



# The Quaero Program

**Edouard Geoffrois**  
Quaero Technical Director

LREC  
Malta, May 20th, 2010

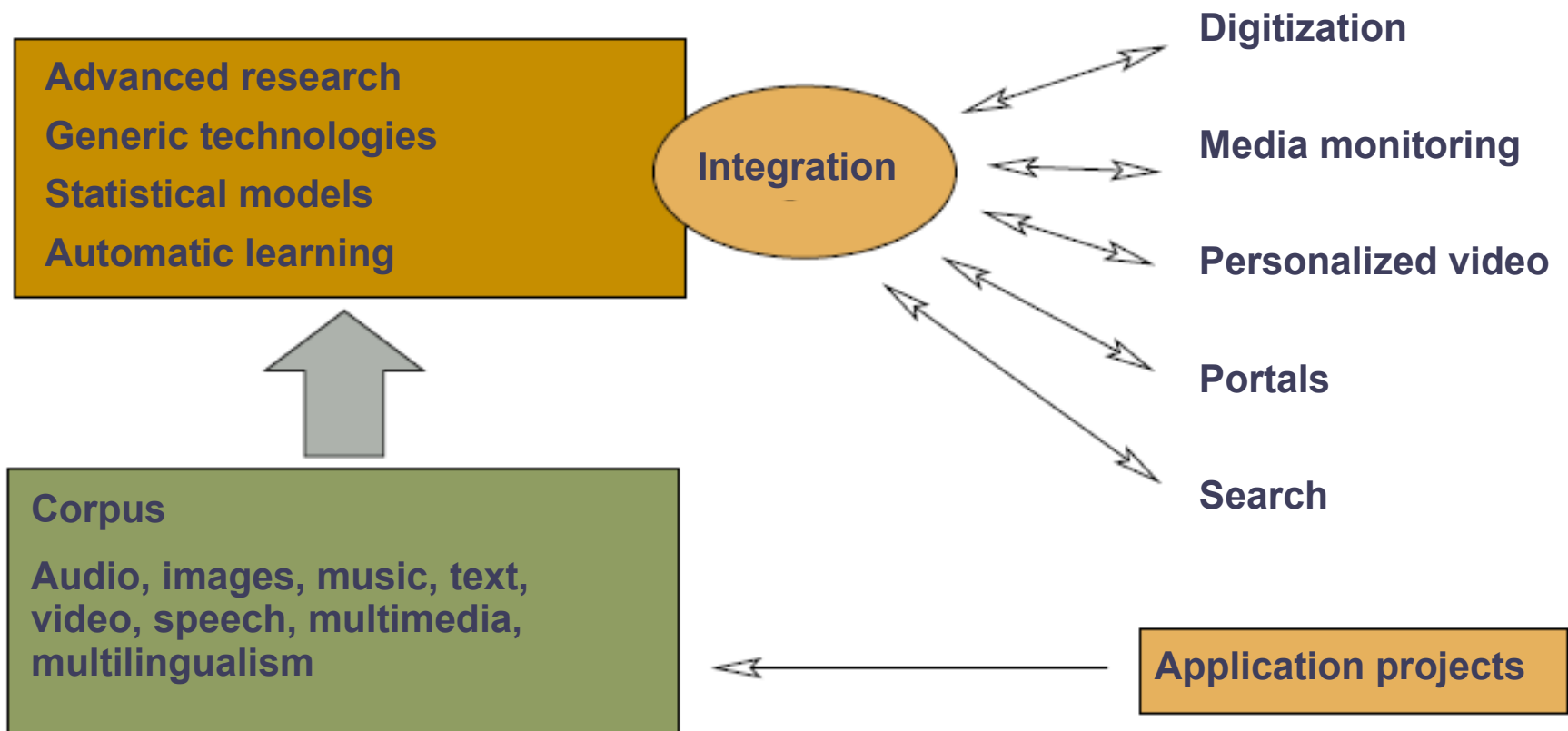
# Quaero in a nutshell

- A program on multilingual and multimodal document processing,
- connecting research and innovation through data and evaluation
- Critical-mass, flexible and extensible
  - 200 M€ budget, 99 M€ funding
  - 5 years starting May 2008
  - 26 partners, public & private, France & Germany

