Multilingual Information Service System for the Beijing Olympics

——— a playground of MT technologies

Presentation on the MT Summit **X**

Liu, Weiquan

Beijing Network and Multimedia Laboratory

September 13, 2005 Phuket, Thailand

Beijing 2008

Topics

- Overview
- Pilot systems
- Multilingual Information Services for Olympics
  - Mission & objectives
  - Goals
  - Roadmap/schedule
  - User population/scale
  - Services specification
- The role of MT technologies

Overview

The Beijing Olympic Games Bidding Committee made this promise:

"During the Olympic Games, any one with any device could receive messages of the games at any time and any where."

In the *Action Plan for the Beijing Olympic Games*, it reads:

"…to ensure that any body at any time and in any place can enjoy the information service that is affordable, comprehensive, personalized and multilingual…"
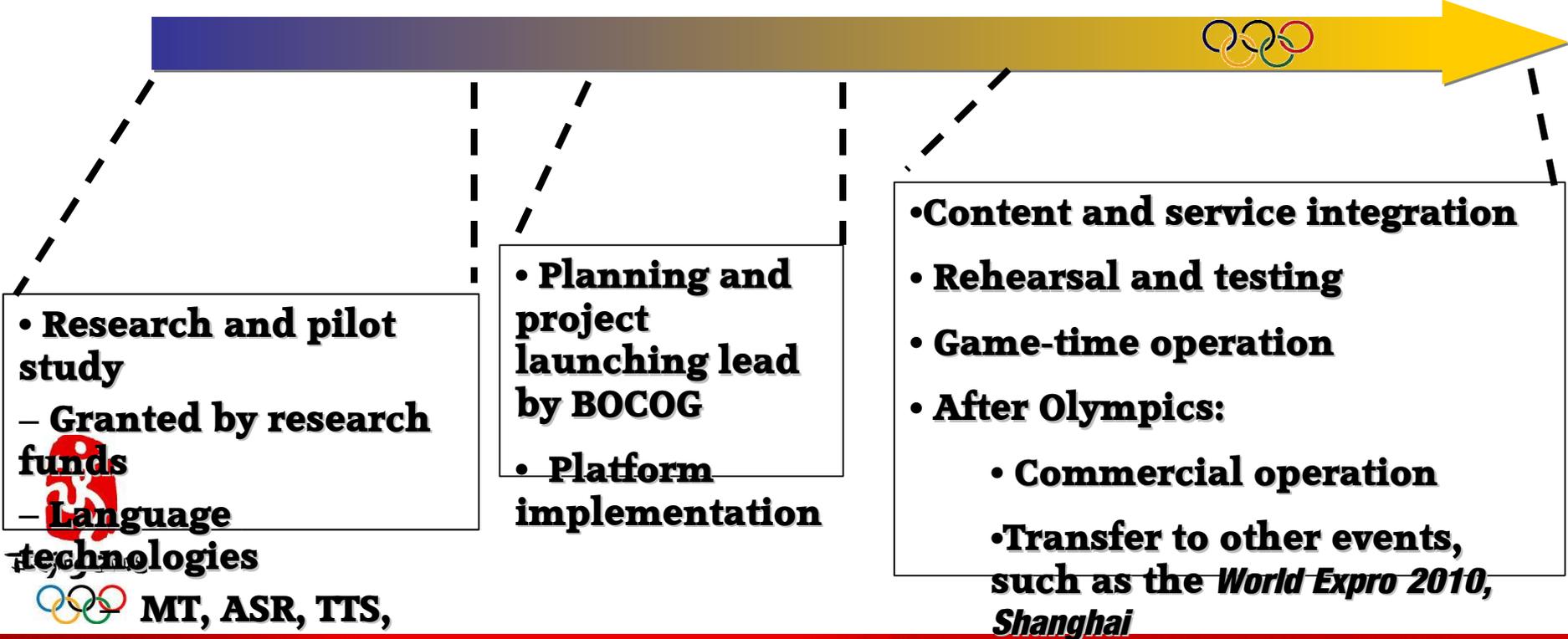
Overview

The roadmap towards ——

Multilingual Information Service System

2002   2003   2004   2005   2006   2007   2008   2009   2010

- **Research and pilot study**
  – **Granted by research funds**
  – **Language technologies**
    MT, ASR, TTS, IR…

