META NET

Preliminary Findings of the Vision Group Translation and Localisation

META VISION

Josef van Genabith

**Centre for Next Generation Localisation CNGL
Dublin City University**

META-COUNCIL 2010, Brussels

About the Speaker

- Josef van Genabith

- Director: Centre for Next Generation Localisation (CNGL)
- Professor School of Computing, Dublin City University (DCU), Ireland
- Founding member of META

Vision Group
Translation and Localisation

❑ **Fields:** Technical documentation, consumer information, official bulletins, user interface localisation, translation services, information gathering, **user-generated content** etc.

❑ **Stakeholders:** Software companies, large users of translation and localisation services, Language Service Providers, translators, customers, consumers, citizens etc.

❑ **Organizers:**
- Hans Uszkoreit (DFKI, Germany)
- Josef van Genabith (DCU/CNGL, Ireland)

❑ **Meetings:**
1. Berlin, 23 July 2010
2. Brussels, 29 September 2010

Needs

Demand and Observations

META VISION

- Translation is an emotional topic that needs to be handled with care.

- Despite progress, the problem of HQMT remains unsolved, often "quality" still is a major problem.

- Translation is not only a linguistic business:
  - Context and common sense knowledge play a decisive role.
  - The value is in the information, entertainment, etc., not in the translation itself.

- Translation workflows and business models are not up-to-date:
  - Data and software migration are very difficult.
  - In some areas too many documents are translated with no real need.
  - *A lot of money is wasted non-productively!*
  - Customers self-service on WEB, user-generated data
  - Lack of personalisation
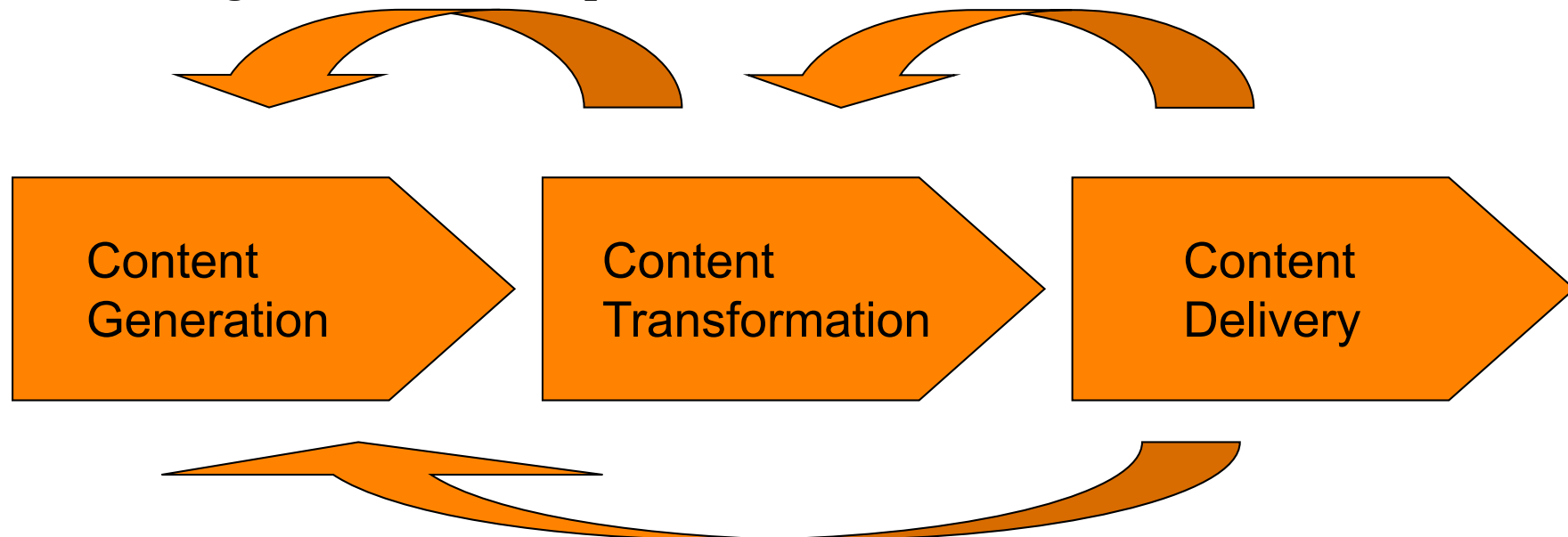 - The mobile WEB

Visions

New Workflows

META VISION

Different demands require **customised solutions**, e.g.,:

	Time	Costs	Trust/ Security	Quality	Personali- sation/ corporate ID
Outbound high quality translation (corporate, manuals, PR material, etc.)	Not critical	Can be costly	Medium	Highest possible	Important
Inbound cross-lingual information gathering	Must be fast	Must be affordable	High	Medium	Not needed
Translation for private users (online content, communication)	Near realtime	Must be cheap	Medium to Low	Medium	Would be a plus

New Workflows ctd.

Example of a new workflow: **going upstream**

- Many (if not most) problems arise from the poor quality of source texts to be translated (spelling and grammar errors, ambiguity, missing context, terminology etc.)