# A close interaction between research and innovation

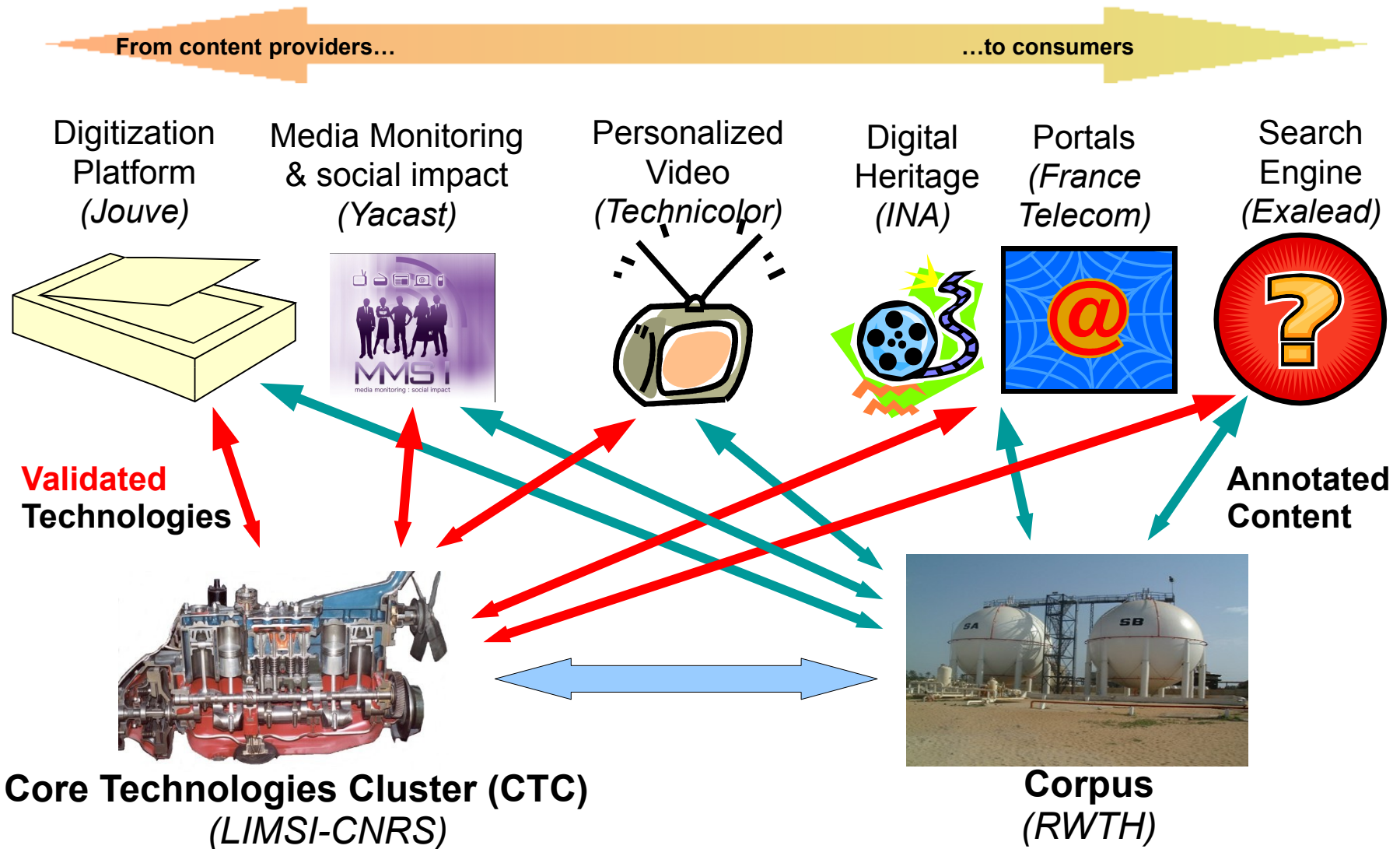
- **Seamless technology transfer** between research and application across partners
  - Develop demonstrators aiming at identified business targets
    - Involve business units in industrial partners
  - Increase the state of the art in associated technology domains
    - Ensure delivery of technology modules
- **Systematic evaluation** of results
  - Provide a dedicated evaluation infrastructure
  - Organize campaigns on a yearly cycle
- Strong investment in **large representative corpora**
  - Directly derive data from application whenever possible
  - Devote significant effort into data annotation