- **Planning and project launching lead by BOCOG**
- **Platform implementation**

- **Content and service integration**
- **Rehearsal and testing**
- **Game-time operation**
- **After Olympics:**
  - **Commercial operation**
  - **Transfer to other events, such as the *World Expro 2010, Shanghai***

Pilot systems

Pilot systems

Research projects and grants:

From May, 2001 to Sept. 2005, supported by

- "863" — <u>National High Tech Research Program</u>
  - 3 consecutive grants
- <u>Ministry of Science and Technology (MOST)</u>
  - 2 grants
- <u>Beijing Municipal Science and Technology Commission</u>
  - 1 grants

Pilot systems

Technology contributing partners:

Machine translation:

- Huajian MT
- ChinaSoft Inc
- Nanjing University
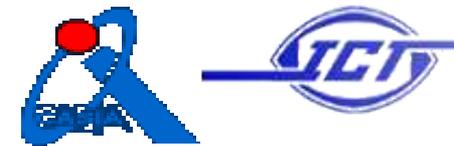- Xiamen University
- Harbin Inst. of Technology

Pilot systems

Technology contributing partners: *(continued)*

Speech recognition:

Automation Inst. , Chinese Academy of Sciences

Acoustic Inst. Chinese Academy of Sciences

TTS:

iFlytek, SinoVoice, Tsinghua Univ.

Linguistic resource:

Computing Tech. Inst., Chinese Academy of Sciences

Harbin Inst. of Technology, Peking Univ.

Information retrieval and Crosslingual IR:

TRS

System integration: CAPINFO

Pilot systems

- **Machine translation,** some of the language pairs:
  - Chinese←→English
  - Chinses←→Japanese
  - Spanish→Chinese
  - German→Chinese
  - Chinese←→French
  - Russian→Chinese
- **Translation Memory:**
  - Chinese→English
  - Chinese→French
- **Available on platforms:**
  - PC, as client/server or standalone
  - Pocket PC
  - Web services
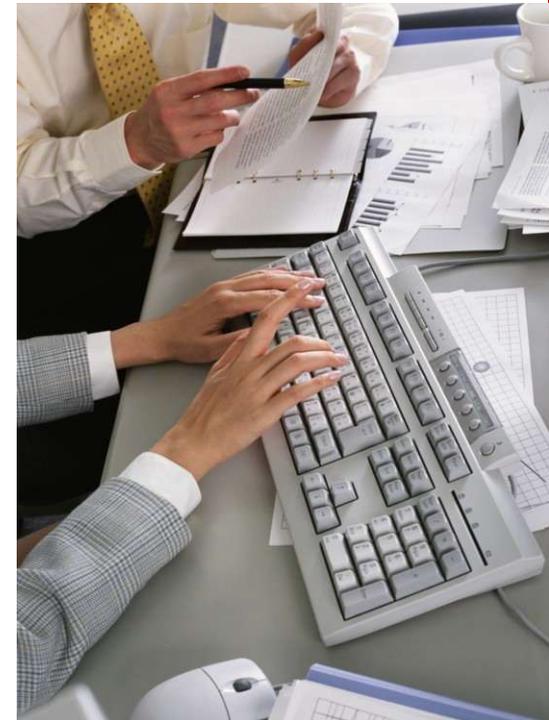
Pilot systems

Other technologies and resources:

➢ **Speech recognition**

- Multilingual: Chinese, English, French and Japanese
- Adapted for non-native speakers (e.g. English)
- Platforms: PC, Pocket PC, Smartphone

➢ **Text-to-Speech**

- Multilingual: 12 languages and dialects
- Mixed language (e.g. Chinese+English)
- Multiple platforms: PC, Pocket PC, Smartphone, Symbian and embedded Linux

Pilot systems

Other technologies and resources: *(continued)*

➢**Information Retrieval**
- Canonical IR
- Question answering, deep information extraction
- Real-time indexing and "0%" space expansion

➢**Cross-language IR**

➢**Linguistic Resources**
- Text and parallel data for ASR and MT
- Speech data for ASR

Beijing 2008

Pilot systems

Demonstration systems:

➢**Multilingual website for Olympics**
- Three official languages: Chinese, English, French
- Technologies:
  – Machine translation (MT)
  – Translation memory (TM)
  – Multilingual content management
  – Cross language information retrieval

Pilot systems

Demonstration systems:

