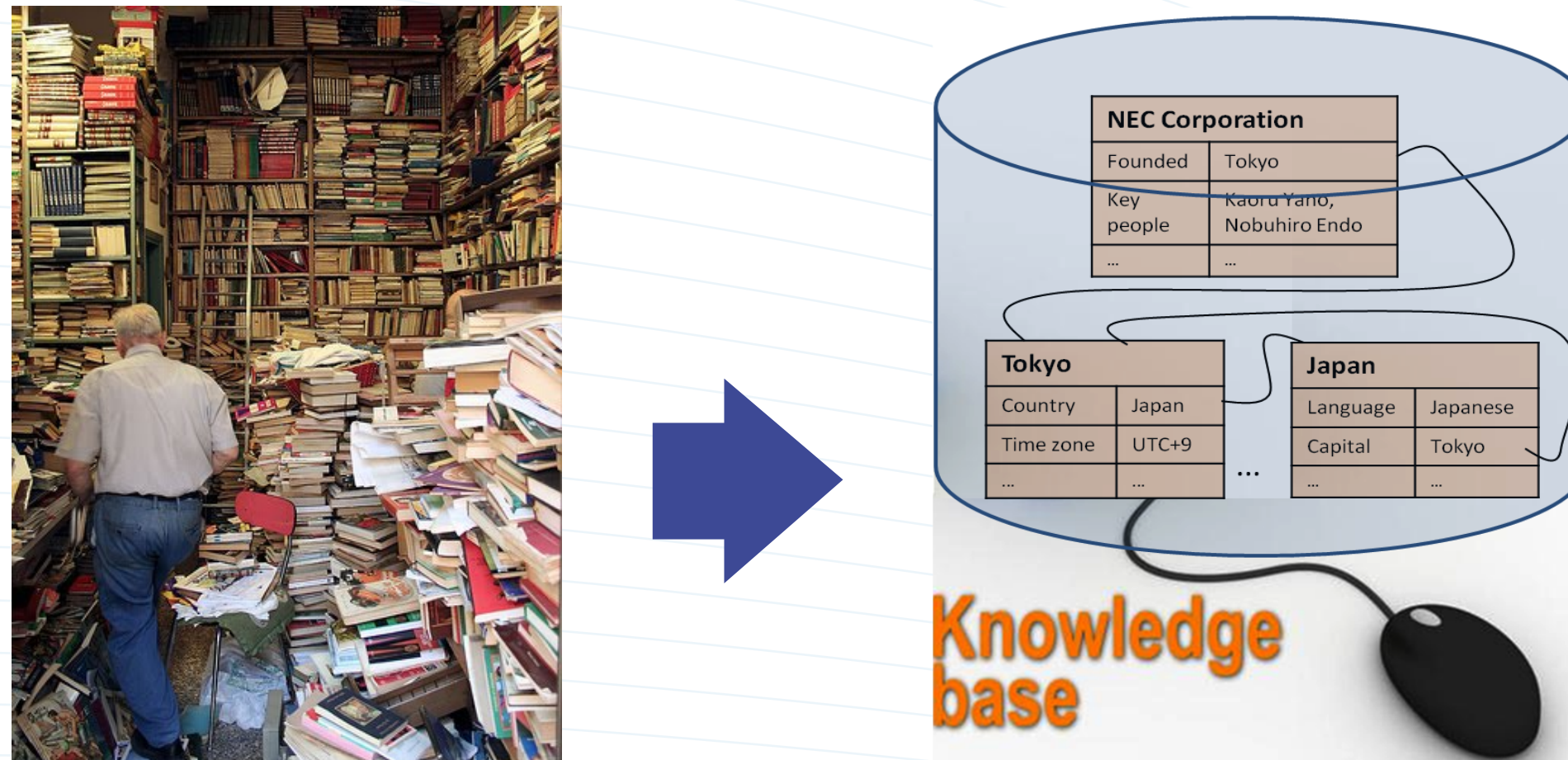
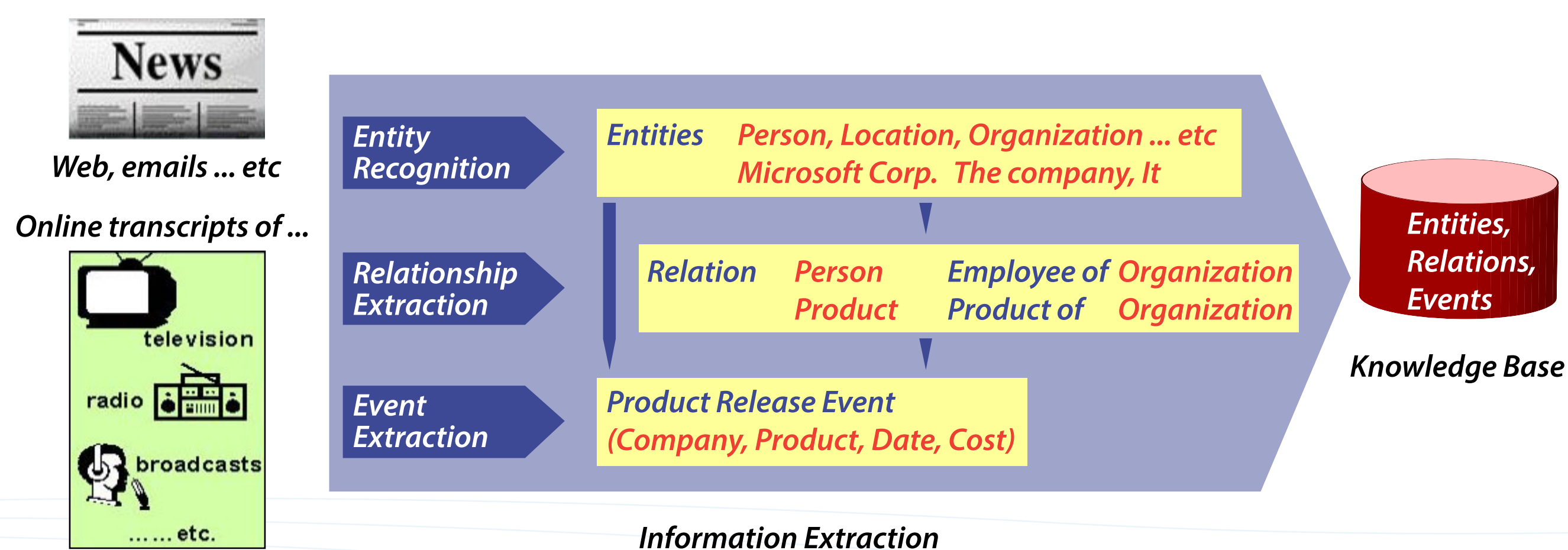


Text Information Extraction for Information Access



Technology Overview

A large amount of unstructured text data, such as news, blogs, forum postings, broadcast transcripts, and etc, is available everywhere in corporate databases and in the Web. Information extraction is a powerful technology that enables effective analysis and exploitation of such data. It extracts critical concepts, tags them with key terms, and presents them in a normalized and exploitable form.



Achievements

- One of the best performing entity linking systems among 22 international teams in Text Analysis Conference 2011 organized by NIST (USA)
- Best text classification system at BioCreAtIve II in 2007 among 19 international teams
- Best named entity recognition system at BioCreAtIve I in 2004 among 12 international teams
- Best named entity recognition system at BioNER shared task at JNLPBA 2004 among 8 international teams

Applications

- Knowledge base population
- Text classification according to entity, relation and event mentions
- Enhance and organize retrieval results
- Visualize entity relation network