



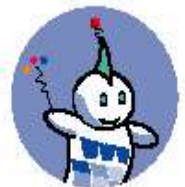
SNOW

Services for nomadic workers



**A framework for
authoring & exploiting
multimodal documentation**

Nicolas CHEVASSUS
EADS Corporate research Center
W3C Seminar - 21st June 2005





SNOW overview



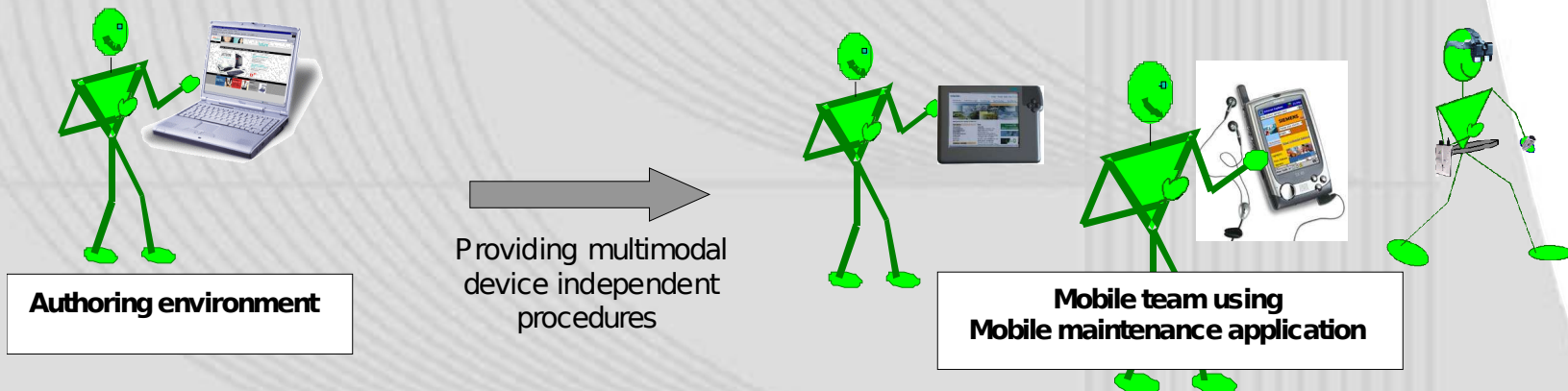
- Objectives
- Aeronautics challenges for maintenance operations
- Multimodal I/O
- Authoring & Execution framework (XTM-P & D3ML)
- Usability
- Conclusions

SNOW objective & challenges

A large-scale industrial **diffusion of multimodal mobile documentation**

for maintenance operations

- How to **author** multimodal mobile maintenance documentation?
- How to **exploit** this documentation through robust interaction modalities?





Aeronautic maintenance business case

Airframe
Engines
Systems
Landing gears
Whole aircraft

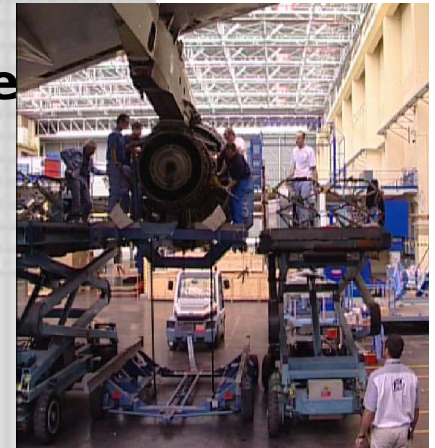
- **Complex products**

- **Aircraft & helicopter**
- **Handfree operations / interactions with test de**
- **Multiple manuals (>30) & complex procedures**

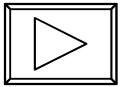
- **Challenges** - A/C manufacturers, MROs, Airlines -

- **Faster, cheaper, safer and more efficient maintenance actions**
- **Reduction of Operation Interruptions**

➔ **Taking benefit of mobile computing**



Procedure Analysis of an inspection



Information access
(Job set up information)



Inspection
(Procedure)



Anomaly detection
(Trouble shooting)



Expertise
(Trouble Shooting)



Non Routine Card
(Trouble shooting)



Repair
(Procedure)



Validation
(Close up)

W3C Seminar – TOULOUSE, 21/06/05

On-the-field survey

- Information access

- No access to technical data sources from working area & large distances
- 20 to 50% of a maintenance task is lost in searching for information

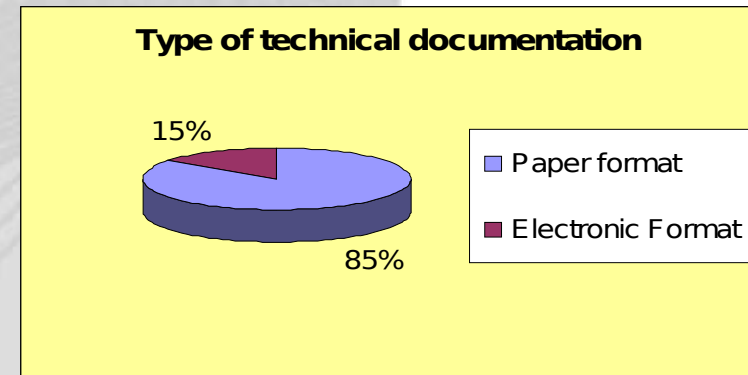
- Communication

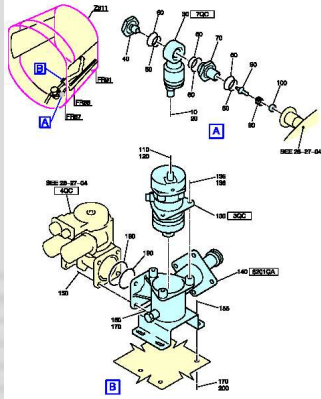
- Poor communication system on working area
- Large distance to cover to communicate with team members
- No direct access to remote expertise on line - which means lost of time
-