➢ **Multilingual website for Olympics**

"Digital Olympics" website

http://www.do2008.gov.cn

"High Tech Olympics" website

http://www.hitech2008.org.cn

Pilot systems

Demonstration systems:

➢ **Multilingual Kiosks (Information booth)**
- Chinese, English and Japanese
- Information services on
  - Real-time weather
  - Beijing Tourism Information
- Technologies:
  - Machine translation (MT)
  - Translation memory (TM)
  - Text-to-speech (TTS)
  - Information retrieval

Pilot systems

Demonstration systems:

➢**Multilingual Kiosks (Information booth)**

Pilot systems

Demonstration systems:

➢**Multilingual Kiosks (Information booth)**

Pilot systems

Demonstration systems:

> **Multilingual Kiosks (Information booth)**

Pilot systems

Demonstration systems:

> **Multilingual Kiosks (Information booth)**

Pilot systems

Demonstration systems:

➤ **Prototype of "Tourist's PDA"**

- English, voice-enabled
- Location-based service
- Machine translation, handwriting input
- Phrase book translation
- Technologies:
  – Machine translation (MT)
  – Voice recognition
  – TTS

Pilot systems

Demonstration systems:

> ### Prototype of "Tourist's PDA"

Pilot systems

Demonstration systems:

➤**Real product––– "Tourist's PDA"**

- English, TTS-enabled
- Information on Hotels, Restaurants, Shopping centers,…
- Indexed phrase book translation
- Available at 10 selected hotels in Beijing (for leasing), since Sept'05

Pilot systems

Demonstration systems:

➢**Real product–– "Tourist's PDA"**

Pilot systems

Demonstration systems:

➢**Real product--- "Tourist's PDA"**

Pilot systems

Demonstration systems:

➢**Real product–– "Tourist's PDA"**

Pilot systems

Demonstration systems:

➢**Real product--- "Tourist's PDA"**
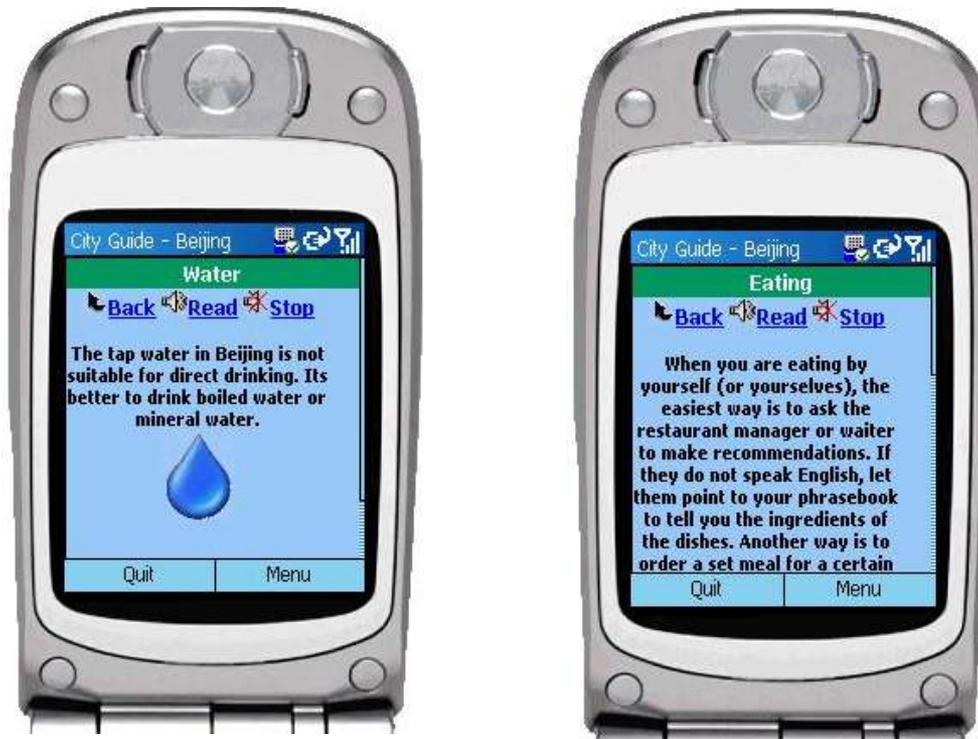
Pilot systems

Demonstration systems:

➢**Real product–– "Tourist's PDA"**

Pilot systems

Demonstration systems:

