

# MOLTO: Results and Future Prospects

Aarne Ranta

Final presentation, Barcelona 23 May 2013

**MOLTO**  
FP7-ICT-247914



# Multilingual Online Translation

Non multa, sed multum not quantity but quality

ABOUT

NEWS

EVENTS

MOLTO's mission is to develop a set of tools for translating texts between *multiple languages* in *real time* with *high quality*. MOLTO will use multilingual grammars based on semantic interlinguas.

FP7-ICT-247914, Strep, [www.molto-project.eu](http://www.molto-project.eu)

U Gothenburg, U Helsinki, UPC Barcelona, Ontotext (Sofia), U Zurich,  
Be Informed (Apeldoorn)

March 2010 - May 2013

EC contribution 2,975,000 EUR

**What we promised**

## Molto's goal

- **high-quality** translation
- between **many languages**
- on the **web**
- with **real-time** updates

## What's different?

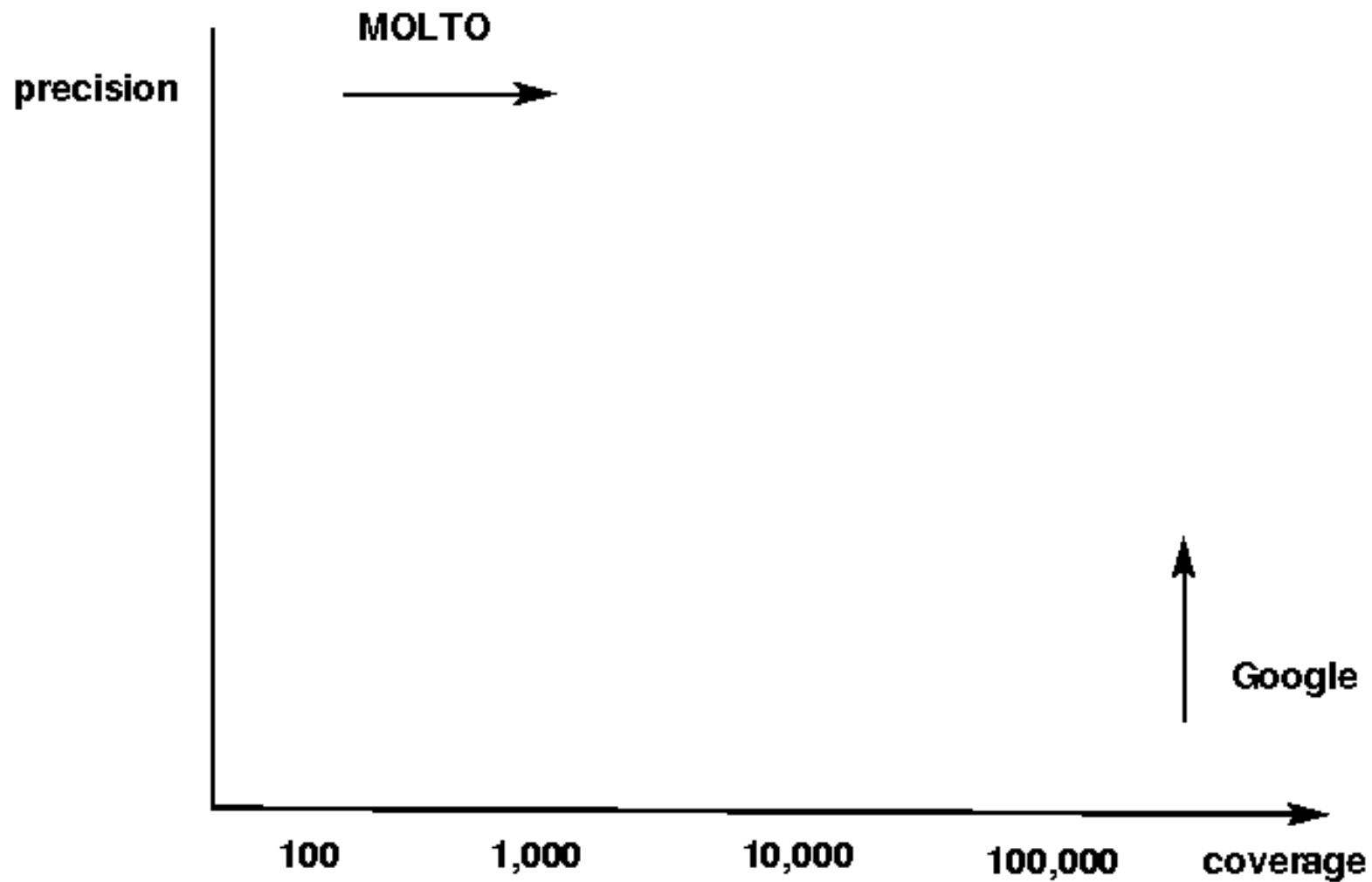
<b>Tool</b>	<b>Google, Bing, Babelfish</b>	<b>MOLTO</b>
target	consumers	producers
input	unpredictable	predictable
coverage	unlimited	limited
quality	browsing	publishing

## **The promise**

Define the content you want to translate...

...and you can do it efficiently and reliable with the MOLTO tools.

# Scaling up



**What we achieved**



# Translation among 15+ languages

Planned in the proposal

- Bulgarian, Catalan, Danish, Dutch, English, Finnish, French, German, Italian, Norwegian, Polish, Romanian, Russian, Spanish, Swedish

Also figuring in final systems

- Hebrew, Hindi, Latvian, Persian, Thai, Urdu

# Translation on many domains

- mathematics
- painting descriptions
- tourist phrasebook
- patent claims and queries
- Wiki: geographic facts

# **Tools available on the web**

Translation tools

System building tools

Collaborative editing (Wiki)

# Rendering semantic content

Ontology

Controlled language

Disambiguation

# Information retrieval

Given a MOLTO translation system...

...a query system is easy to build on top.

# Grammar engineering productivity

**Anyone can build a translation system** - with a couple of days' training.

Building a system is a matter of **days**.

Adding a language is a matter of **hours** - and needs only a couple of hours' training.

# Scaling up system size

Automatic lexicon extraction, from

- Wordnets
- Wikipedia
- terminologies
- SMT phrase tables

Hybrid SMT + GF

- the best patent translator of today?

Efficient, robust, statistical GF parser

- already competitive for less resourced languages (Urdu, Bulgarian, Finnish)

# Impact

## Research:

- controlled languages
- semantic web
- natural language generation
- (less in main-stream MT)

## Industry:

- in MOLTO: Ontotext (Bulgaria), Be Informed (The Netherlands)
- outside: Galois Connection (USA), Lingsoft (Finland)



# The presentations

## **Core technologies**

Grammar Engineering Tools (Gothenburg)

The GF Cloud (Gothenburg)

Lexicon Extraction (Helsinki, Gothenburg)

Query Technologies (Ontotext, Gothenburg, UPC)

## **Emerging technologies and research aspects**

Robust and Statistical Parsing in GF (Gothenburg, Helsinki)

Hybrid SMT/GF Patent Translation (UPC, Gothenburg)

# Applications of MOLTO Technologies

Multilingual Mathematics Online (UPC)

The Multilingual Online Museum (Gothenburg, Ontotext)

Multilingual CNL-based Semantic Wiki (Zurich)

Multilingual Business Logics (Be Informed)

# Evaluation and Future

Evaluation of MOLTO Technologies (Helsinki)

MOLTO Exploitation Perspectives (Ontotext, Be Informed)