

A collection of historical artifacts is displayed on a light-colored surface. On the left, there is a blue leather-bound book or album with a checkered pattern of small circular objects. Several medals are pinned to the book, including a red ribbon medal with a circular emblem, a silver star-shaped medal with a central emblem, and a blue ribbon medal with a circular emblem. A pair of gold-rimmed glasses with a thin wire bridge is placed over the book. In the bottom left corner, a circular compass with a white face and black markings is visible. The background is a plain, light-colored surface.

Jumping Off Points

Ideas of possible tasks

Examples of possible
tasks

Categories of possible
tasks



Ideas for Evaluation Tasks

- ◆ Possible Outcome-based Tasks
- ◆ MT in the Real World
- ◆ Evaluation of MT Components
 - Linguistic components
- ◆ Evaluation of MT as Software



Possible Outcome-based Tasks

- ◆ Combining MT and NLP technologies
- ◆ Measuring ability to use translated documents for specific tasks
- ◆ Combination of MT and Translation Memory
- ◆ Finding all relevant and important information on topic X



Other Toys to Play With

- ◆ Information extraction
 - Named entity extraction
- ◆ Information retrieval
- ◆ Summarization
- ◆ Topic detection and tracking
- ◆ Speech translation
- ◆ Process ordering with the above



Use of Translated Documents

- ◆ Binning
- ◆ Gisting
- ◆ Other DARPA Task Proficiency Scale tasks
 - Question answering / reading comprehension



MT in the Real-World

- ◆ MT with inconsistent data
 - What are the effects of data clean-up?
- ◆ Looking at data normalization
 - Are some more important to handle pre-processing?
 - Can some be handled lexically
- ◆ How can we evaluate the marriage of users and systems?



Evaluating Components

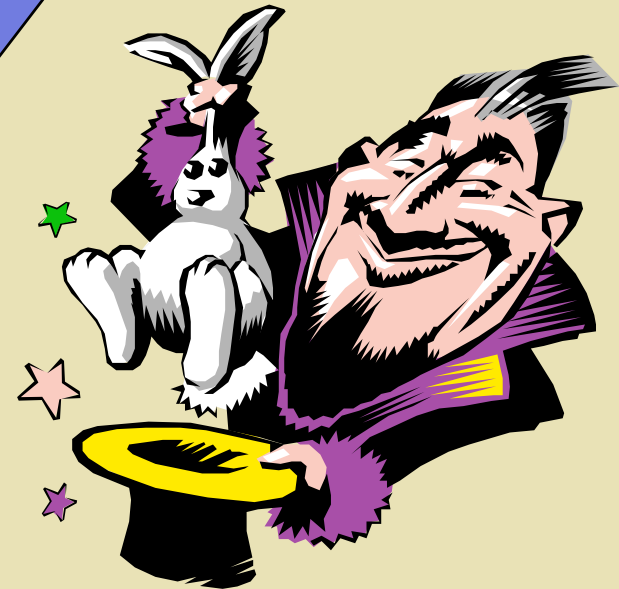
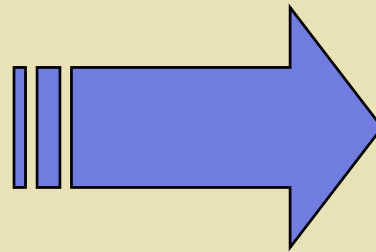
- ◆ MT across language families
- ◆ Evaluation of alternative architectures / components
- ◆ Using language learner evaluation techniques on system output
 - Measuring vocabulary coverage
- ◆ Foreign language term extraction to evaluate MT system output
- ◆ Finding appropriate corpora and methods for feedback



This Is Software After All

- ◆ Ease of use
 - i.e., Lexical update
- ◆ Variety of input types, output modes
- ◆ Failure boundaries

This is All Wrong?





Possible Data Sets

- ◆ Chemical Weapons Treaty
 - French, Spanish, Arabic, Chinese, English, Russian
- ◆ DARPA 1994 Data
 - Spanish => English :: 2 HT + 6 MT
 - French => English :: 2 HT + 5 MT
- ◆ Edupage - German source, can find English and other languages
- ◆ Some things we can find
 - CDC Immunization Sheets
 - More Edupage
 - TELRI Corpus

Special Requests?