➢**Prototype of "Olympic PDA"**
- Targeting users —— "Olympic Family"
- English, Chinese and French
- Both speech recognition and TTS enabled
- Location-based service
- Machine translation
  - Chinese to English and French
  - English to Chinese
  - French to Chinese

Pilot systems

Demonstration systems:

➢**Prototype of "Olympic PDA"**

- 10 major functionalities:
  - Language assistant
  - Game Schedule
  - Medal tally
  - Venue's info
  - News
  - Bulletin board
  - Weather
  - Transportation
  - Map and location
  - "One-button" to Call center

Pilot systems

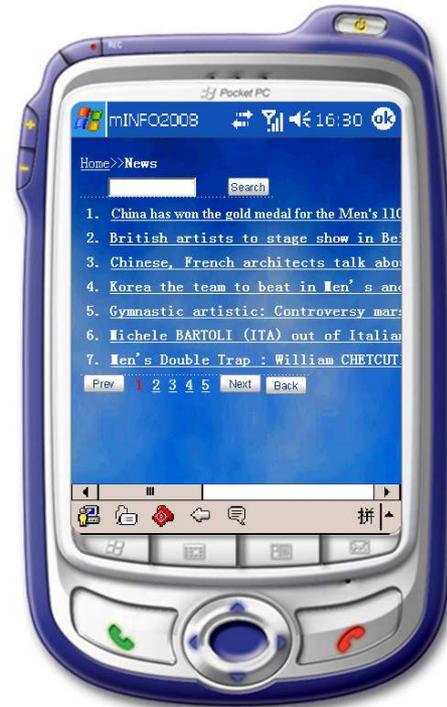**Demonstration systems:**

➤ **Prototype of "Olympic PDA"**

Pilot systems

Demonstration systems:

➤ **Prototype of "Olympic PDA"**

Pilot systems

Demonstration systems:

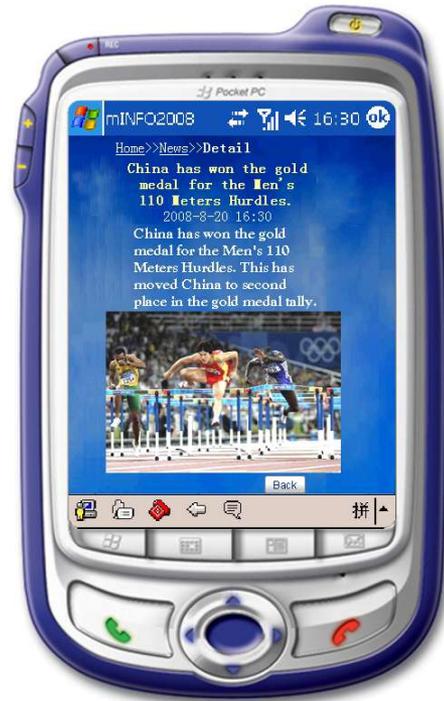➢ **Prototype of "Olympic PDA"**

Pilot systems

Demonstration systems:

➢ **Prototype of "Olympic PDA"**

Multilingual Information Services
# for Olympics

Mission and Objectives

"**Provide state of the art, yet mature, robust and tested solutions to the Olympic Games**"

➢ **Before and during Beijing Olympics**
- — Provide information services for "anyone" at "anywhere" and "anytime"
- — Integrate services from both the game and the city
- — Multilingual when needed
- — Support all possible means of access

➢ **After Olympics, secure the values to the city and society**
- — Benefit citizens, tourists and government
- — Mature technology solutions transferred to other cities and events

Beijing 2008
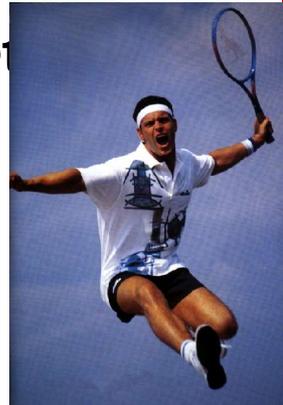
Project Goals

➢ **Build comprehensive information service system for the general public and Olympic family**

- Establish **<u>Olympic Call Center</u>**, provide language service and CTI services by staff and IVR

- Partner with the official Olympic website, build dedicated Internet websites, support existing city kiosks

- Support wireless devices (PDA and smart mobile phone) with LBS, WAP, SMS and MMS services