- Work conditions

- In 85% of maintenance tasks, operators need to be hand-free
- 100% of operators regularly perform their task in uncomfortable positions or environment

➔ Improve work conditions, increase productivity and reactivity





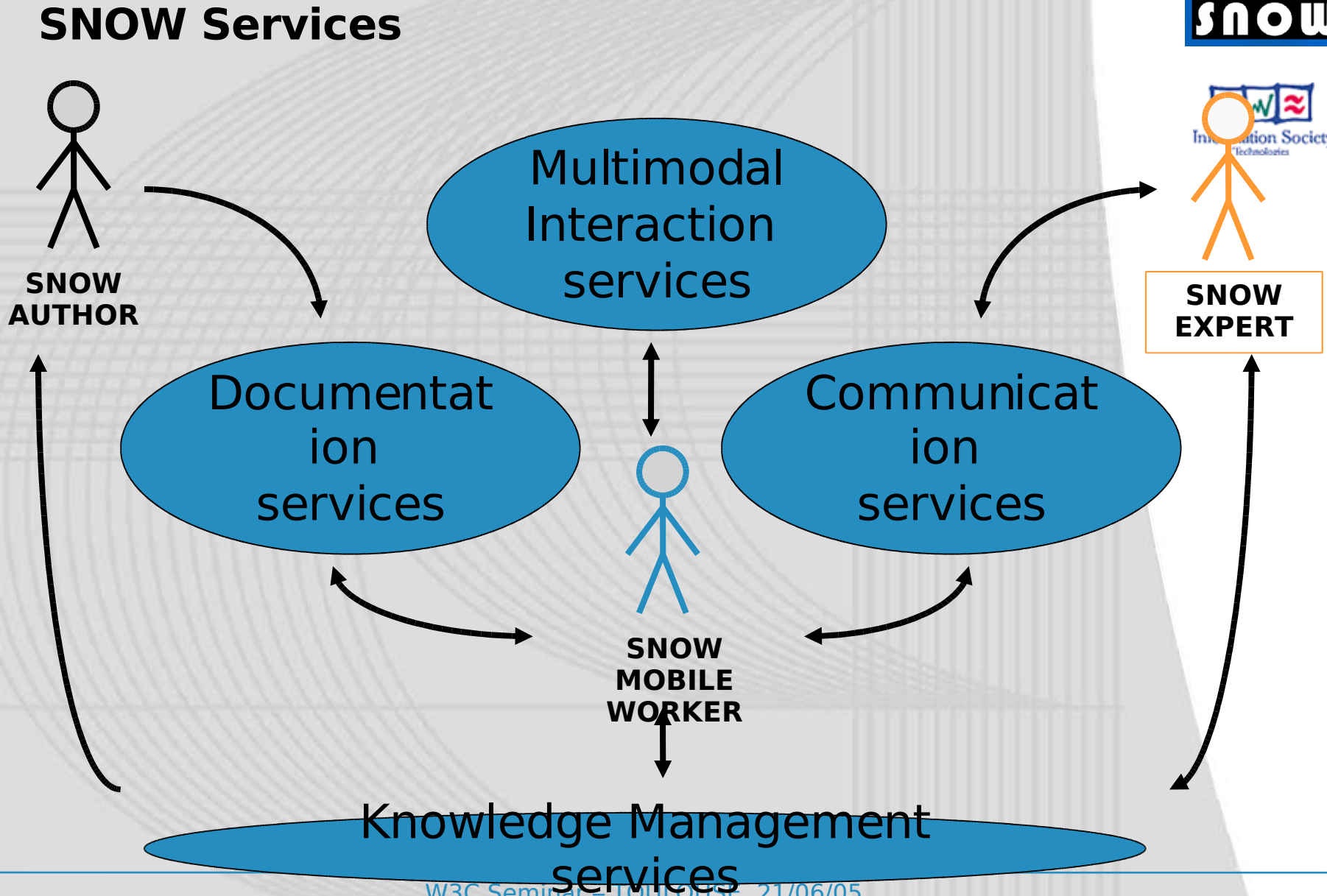
Procedures

- Structured document
- Texts & graphics

E-multimodal procedures

- **Walk around visual inspection**
 - wearable computer + voice command + Text-To-Speech
 - **Pre/post flights checks**
 - pen tablet computer + Text-To-Speech
 - **Maintenance tasks**
 - see-through display + wearable computer + voice command
- You author once and use anywhere, anytime with any device







SNOW Approach Towards Content Authoring

- Separation of concerns – Using two languages to describe **content** and **user interface (UI)**:
 - Topic-map-based description of procedure content **XTM-P**
 - Extensible mark-up language for device-independent and multimodal UI description **D3ML**