# Development cycle



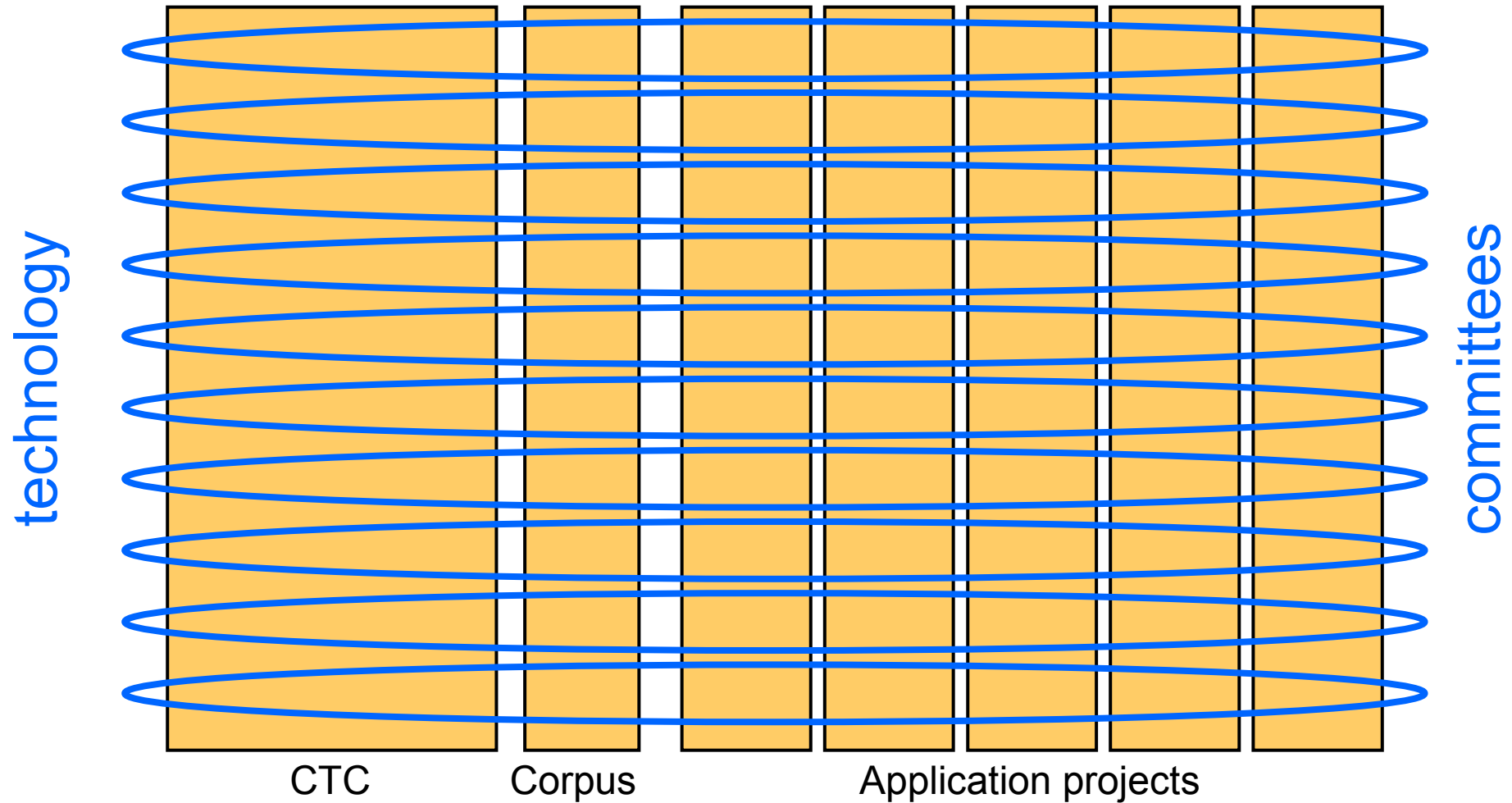
# Overall structure

several application projects sharing core technologies and resources



# Technology committees

to coordinate work along the value chain for each technology domain



# Covered technologies

Search and navigation

Written language processing

Translation

Speech processing

Audio & music processing

Scanned document processing

Image processing

Video processing

Multimodal fusion

Coding & protection

# Partners

- Private enterprises

- Bertin technologies, Exalead, France Telecom, Jouve, LTU Technologies, Synapse Développement, Technicolor, Vecsys, Vecsys Research, Yacast



- Public Research Labs

- CNRS-IMMI, CNRS-LIMSI, CNRS-INIST, INRIA, IRCAM, IRIT, Institut Telecom, LIPN, INRA-MIG, Université Joseph Fourier, Karlsruhe Institute of Technology, RWTH Aachen



