
MuST: A Workshop on Multimodal Summarization for Trend Information

Tsuneaki Kato*1 Mitsunori Matsushita*2 Noriko Kando*3

*1 The University of Tokyo

*2 Nippon Telegraph and Telephone Corp.

*3 National Institute of Informatics



Characteristics

- To encouraging cooperative and competitive studies
 - Promoting discussion
 - Conforming communities
 - Constructing and accumulating resources
 - Non-evaluation workshop
 - Objective and quantitative evaluation is outside the scope
 - Shared research resource as a centripetal force
 - Participants address common or related themes using the research resource
 - (Loosely) shared theme of research
-

Talk Overview

- Subject of the MuST workshop
 - Trend information
 - Multimodal summarization
 - Research framework
 - Related research fields
 - Data set
 - Design principle
 - Specification
 - Course, current situation and future plan
 - Conclusion
-

Subjects: Trend Information

- Summarization of temporal statistical data, obtained through synthesis rather than enumeration
 - changes in product price and sales
 - public approval rating of political parties
 - Not always single-dimensional temporal information; can be multi-dimensional
 - Market share of a given product
 - Land price
-

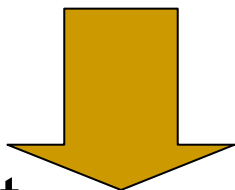
Multimodal summarization for Trend Information

Queries on trends

“How the price of gasoline shifted during the year?”

“What the situation has been in the PC market?”

“How terrible the typhoons were last autumn?”