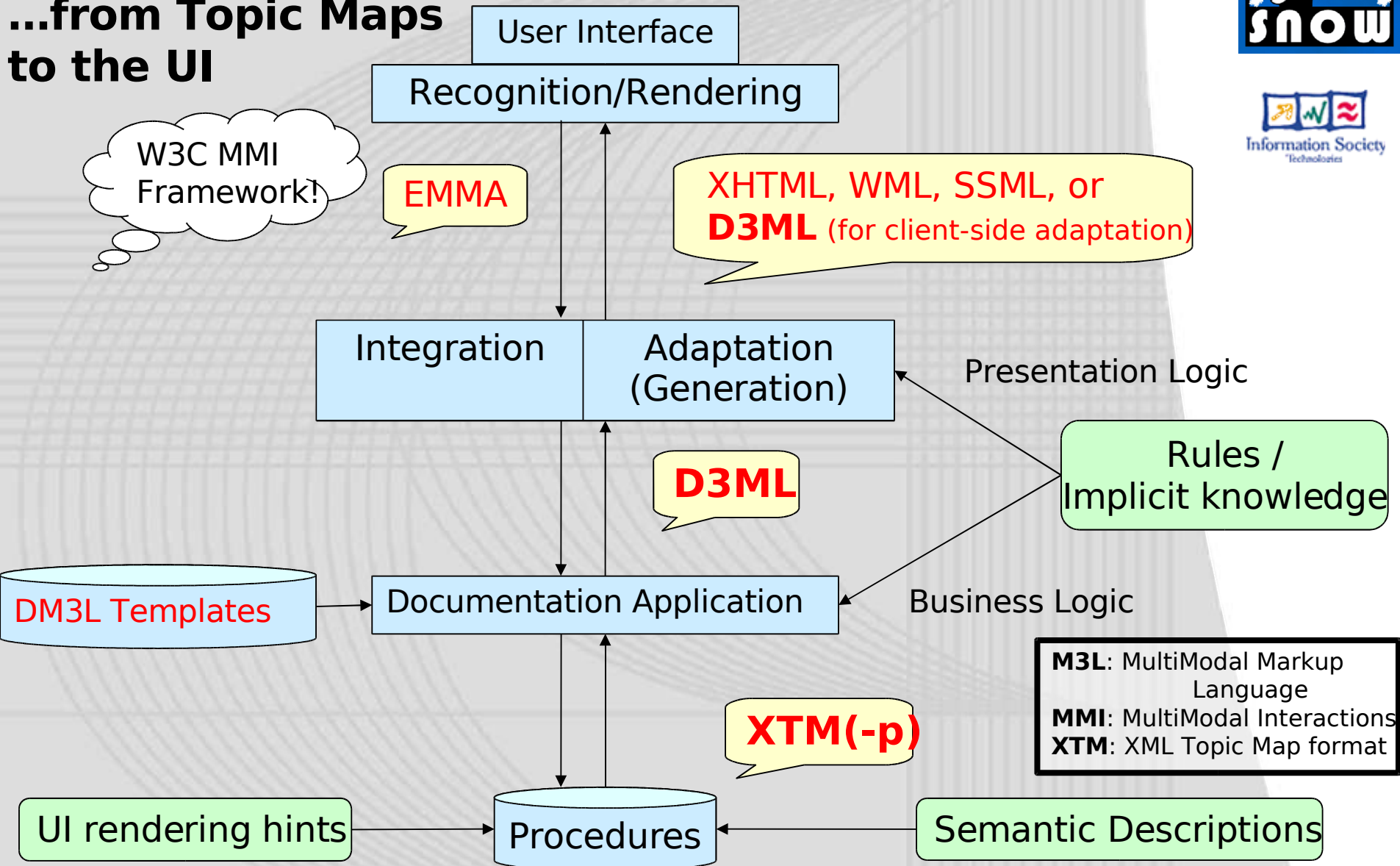
Multimodal commands

- EMMA (Extensible MultiModal Mark-up Language) for describing the interpretation of user input
- Follows standardization of W3C Multimodal Interactions Framework





