

# Machine Translation (MT)

## What do we do at NIST?

### Evaluate Machine Translation Technology

*the use of software to automatically translate from one language to another*

#### *technologies evaluated*

##### Text-to-Text



##### Speech-to-Text & Speech-to-Speech



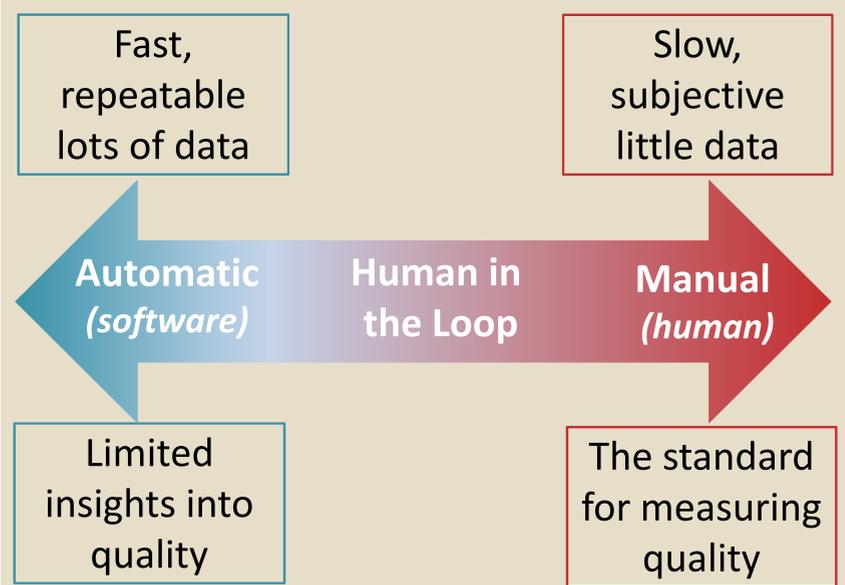
##### Image-to-Text



### Improve MT Measurement Methods

*the degree to which resulting translations are appropriate for their intended use*

#### *continuum of metrology approaches*



Accomplished through open evaluations using common data, shared resources, and appropriate metrics

### **The problem**

- Measuring technologies that automatically translate from a source language into a target language (*output must be both accurate and fluent*)
- Developing and assessing metrology techniques that are used for MT evaluation

### **This is hard because**

- Measuring MT quality is not trivial (*creating reference translations is not deterministic and results in subjective measures of quality*)
- Consistency and repeatability in measurement are issues
- Measurement techniques differ for different language pairs, input mediums, and data styles

### **NIST continues to work on this problem, we**

- Design and implement open MT evaluations using standard data sets and appropriate metrology (*NIST Open MT and NIST Metrics for Machine Translation Challenge*)
- Serve as technical experts for DoD MT-related programs, for evaluation design and implementation (*DARPA GALE, TRANSTAC, and MADCAT programs*)
- Chair an MT metric development working group (*U.S Army's MFLTS program a.k.a. Sequoyah*)

### **Demo**

Speech-to-Speech translation technology

Developed through the DARPA TRANSTAC evaluations

Enhances communication between U.S. Soldiers and Iraqi Civilians

Primary evaluation metric developed by NIST – bi-lingual judges assess “the odds of successful concept transfer”