



# 2010 the Industrialisation of MT

**Fred Hollowood**

Research Director (SES)

# Sample Agenda

- 1 Symantec – MT History
- 2 Choices
- 3 Resources
- 4 Customization
- 5 Measurement

## Symantec – MT History

- Engaged with RBMT in 2004
- Worked on Controlled Language to prepare our English source
- Use post-editing services to bring content to commercial quality
- Publish product documentation through MT in seven languages
- Engaged with SMT in 2008
- Run a small research team in collaboration with DCU and CNGL
- Stress the importance of research to further the application of the technology

**MT is not for the faint hearted**

# Choices available now

- Post-editing Services
  - Moved from the unknown to a commercial offering
  - Both process and technology need refinement
  - Educational engagement is still nascent
- MT Systems
  - “Free” systems appear regularly
  - Custom systems offered commercially
  - You pay for what you get

# Resources

- Hardware
  - MIPS are cheaper than ever before
  - Available for hire in the Cloud
- Finding Specialists
  - Internal Staff
  - Vendors
  - Crowd-sourced



# Customisation

- Domain
  - Machine domain tuning necessary for commercial viability
  - Defining and cleaning training material needs to be automated
  - Flexible leverage of internal tags for disambiguation
  - Broad domain ability required for some applications
- Customer
  - Personalised domain tuning for multilingual community interaction

# Measurement

- Human versus Automatic
  - Automatic measurement gives us precision (Bleu, Meteor, GTM,)
  - False sense of security
  - Human perception of language improvement is a requirement
    - Coarse
    - Individual
    - Expensive to survey
- Challenges
  - Cost of provision Words/\$ ,Words /CPU cycle
  - Measuring real-time MT provision      Words/S
  - Part of an industrial process – embedded in a workflow
  - Scalable, Affordable, Adaptable





# Thank you!

Fred Hollowood

[fhollowood@symantec.com](mailto:fhollowood@symantec.com)

+353 862634817

**Copyright © 2010 Symantec Corporation. All rights reserved.** Symantec and the Symantec Logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

This document is provided for informational purposes only and is not intended as advertising. All warranties relating to the information in this document, either express or implied, are disclaimed to the maximum extent allowed by law. The information in this document is subject to change without notice.