...from Topic Maps to the UI





XTM-P : XML Topic Map format for maintenance Procedures

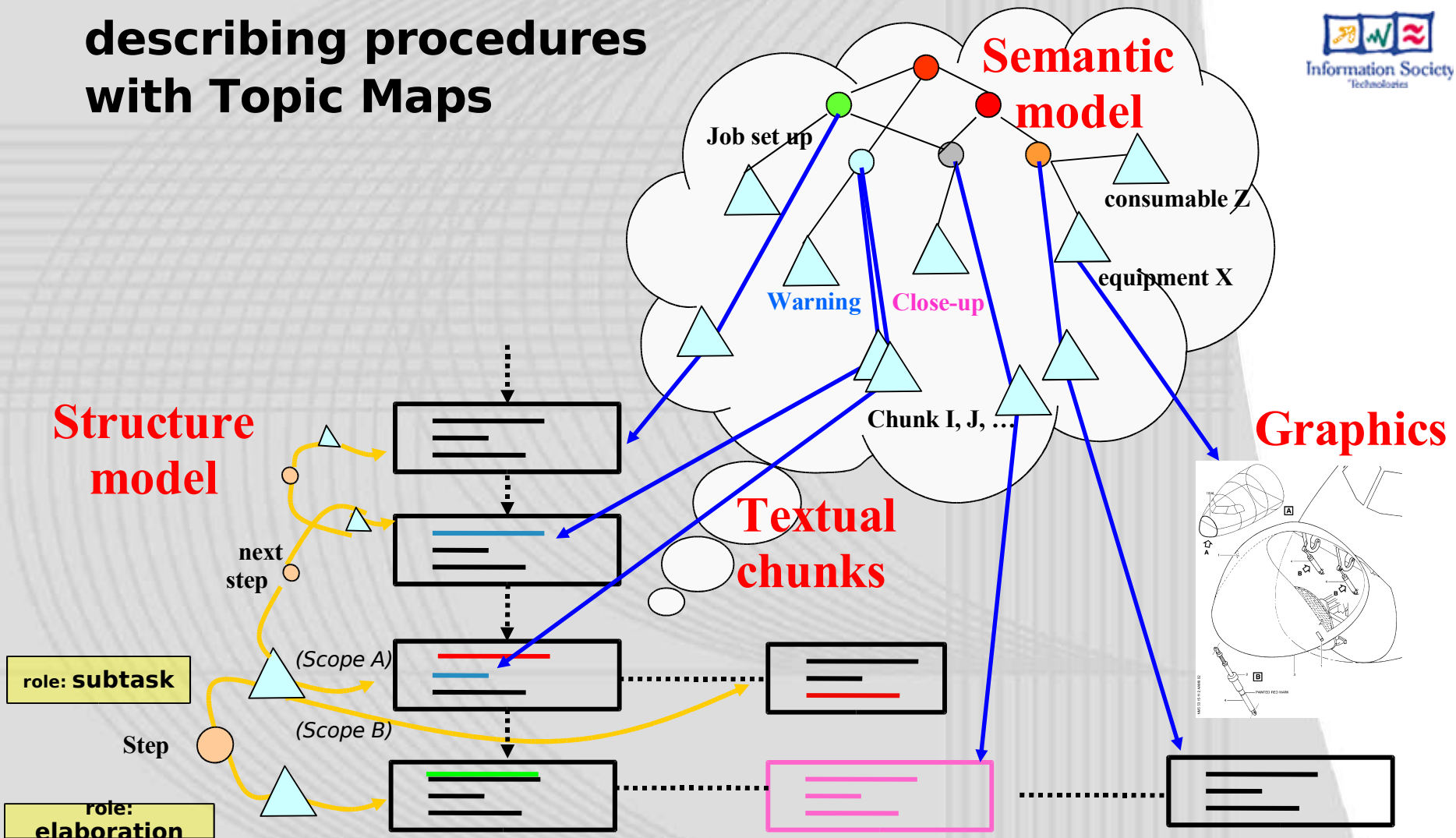
- Technically: XTM 1.0 use specification ; RDF export
- WHY XTM-P ?

Rich normalized syntax to describe taxonomies, rhetorical structures, contextual aspects, lexicons...

Benefit from available components both to :

- build resources: reuse published set of types and other on-line logical primitives as XTM topic map usable in creating new instances of procedures
- optimize the information structuring and navigating: decomposition of relations between text, figures, reference of tools, equipment, consumable
- align the semantic representations with an object-oriented model: instantiation is relying on XTM-P features such as:
 - constraints declaration on types and cardinalities
 - inference rules on conceptual relations