Project Goals *(continued)*

- **Provide services in the means of:**
  - Bundle existing city information with Olympic information
    - e.g. game-watching guidelines: ticket selling, traffic and bus schedule, game results notification
  - Integrate existing services in the city with new services
    - e.g. Call center's language service helps tourism office in serving foreign tourists

Illustration of the goals

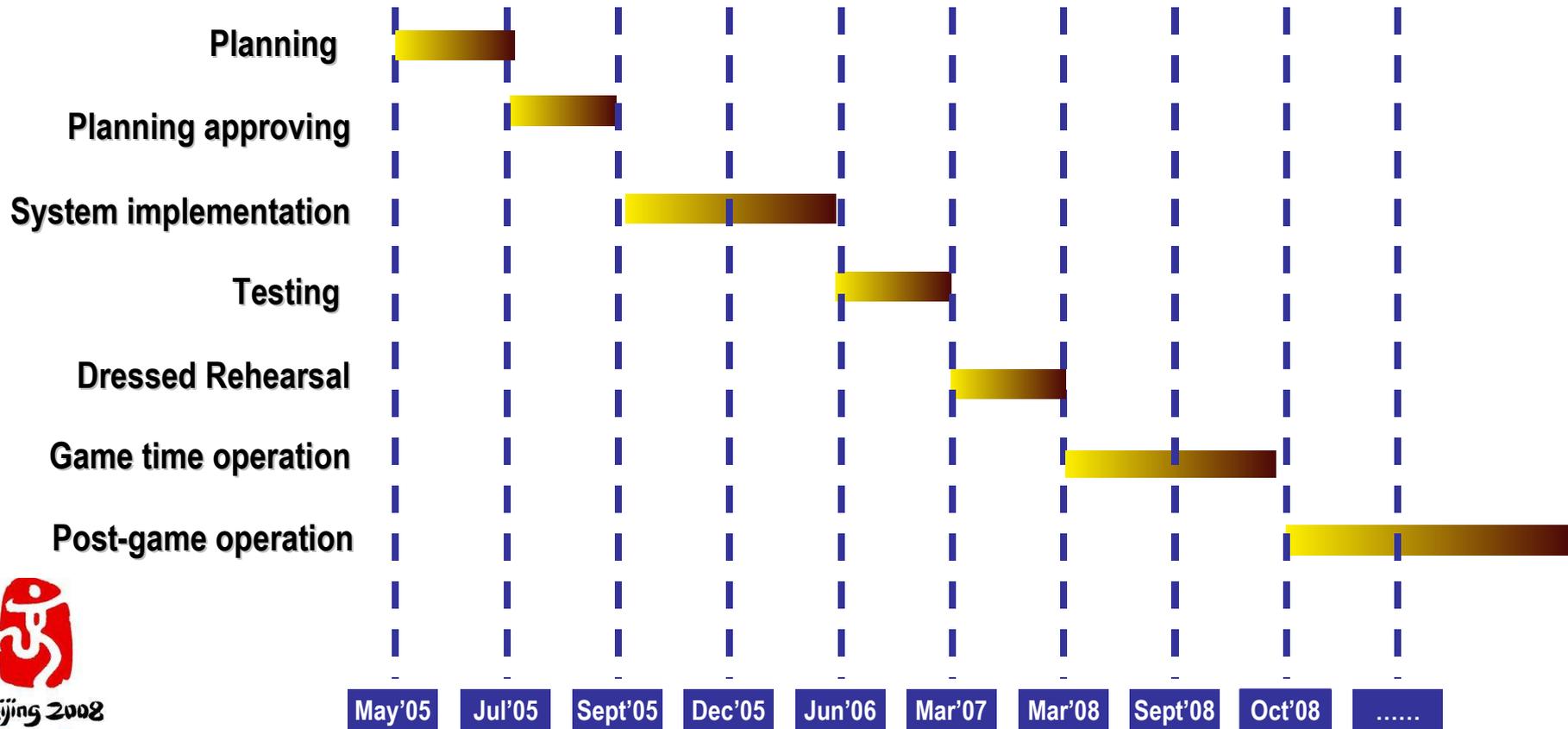Multilingual Information Service System for the Beijing Olympics
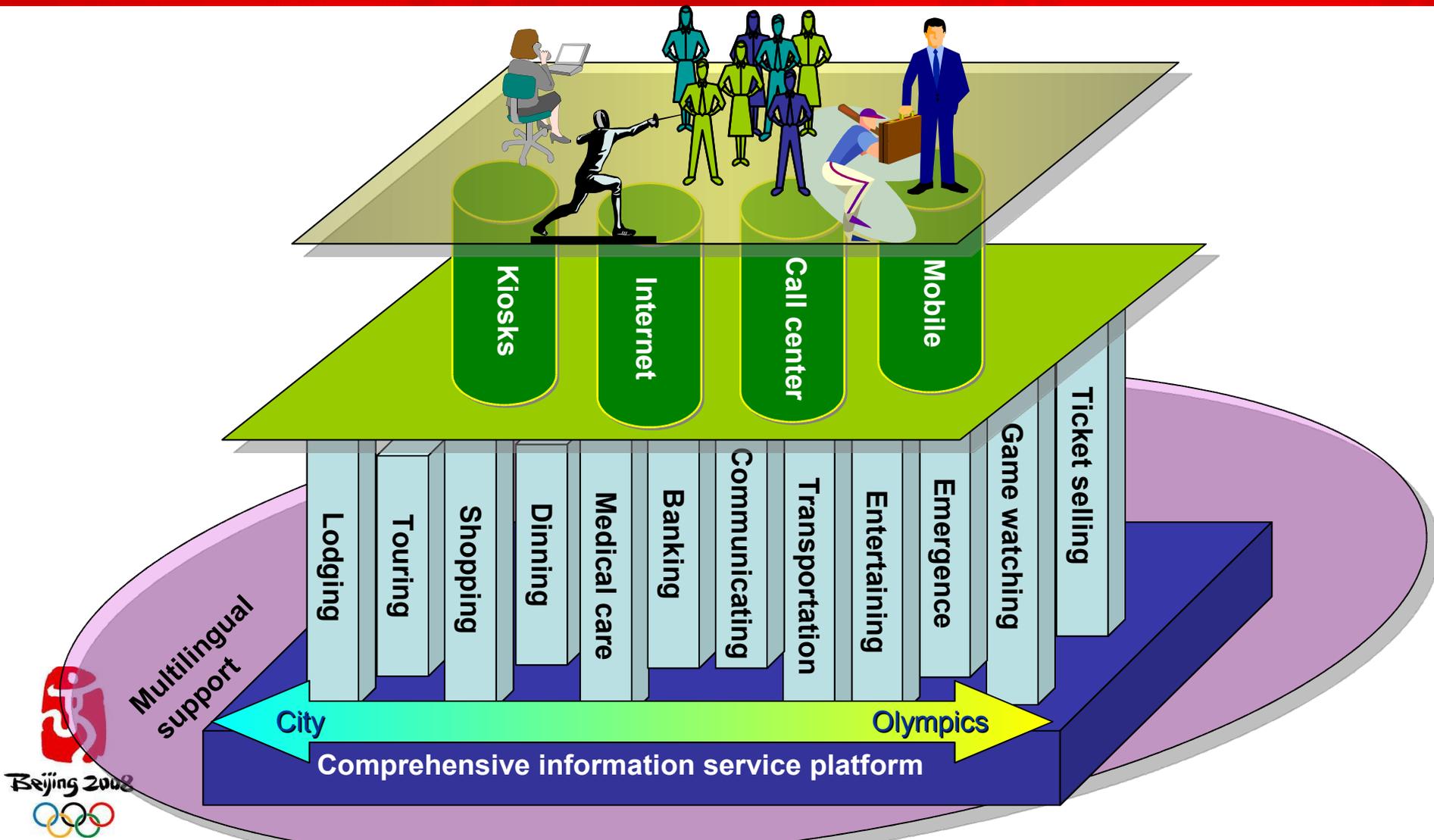
The project consortium

