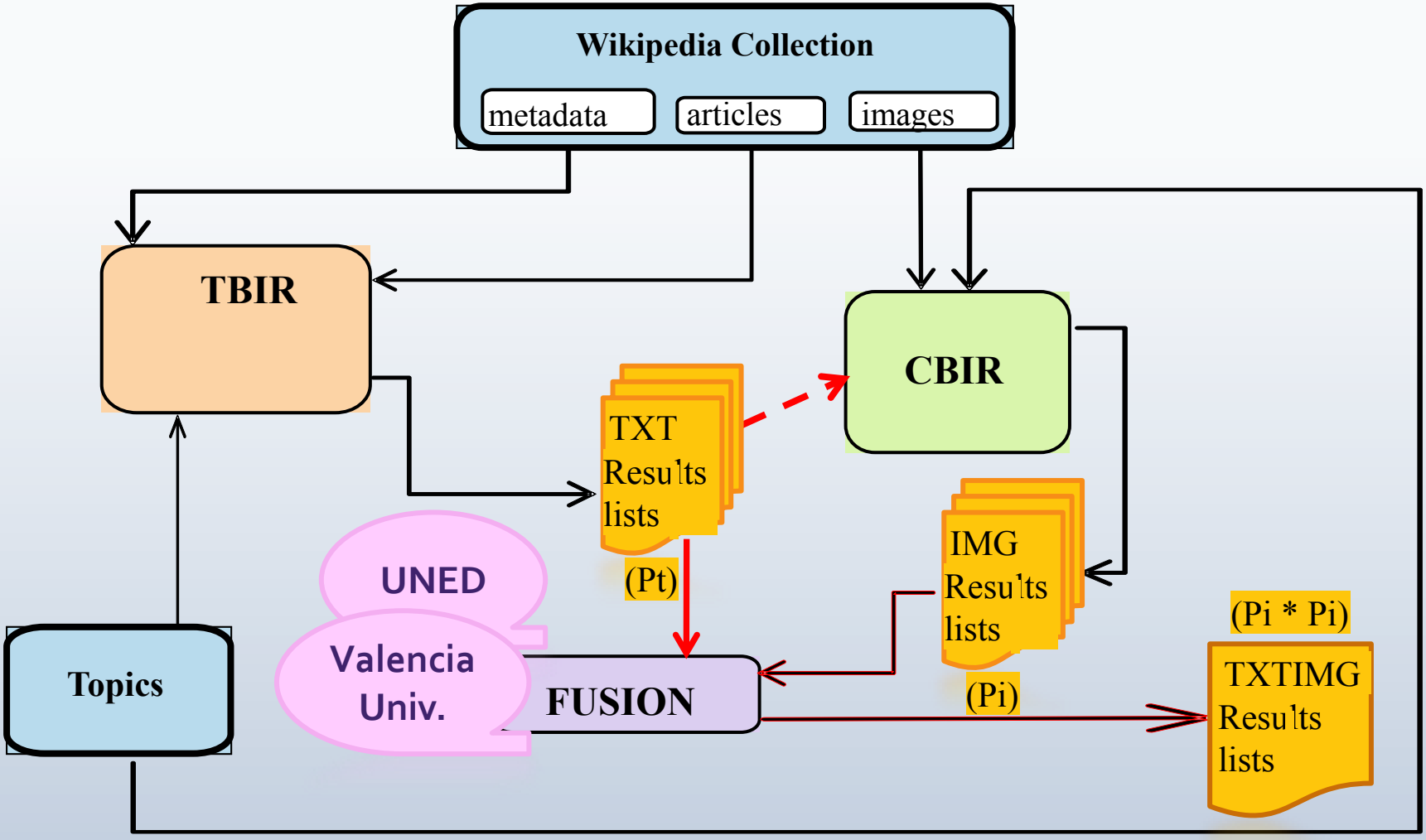
1 Universidad Nacional de Educación a Distancia, UNED
2 Universidad de Valencia, UV

R. Granados[1], J. Benavent[2], X. Benavent[2], E. de Ves[2],
Ana García-Serrano[1]

Index

1. Objectives UNED-UV @ ImageCLEF2011
2. System overview
3. TBIR: Textual approach
4. CBIR: Visual approach
5. Fusion: merge and reordering
6. Runs and Results
7. Conclusions and future work