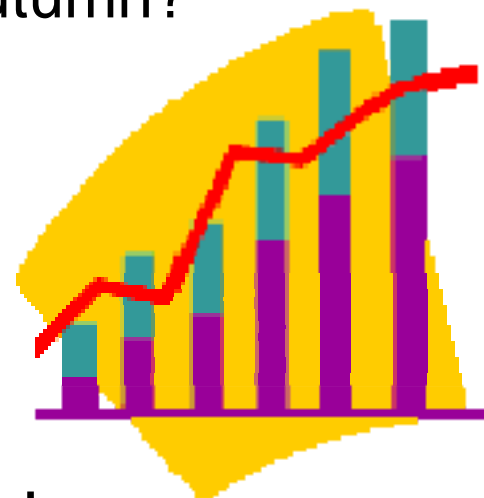
Concise, plain text

Information graphics

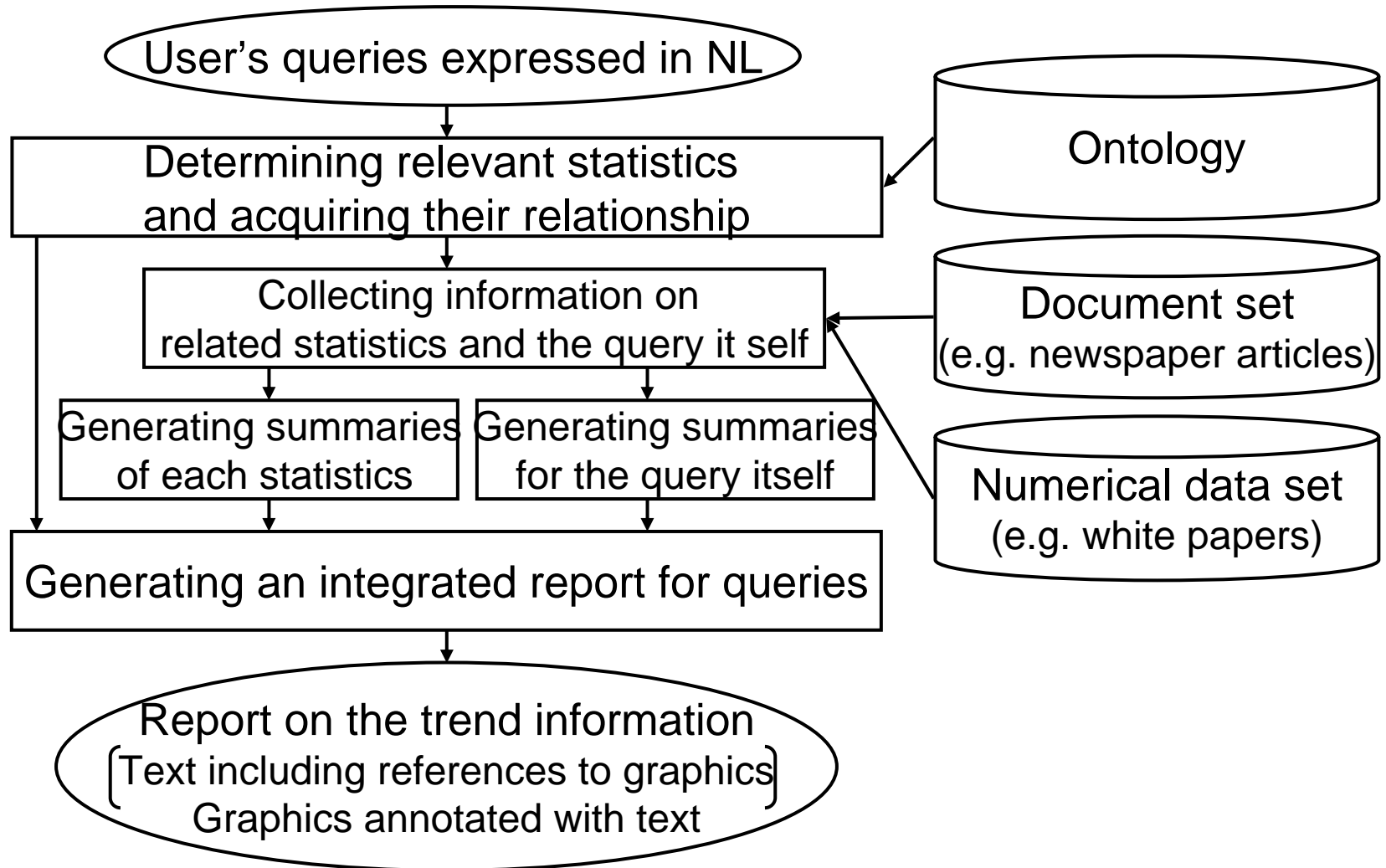
Multimedia presentation

text including references to graphics

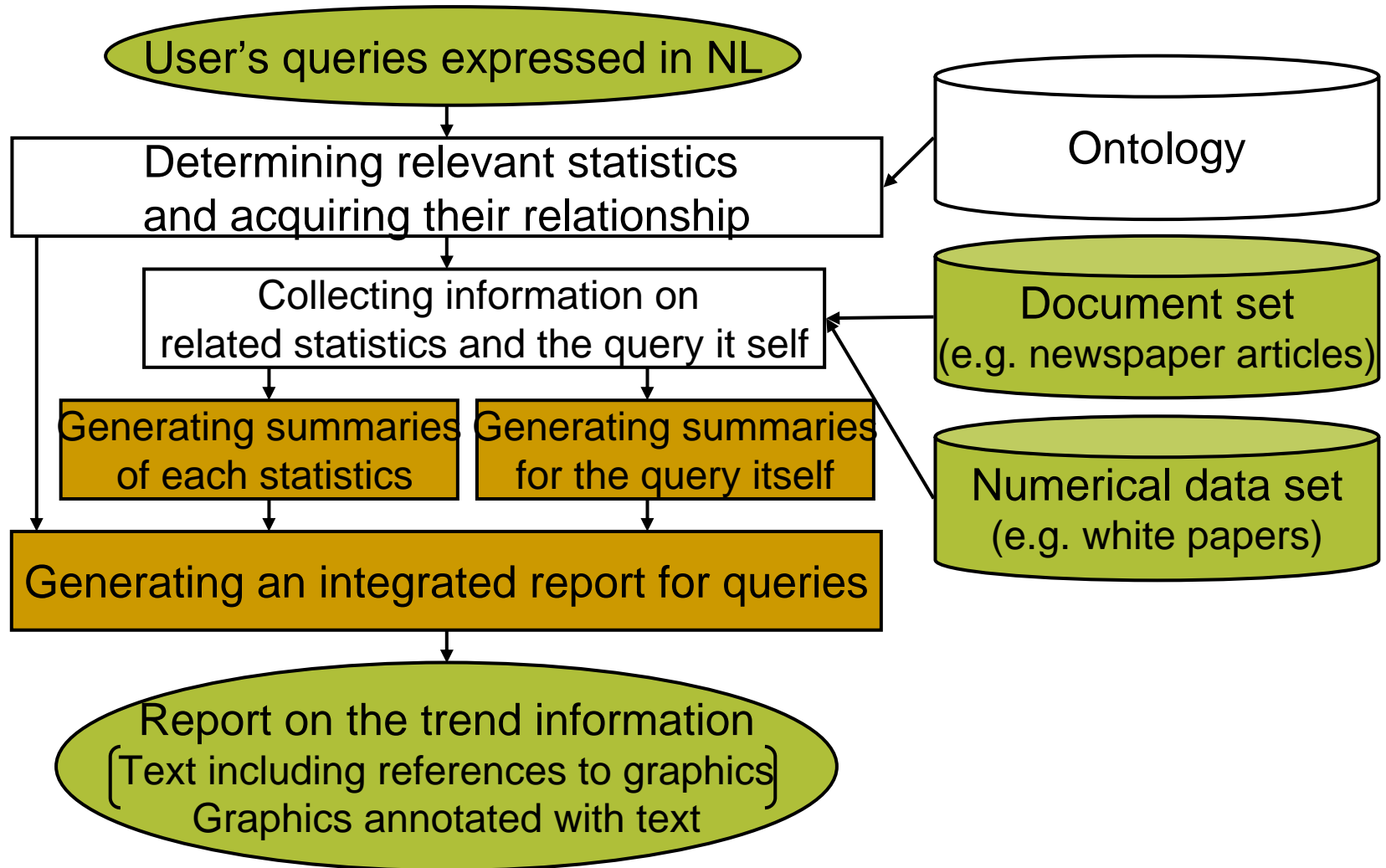
graphics annotated with text



Framework



Framework



Related Fields

- Multimedia presentation generation
 - Multimodal dialogue processing
 - Question answering
 - Text summarization
 - Information extraction
 - Information visualization
 - Human-computer interaction
 - Several elemental technologies
-

Research Resource: Data Set

- Materials to be processed
 - Examples of users' queries
 - Text information and numerical data used in order to answer those queries
 - Intermediate results
 - Extracted portions of text accompanied by a proper level of analysis
 - References for the output design
 - Sample summaries and information graphics
-

Data Set

For 20 topics,
incl. 58 statistics and event types related

Gasoline Price (Pump price of gasoline, Dubai oil price)
SONY (Consolidated net sales, Consolidated operating profit, ...)
Typhoon (Landfall event)
Home run tallies of MLB (Hitting a home run, ...)

- List of newspaper articles
 - Annotations of those articles
 - Examples of textual summaries
 - Tables or charts
-