- Department of Technology, BOCOG
- Capinfo Company Limited
- China Netcom
- China Mobile
- Chinese Academy of Sciences:
  - Institute of Software
  - Institute of Computing Technology
  - Institute of Automation

Roadmap

> **Launched in April, 2005 by BOCOG, as one of "High-tech Olympic" projects**

Task	Timeline
Planning	May'05 – Jul'05
Planning approving	Jul'05 – Sept'05
System implementation	Sept'05 – Dec'05
Testing	Jun'06 – Mar'07
Dressed Rehearsal	Mar'07 – Mar'08
Game time operation	Mar'08 – Oct'08
Post-game operation	Oct'08 – ……

Timeline markers: May'05 | Jul'05 | Sept'05 | Dec'05 | Jun'06 | Mar'07 | Mar'08 | Sept'08 | Oct'08 | ……

Targeted users

- **Spectators** (estimated 7 million people)
  - Beijing citizen 6 million
  - Domestic 1.2 million
  - Foreigner (including HK, Macao, Taiwan) 0.6 million

- **Registered members** (300 thousands)
  - Olympic Family (55k)
    - IOC, IPC, NOC, NPC, IF
    - Athletes, Coaches, accompanying staff
    - Sponsors
  - Registered media person (25k)
  - Marketing partners (110k)
  - Staff (100k)
   - Paid staff (5,000)
    - Volunteers (80-100k)

From inside BOCOG, different functional centers:

➢ Game Services——for all the spectators
- Spectators service: hotel reservation, transportation, venue gate entering, watching, venue quitting

➢ International Liaison——for VIPs and ceremonies
- Multilingual Switchboard, 40-50 languages

➢ Human Resource——for volunteers
- Volunteers recruiting, training and management

➢ Sports service——for game participants
- Schedule, scores, athletes info

➢ Media operation——for journalists
- Registered media service, real-time broadcasting

Requirements and Resources

- ➢ Cultural events
  - Event schedule and service
- ➢ News service
  - Real-time news reports
- ➢ Marketing
  - Ticket selling, souvenir
- ➢ Medical service
  - Medical care centers
- ➢ Security service
  - Security control information

Beijing 2008
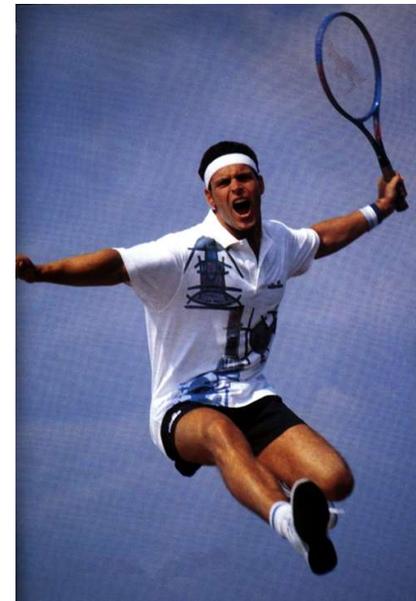
Requirements and Resources

- ➢ **Game related information includes:**
  - – Ticket sales information
  - – Schedule of the games
  - – Introduction to the athletes and teams from every nations and for each competition
  - – Description of the venues, their locations, the means of transport to get there, what needs to know for entering and watching a game
  - – Real time release of the competition results and the medal tally
  - – Knowledge of the sports, such as competition rules

From the city:

➢ Tourism Bureau

- Spectator's hotel reservation center
- Travel and touring info

➢ Commerce Bureau

- Restaurants, shops, entertainment info

➢ Transportation commission

- Public transportation helpdesk

Requirements and Resources

➢ **City information includes**:

– **Information for tourists**:

- lodging, site seeing, shopping, entertainment, restaurant and transportation

– **Basic information** and common sense

- weather, currency exchange, emergency numbers, medical care services

– **Location based information**

- GIS+ dynamic content from the city, such as the guiding introduction to a tour site, or the description and location of a restaurant etc.

Defining the system

Multilingual information services:

- Pre-game information dissemination
- Hotel reservation service
- City information service (including e-commerce)
- Public transportation helpdesk service
- Game information dissemination (schedule, venue, scores, medals…)
- Emergency helpdesk service
- Tourist consulting and reservation service
- Location based service on mobile devices
- Multilingual language assistant service

The role of MT

The role of MT in the system

- **Translation from Chinese to English, French, Spanish, Japanese,...**
  - Online, for real-time publishing
  - Offline, working with Translation Memory
- **Translation from other languages to Chinese**
  - Indexed phrase book
  - User's input sentences

Beijing 2008

The role of MT in the system

- **Domain**
  - Chinese specific and Olympic related domain,
    - Chinese entity names: people, place names
    - Athletes and sports terminology

- **Scenario**
  - MT: Structured information, e.g. bus route, weather, currency exchange rate,….
  - TM: for multilingual versions synchronization

- **Other related technologies**
  - Speech recognition
  - Text-to-speech
  - Information retrieval and crosslingual IR
  - Content management System (CMS)

Thank You !