2. System Overview



3TBIR

- IDRA Tool (InDexing and Retrieving Automatically)
  Our own implemented tool (under GPL 3.0)

- Architecture in several modules

  - **Preprocessing** (language-dependent):
    - Stopwords, Stemming (Snowball)
    - NER (Stilus - DAEDALUS)

  - **Indexing**: Vector Space Model approach (IDRA) and Basic Lucene executed from IDRA

  - **Merging Module to Multilinguality and textual enrichment**:
    - Join (features)
    - Enrich (merge of two results lists)
    - MAXmerge (merge of two results lists)

6. TXT Runs and Results - Baseline

ID	Lang	Details				
		Sub-System	Fusion	MAP	Run Pos.	Group Pos.
run1	EN	IDRA	-	0.1727	44	9
run2	EN	IDRA + stemming	-	0.2056	26	8
run3	EN	IDRA + stemming + Cat	-	0.2243	21	6

- Stemming (EN, DE, and FR)
- WP Cat and TITLE are joined to DE, EN, FR

3TBIR

- Differences between **IDRA** indexing and ranking and **Basic Lucene**

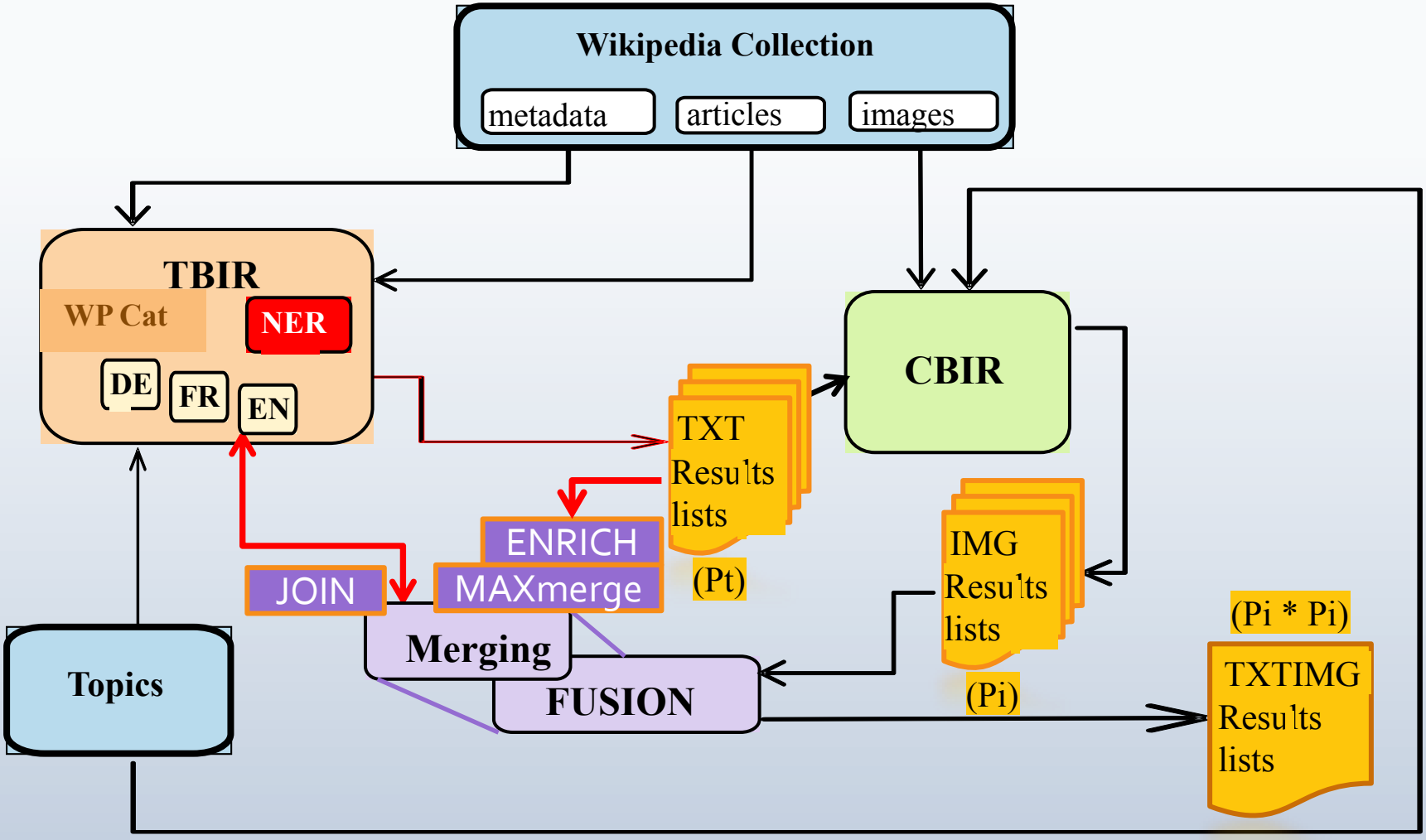
ID	Lang	Details				
		Sub-System	Fusion	MAP	Run Pos.	Group Pos.
run3	EN	IDRA	-	0.2243	21	6
run4	**DE+FR +EN**	IDRA	MAXmerge (3en, 3fr, 3de)	0.2489	14	4
run5	EN	IDRA+Lucene	-	0.2601	10	3
run8	**DE+FR +EN**	**IDRA+ Lucene**	**MAXmerge (5en, 5fr, 5de)**	**0.3044**	**3**	**2**