- Public Institutions

- BnF, DGA, INA, LNE



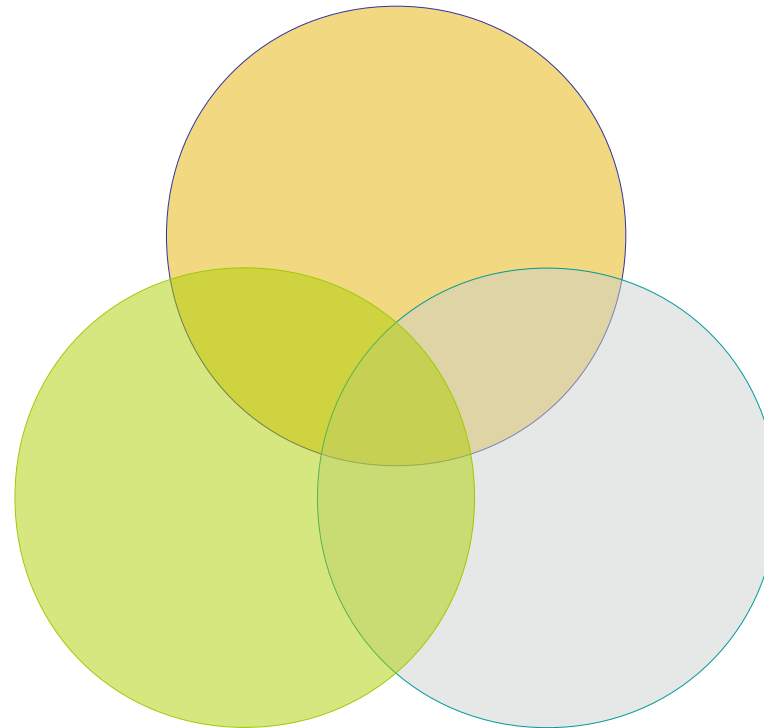
# Integrated program management team



technicolor



**Program Leader  
Program Office Manager**

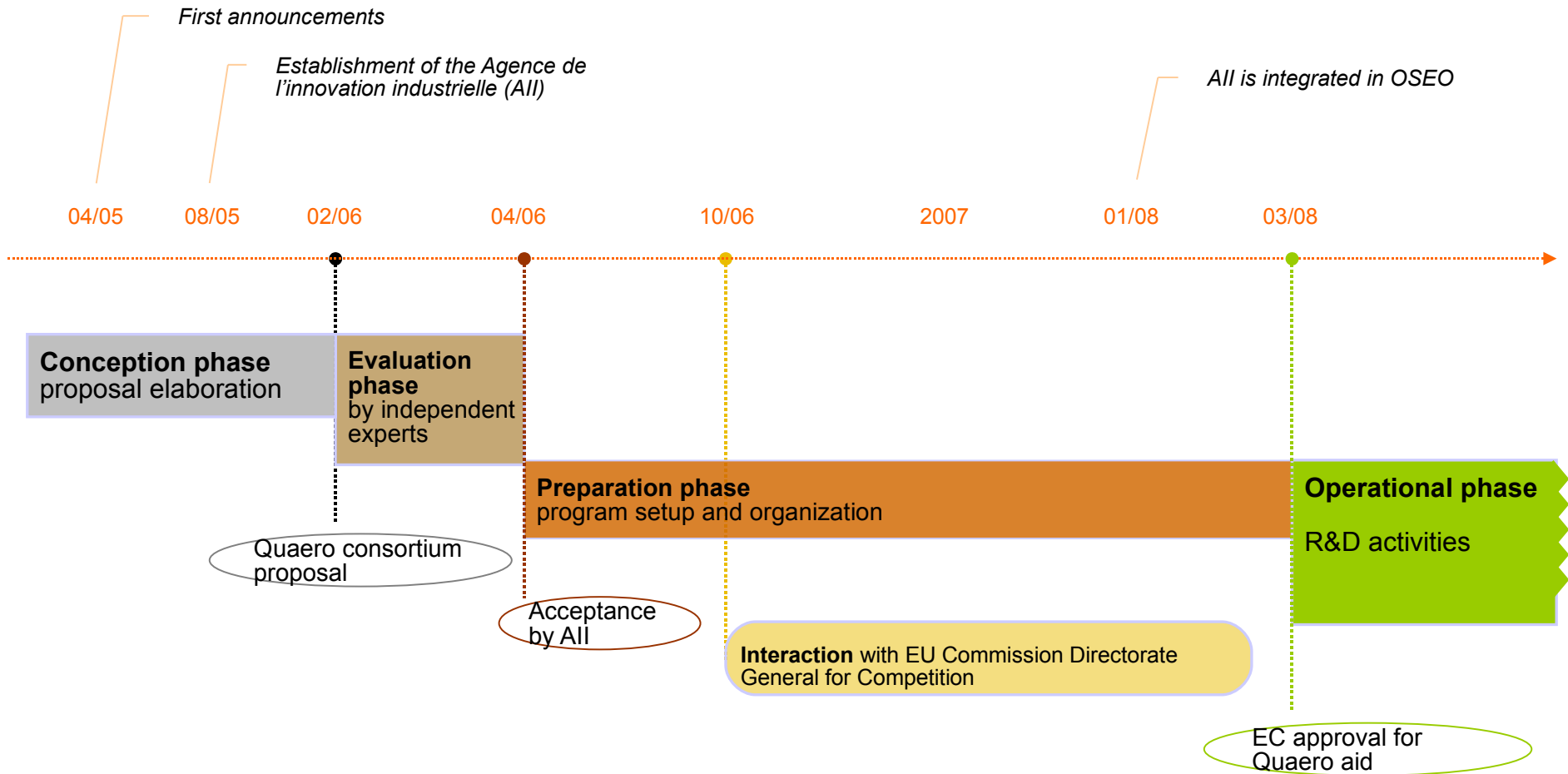