XTM-P describing procedures with Topic Maps





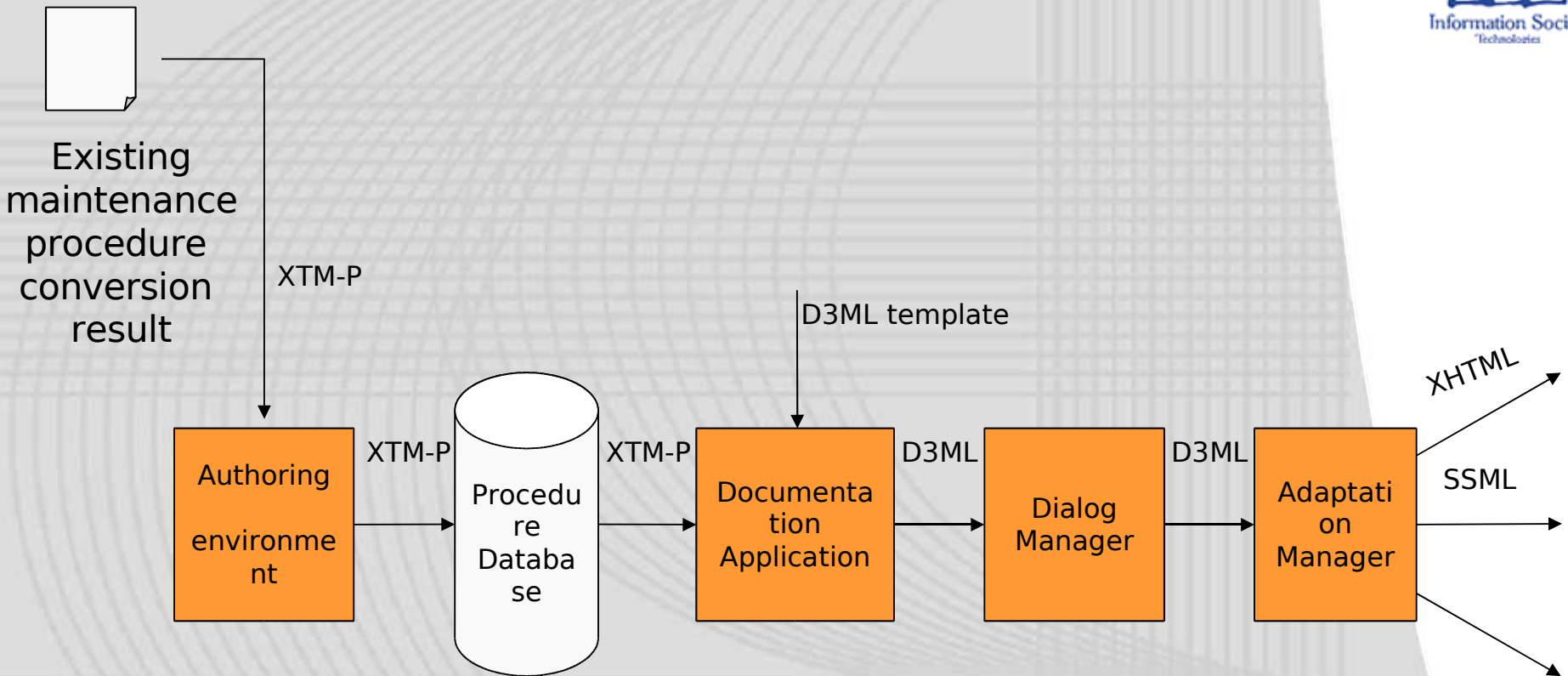
XTM-P procedures

Following **W3C** Workshop on *Metadata for Content Adaptation*:

- Semi-automatic generation of XTM-P procedures, derived from existing markups in Maintenance Manuals
- Metadata definition guide how the adaptation engine can react to the information
- Independent evolvement of the metadata definition and enrichment with reasoning rules for XTM-P / DM3L alignment purpose



Processing of XTM-P and D3ML





D3ML: Scope and features

- Device-independent Multi-Modal Mark-up Language
- **Generic description of arbitrary device and modality independent Web content**
- Inspired by Renderer Independent Mark-up Language (RIML) of the EU project CONSENSUS
- Scope:
 - *User interface templates* with hooks to business logic
 - *Transfer format for intermediary representations* of user interfaces, containing all information necessary for adaptation to specific end devices and modalities
- Built as an XHTML2 language profile using various modules from XHTML2, XForms, and DSelect, among others

