KVASIR Project – Knowledge ViA Semantic Information Retrieval

Liang Wang, Jussi Kangasharju

Department of Computer Science, liang.wang@cs.helsinki.fi

UNIVERSITY OF HELSINKI

Future Content-Aware Network

- **Motivation** – Content consumers only care about "what they need".

- **Question** – How to provide smart content provision and recommendation?

- **Objective** – Content are interconnected semantically in the network.

Information-Centric Architecture

- Access content by name instead of location.
- Push popular content to network edge.
- Open network interface to applications.

Research Challenges

- How to efficiently reduce the dimensionality?
- How to perform fast search in high-dimensional space?
- How to effectively minimize the index file size?
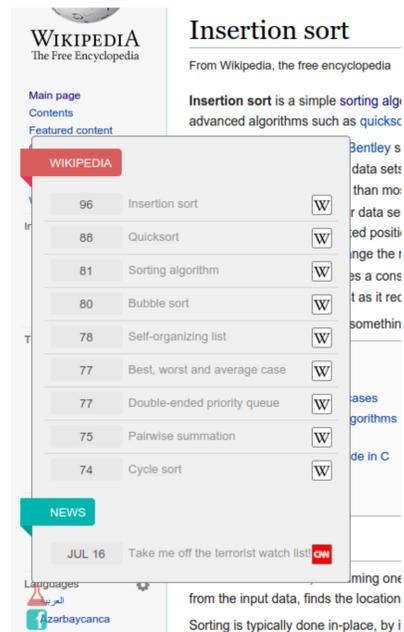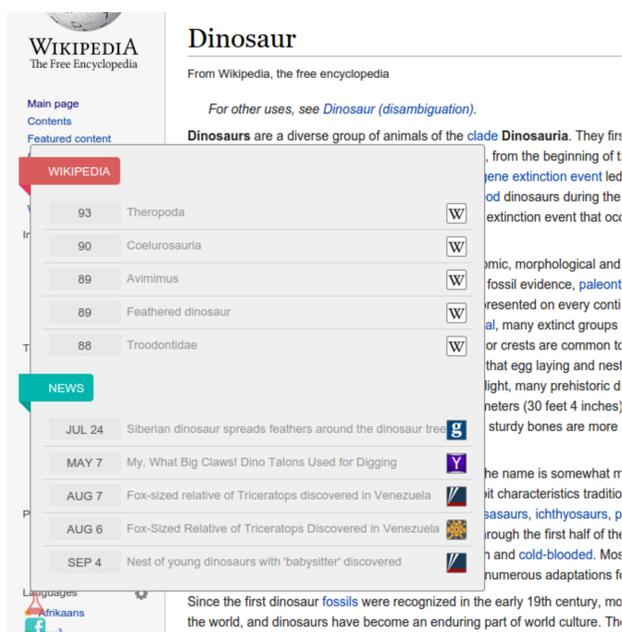- How to smartly recommend based on one-class SVM?

KVASIR

Big Data Framework

- Fast proccess large text corpus.
- Accelerate design → develop → deploy cycle.
- Cope with fast rate of user-generated content.

Machine Learning Techniques

- Latent Semantic Analysis → Content patterns.
- Random projection → Accuracy vs. Efficiency.
- Approx. nearest neighbor → Fast search.

Flexible Frontend Design

Join Us! If You Want To …

- Develop a high performance server backend …
- Contribute to the frontend enhancement …
- Learn how to build a realistic Internet service …
- See how to apply machine learning in the wild …
- Write an interesting master thesis …

- Please contact us →