# Evaluation infrastructure

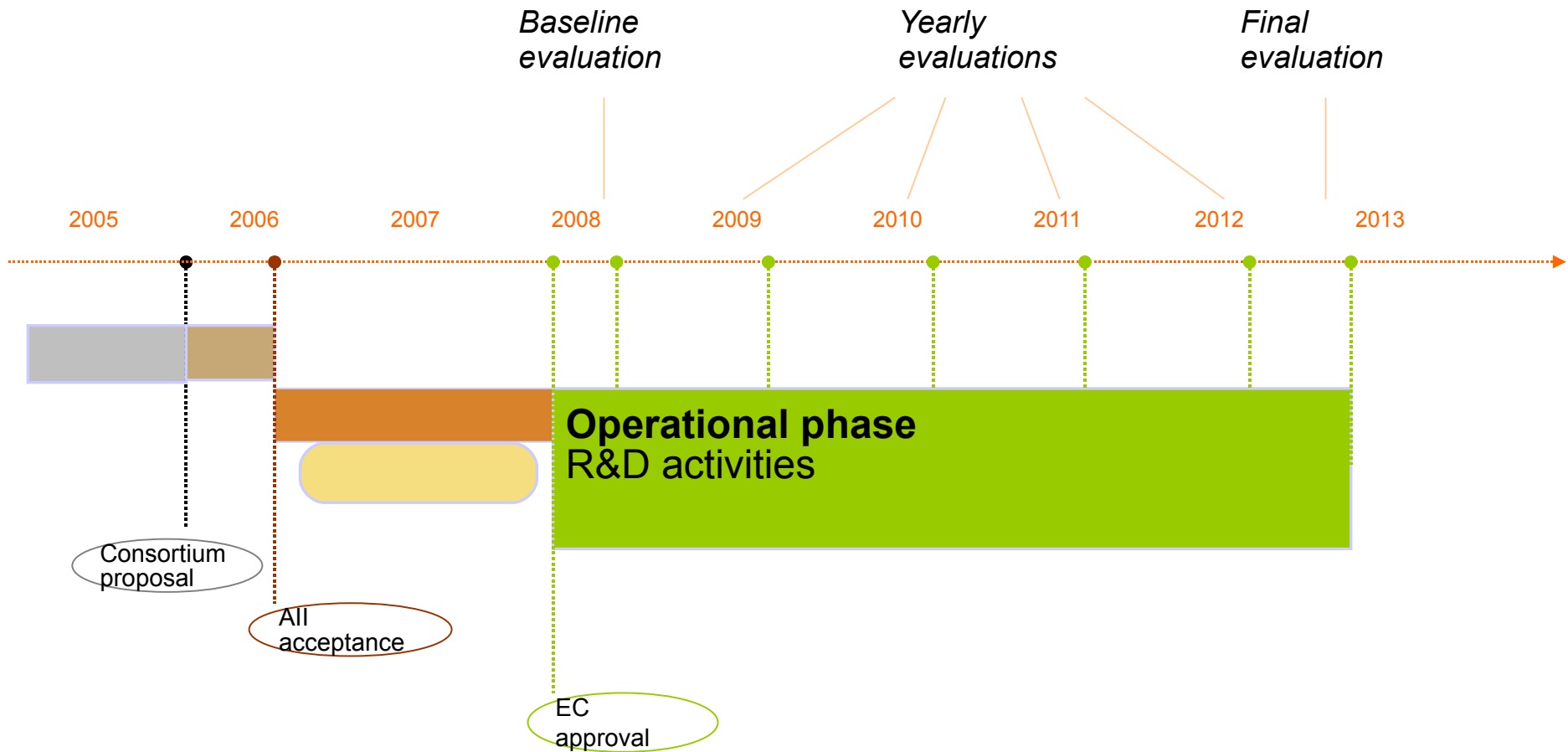


- All technology developments organized by tasks and sub-tasks
- About 20 persons involved (~ 8 FTE) for the program duration, distributed over 6 partners (DGA, LNE, IRIT, UJF/LIG, CNRS/GREYC, LIPN)
- Tightly coordinated with corpus creators, technology developers and integrators with the technology committees
- Organizing evaluation campaigns on a yearly basis