3TBIR

- **NEs** are indexed separately or Joined with EN metadata

ID	Lang	Details				
		Sub-System	**Fusion**	**MAP**	**Run Pos.**	**Group Pos.**
run5	EN	IDRA+Lucene	-	0.2601	10	3
run7	EN	IDRA+ Lucene +NER	Enrich (5en,NEs)	0.2515	13	4
run10	EN	IDRA+ Lucene +NER	JOIN	0.2403	16	4

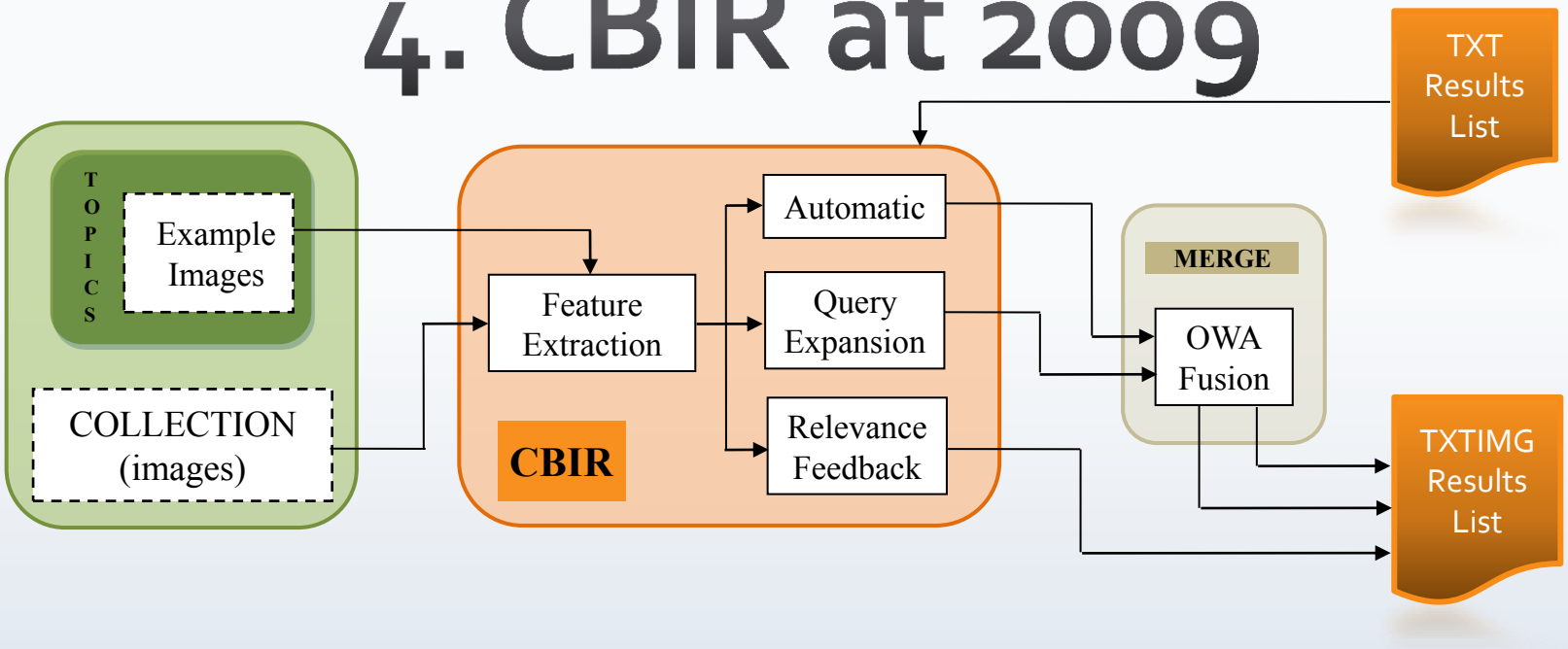
2. Multilingual Search



6. Runs and Results - Multilinguality

ID	Lang	Details				
		Sub-System	Fusion	MAP	Run Pos.	Group Pos.
run5	EN	IDRA+Lucene	-	0.2601	10	3
run8	**DE+FR +EN**	**IDRA+ Lucene**	**MAXmerge (5en, 5fr, 5de)**	**0.3044**	**3**	**2**
run 9	ALL	IDRA+ Lucene	JOIN	0.2758	9	3

4. CBIR at 2009



Low level Features:

- Color
- Texture

Algorithms:

- Automatic
- Query expansion: N first TXT Results List
- **Relevance feedback:** Logistic Regression algorithm

Best results at 2009

4. CBIR (Relevance Feedback)

- The logistic regression model is used to estimate the relevance probabilities of each image in the collection (10.000).

$$x = (x_1, \ldots, x_p)$$

$$\pi(x) = P(Y = 1 \mid x)$$

$$logit[\pi(x)] = \alpha + \beta_1 x_1 + \ldots + \beta_p x_p = \ln \frac{\pi(x)}{1-\pi(x)}$$

$$\pi(x) = \frac{\exp(\alpha + \beta_1 x_1 + \ldots + \beta_p x_p)}{1 + \exp(\alpha + \beta_1 x_1 + \ldots + \beta_p x_p)}$$

- 2011: Manual relevance feedback into automatic one
  - The relevant images set: query images plus the first 3 ranked images of the text filter.
  - The non-relevant images set: first we select # random images out of the text filter, then it is ordered by the Euclidean distance, and finally the negative examples are de M latest ones.

4. CBIR (ImageClef@Wikipedia 2011)