- User-generated content …

- Integration of content production and translation on demand

| Content Generation | Content Transformation | Content Delivery |

New Architectures

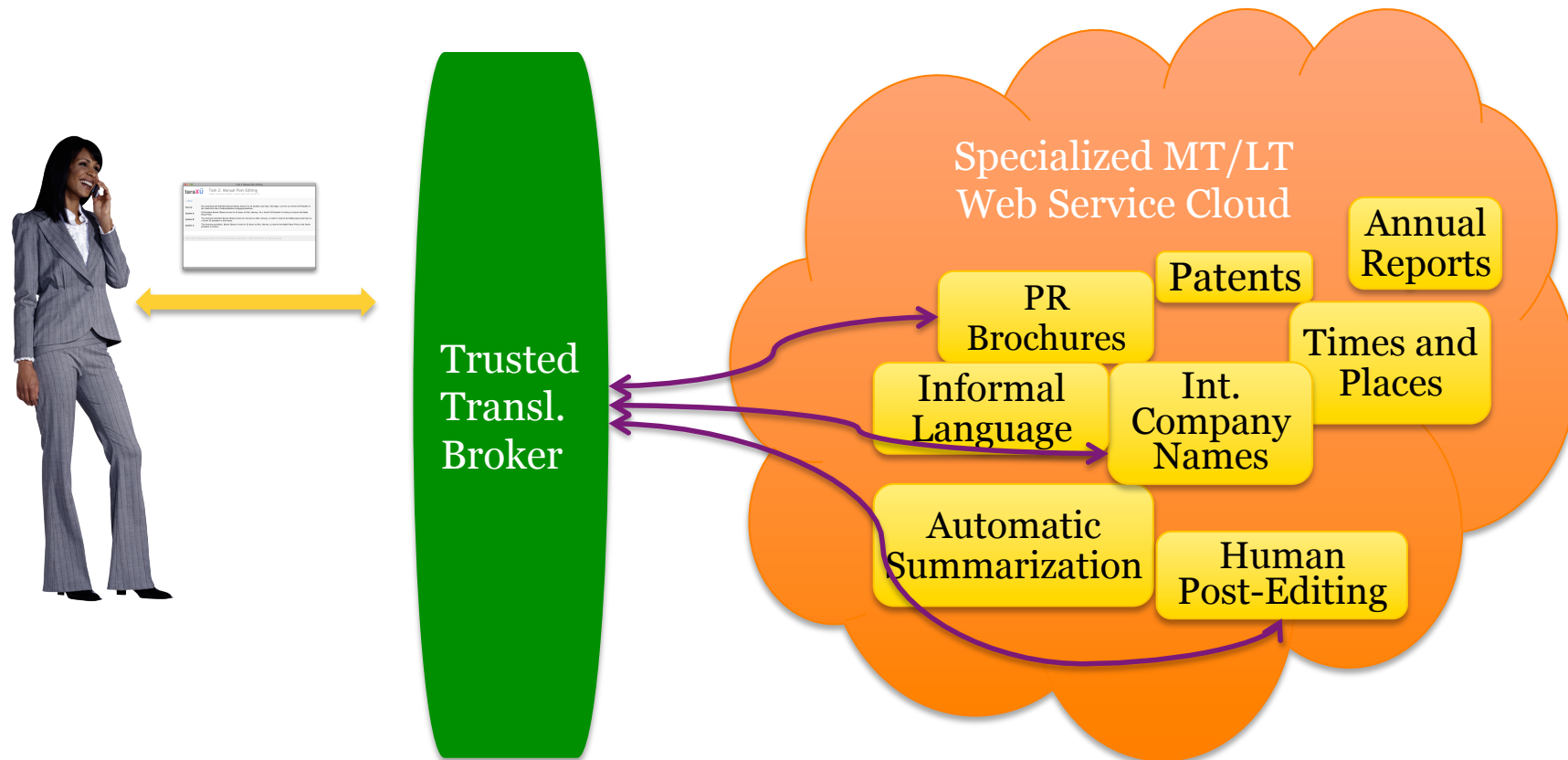- **New MT technology and infrastructure**
  - Drawing on (artificial) knowledge heavily (user needs, context, task, ambient semantics, web, etc.) with machine learning
  - Sharing of data and resources: Standards, interoperability, agile workflows!
- **Translation brokering**
  - Brokering service sends input to domain/language/style/genre/corporate/user/etc. adapted services
  - Cloud of highly specialised LT/MT services
  - Interoperable components, connected by highly adaptive workflows
  - SMEs as well as big companies offer/use specialised LT/MT services
  - On demand
  - Transparent cost calculation

Translation Brokering

META VISION

Trusted Transl. Broker

Specialized MT/LT Web Service Cloud

Annual Reports

Patents

PR Brochures

Times and Places

Informal Language

Int. Company Names

Automatic Summarization

Human Post-Editing

Research and Training

- **New Human-Centred Research Paradigm in MT:**
  - Human Factors:
  - Truly hybrid processes of machines (robots?) and humans
  - Humans as providers of data, insights, quality judgements, critique, etc.
  - Humans as test users and evaluators of early MT prototypes

- **New training for pre- and post-editing** of MT texts (missing in current curricula for translators)

- MT/LT students should be trained in translation.

Topics with Visionary Potential META=VISION

- **Domain specific**
  - More knowledge sources in MT (and machine learning)
  - New modes of research studying and involving translating humans
  - A new role for the human in the machine translation process

- **Domain independent**
  - New, transparent business models: mobile WEB, user-generated content
  - New computing/deployment models/workflows:
    - Interoperable plug & play components
    - Adaptive, agile workflows
  - New ways of sharing resources and knowledge in R&D

A Final Vision:
Ambient Translation Projection