# History




# Timeline



# Funding

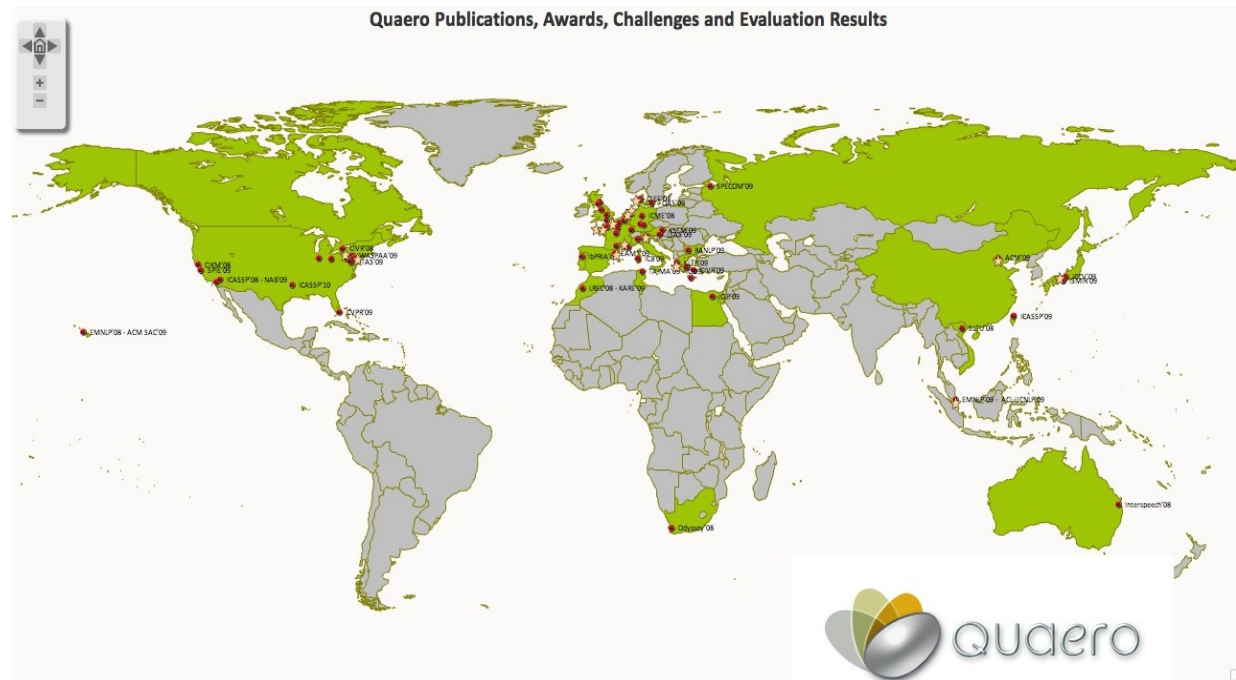


- 200 M€ budget over 5 years (~40 M€ / year)
  - 99 M€ of subsidies and reimbursable loans from French agency for industrial innovation 
  - Complemented by the partners
- Flexibility
  - Possibility to adjust the workplans
  - Continuous monitoring
- Extensibility
  - Scheme to collaborate with associate partners bringing in new applications with corresponding data or new technology capabilities
  - First associated partner joined in 2009 (Priberam, Portugal), contacts under way with other potential associated partners

# Some scientific and technical achievements

- 40 tasks evaluated internally
- 12 participations of partners to international evaluations, always in the top 3

- More than 200 publications



- 19 awards

# Some applications integrating speech & language technologies



- VoxleadNews
  - <http://voxleadnews.labs.exalead.com>
- France 24 HD Player
  - <http://lab.france24.com/hdplayer>
- Presidency web site
  - <http://www.elysee.fr>
- FM Premium



# Wide media coverage





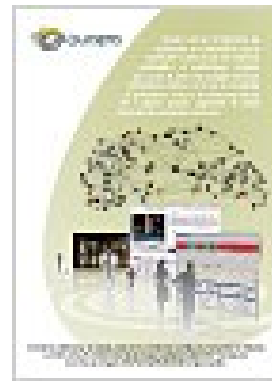
# Further information



- Web site : <http://www.quaero.org>



- Brochure



- Technology catalog



# Thank you for your attention!