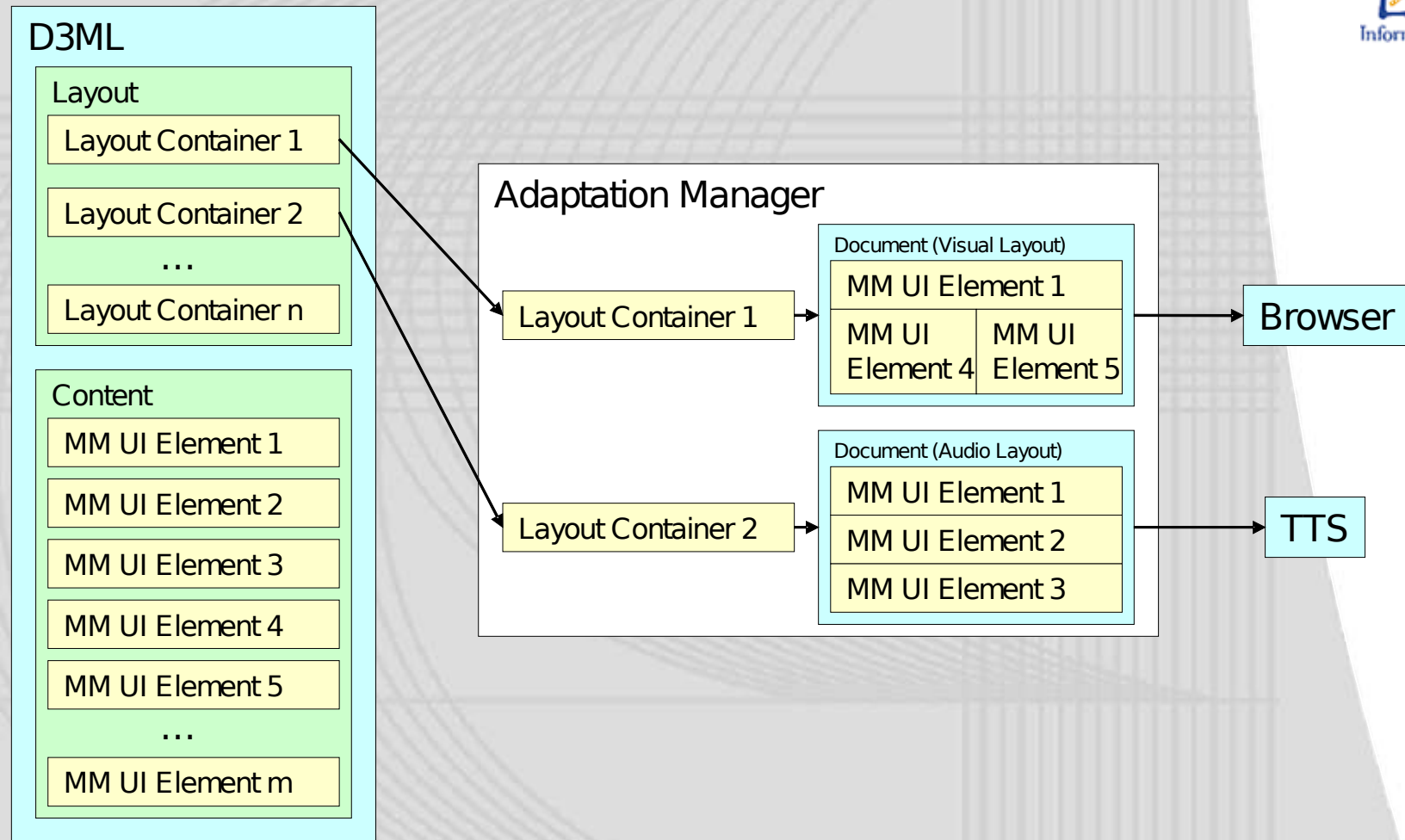


D3ML concepts

- Clustering & combination of possible end-device-modalities according to device capabilities
- Multiple layout containers (alternative presentations)
- Content control (conditional presentation depending on preferences and properties)
- Pagination (splitting of content to separate pages/units)
- Binding of multimodal commands to application-specific action handlers
- Identification of UI elements (for relationships etc.)
- Specification of meta-information for adaptation (e.g., non-scalable images)
- Input data validation

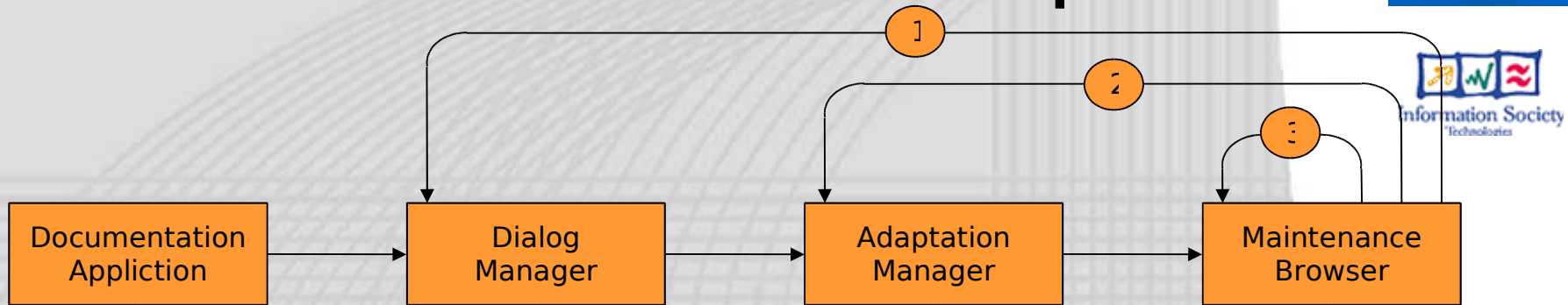


Multiple layout containers





Multimodal commands and their scope



1 – Application level command 2 – Page navigation command 3 – Browsing control command

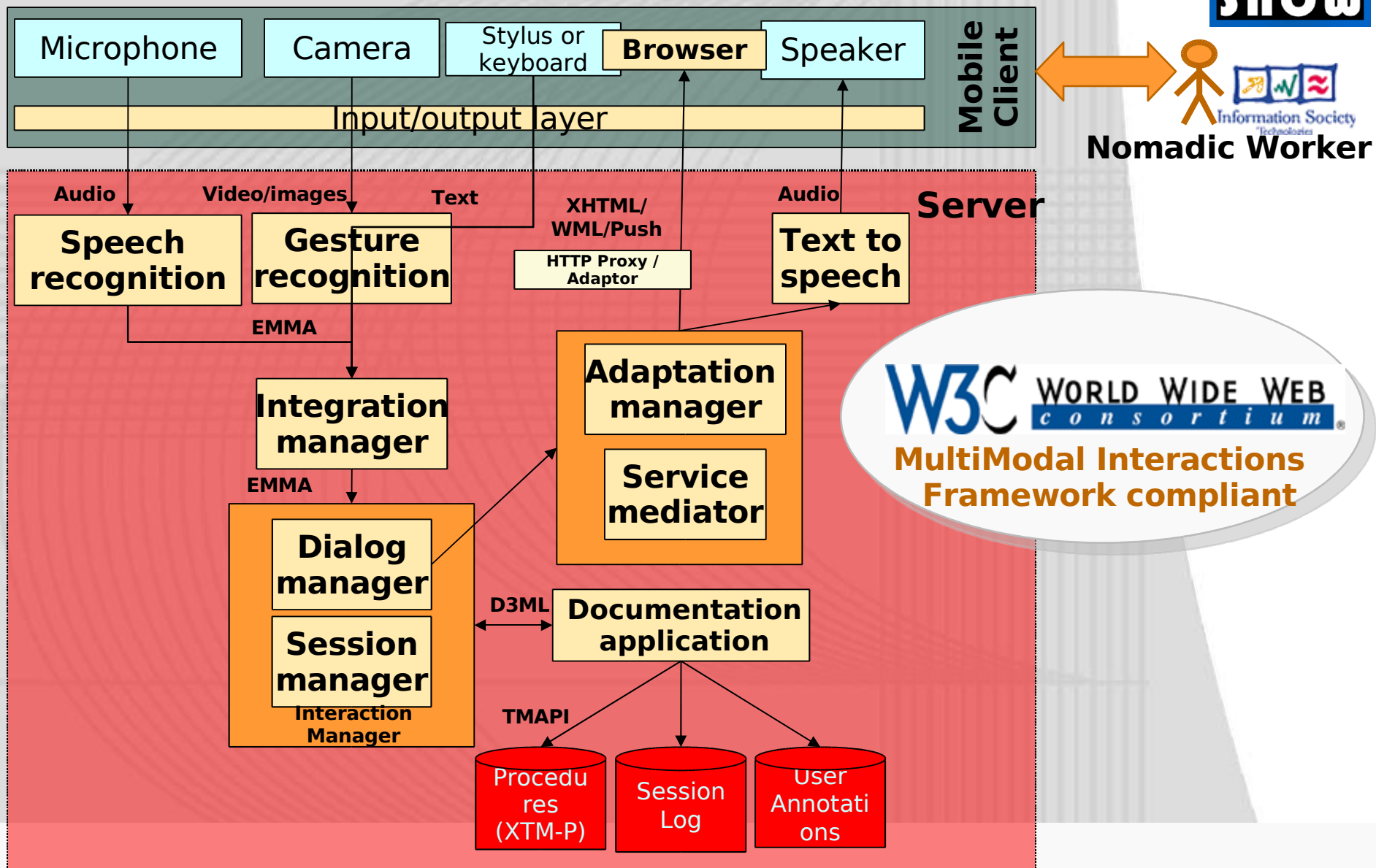
- Application level commands
 - Directly related to business logic (e.g., fetch data)
- Page navigation commands
 - Necessary due to content adaptation to end-device size and output modality restrictions (i.e., pagination and voice dialog steps)
- Browsing control commands
 - E.g., scrolling, input focus navigation



