

WebCLEF breakout

2005 and beyond

Agenda

- **Issues related to 2005**
- **Methodological issues**
- **Forward looking**
- **Help!**

Issues related to 2005

- Thank you's
- Feedback
 - Why did you/did you not take part?
 - Effort on topic development
 - Instructions, tools, procedures, deadlines,...
- Results for multilingual

Methodological points

- Please note: these are **challenges**
 - User task/model
 - Incompleteness of “relevant in translation”
 - Dominant language issue?
 - Do we want ad hoc?

Looking forward

- Policy on sharing WebCLEF 2005 data
- More of the same?
 - Mixed monolingual, multilingual
 - New tasks?
 - New crawl?
- Other corpus, other type of web documents
 - Blogs

Help!

- **Crawling**
- **Encoding normalizer**
- **Link analysis**
- **Result analysis**
- **...**

WebCLEF 2006 & beyond

- Run same task as WebCLEF 2005, with patches to corpus, and markup as requested by participants
 - of topics
 - pilot: exploring blogs