- Testing the behavior of our low-level features (UV) with the CEDD ones
  - UV Low level features:
    - Color information (four fragments of same size): A feature vector of 222 components
    - Texture information (granulometric distribution and Spatial size distribution): A feature vector pf 71 components

ID	TXT run	Details			
		MAP	Fusion	Run Pos.	Group Pos.
run11	run9+UV	0.0553	Pi	56	8
run12	run9+CEDD	0.0516	Pi	57	8
run13	run9+CEDD	0.2869	Automatic	18	4
run14	run9 +CEDD	0.3006	Pt*Pi	12	3
run17	run9+UV	0.3006	Pt*Pi	13	3

5. Late Fusion

Main proccess:

- First TBIR works over the whole Image Collection
- CBIR works from the TXT results list (10.000)

  TBIR acts as a filter based on the assumption that the textual module better captures the meaning of a topic.

- Approaches (Pt, Pi ranges are not normalized):
  - Pt*Pi, Pt = 1 (only visual ranking)
  - Pt*Pi
  - OWA: average(Pt, Pi) , orness(Pt, Pi) with 0.3

6. Runs and Results

ID	TXT run	Details			
		MAP	Fusion	Run Pos.	Group Pos.
run14	run9 +CEDD	0.3006	Pt*Pi	12	3
run15	run9 +CEDD	0.2869	OWA(Med)	17	4
run16	run9+CEDD	0.2980	OWA(Orness0.3)	14	3
run18	**run8**+CEDD	**0.3405**	**Pt*Pi**	**8**	**2**
run19	run8+CEDD	0.3233	OWA(Med)	9	2
run20	run8+CEDD	0.3367	OWA(Orness0.3)	10	2

0.04

7. Future Work

- Good way of working fusing textual and visual list
  - Refine the identified differences between the three ways of calculating the score of the final list
- CBIR: Non Automatic Relevance feedback algorithm based on logistic regression
- TBIR: Why NEs does not improve previous results?
- Textual enrichment (WP articles summaries)
-
-