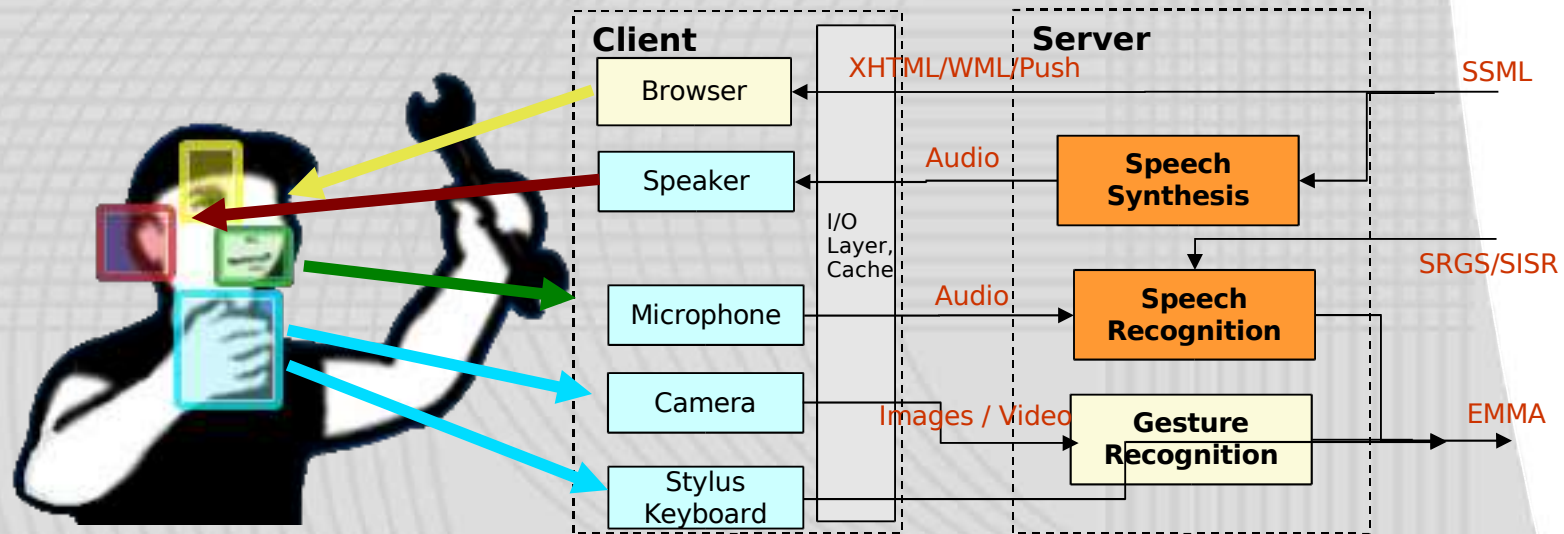
Why EMMA for commands?

- Extensible MultiModal Mark-up Language
- XML markup language for describing the interpretation of user input
- Seems to replace NLSML in IETF's MRCPv2 because of more general focus on arbitrary modalities
 - Natural Language Semantic Markup Language (NLSML), an XML-based markup for representing the meaning of a natural language utterance
 - Media Resource Control Protocol version 2 (MRCPv2), a network protocol between speech media servers and application platforms