The Roles of Data Set

Information Collected

Articles, Tables and Charts

Multimodal
Summarization



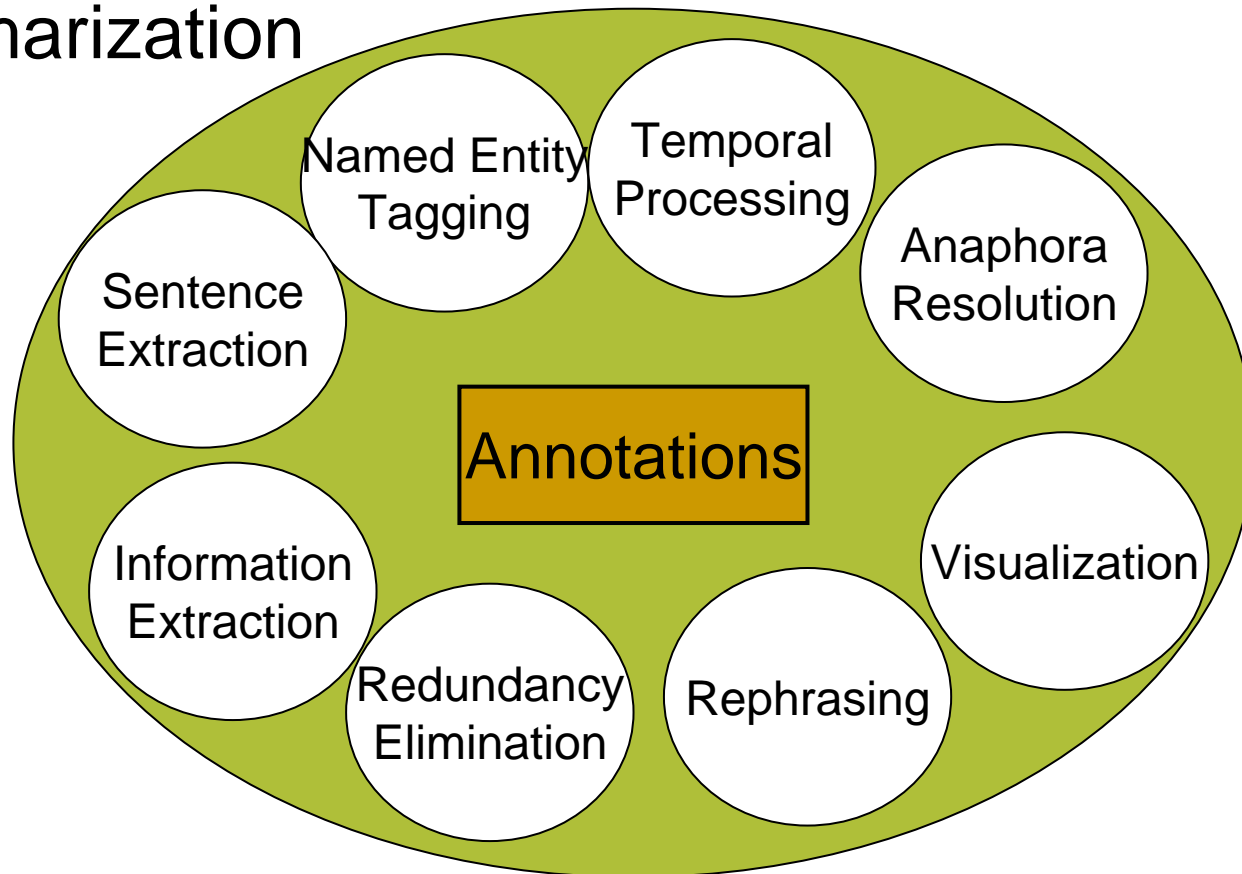
Annotations

Summaries, Reports

Textual summaries, Charts and Tables

The Roles of Annotation

Multimodal Summarization



Examples of Annotation

<unit stat="レギュラーガソリン全国平均店頭価格">石油情報センターが16日発表した給油所石油製品市況調査によると、<date gra="週" abs="19990617">今週</date>調査の<name part="head">ガソリン価格(レギュラー1リットル)</name>は<name part="foot">全国平均</name>で<val>92円</val>となり、<date gra="週" abs="19990610">前週</date>の<name part="foot">平均</name>に比べ<rel>1円</rel>上昇した</unit>。<unit stat="レギュラーガソリン全国平均店頭価格"><name part="foot">全国平均</name>の上昇は<date gra="旬" abs="19990201">2月上旬</date>以来、<dur gra="月">4カ月</dur>ぶりだ</unit>。

Annotation on English Text

<unit stat="nationwide average of pump price of gasoline"><name part="head">The price of gasoline (one liter, regular)</name> reached <name part="foot">a national average of</name> <val>92 yen</val>; <rel>1 yen</rel> higher than the <name part="foot">average price </name> of <date gra="week" abs="19990610">last week </date></unit>.

<unit evet="typhoon landfall">Medium-strength <par>typhoon No. 10 </par> struck <par>Makuraszaki-shi, Kagoshima</par>, at <date gra="hour" abs="19980917">about 4:30 pm on the 17th</date>, and will strike in <par>the vicinity of Shukumo-shi, Kochi</par> <date gra="hour" abs="19980917">the same night</date></unit>.

Course, Current Situation, Future Plan

- Announced in Nov. 2004
 - CFP started at the beginning of 2005
 - Data set was started to deliver in Jan./April
 - Round table meeting in Nov. 2005

 - 14 participants
 - Two MLs:
 - closed to the participants
 - open widely for interested persons (about 90)

 - Closed workshop on March and several opportunities for research presentations
-

Conclusion

- MuST workshop was designed to encourage cooperative and competitive studies on summarization and visualization for trend information
 - Participants share the same research source, whereby they address common or related themes
 - MuST started at the beginning of 2005, and the first period come to an end in March 2006
 - Looking toward the next period, efforts are being made to establish the course for encouraging further research
-