Architecture of Mobile Working Environment



Mobile Worker Interaction Concept



SRGS Speech Recognition Grammar Specification

SISR Semantic Interpretation for Speech Recognition

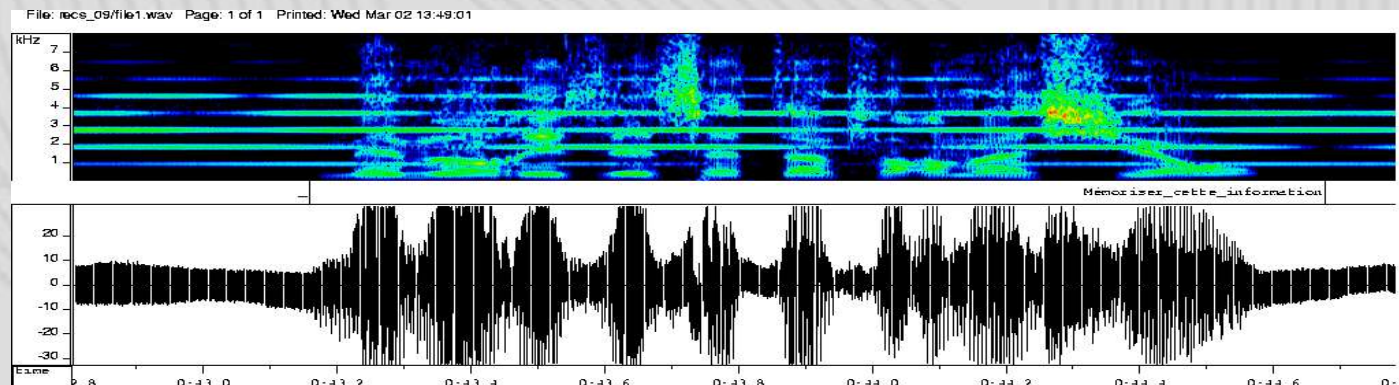
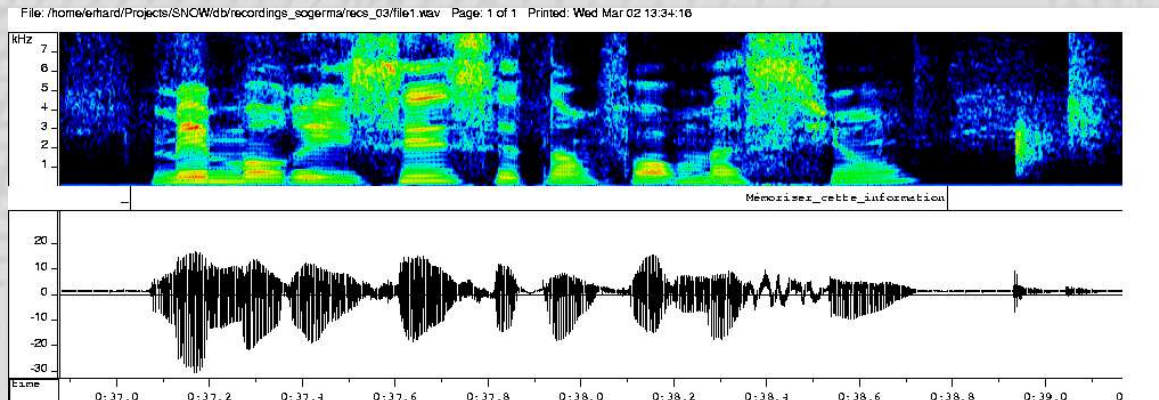
EMMA Extensible Multimodal Annotation Markup Language

SSML Speech Synthesis Markup Language



Acoustic environment of the mobile worker

- Non-stationary/impulsive and harmonic distortion
- High dynamic range of speech and distortion





Introduction: interaction by use of hand gestures captured by body-mounted camera system

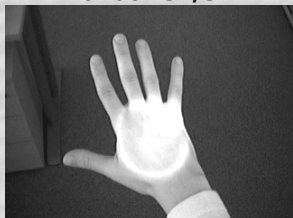
- **Modality:** Gesture recognition for interaction when other input modalities fail (e.g. because of environmental noise, wearing gloves)
- **Gestures vocabulary:** should be easy to *learn* by the worker and easy to *recognize* by the software
- **Challenges:**
 - unknown illumination
 - unknown, textured, and moving background
 - unknown skin color and/or gloves
 - robust real-time recognition
 - low processing power and bandwidth



Selected gestures

- Static gestures only (lower bandwidth and CPU requirements)
- Set of command/gesture associations

validation/ok



close



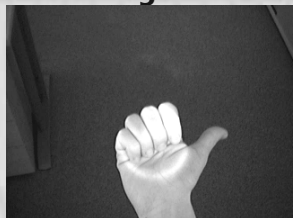
up/more



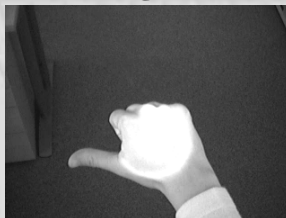
down



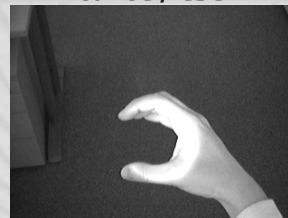
right



left



cancel/reset



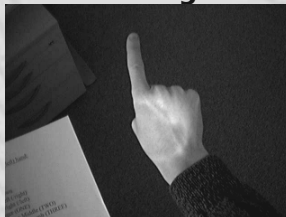
zoom



repeat



counting 1



counting 2

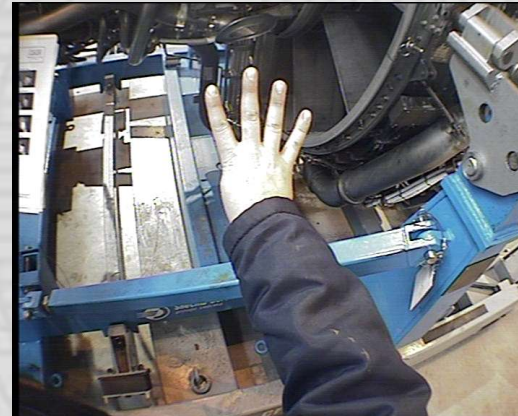


counting 3



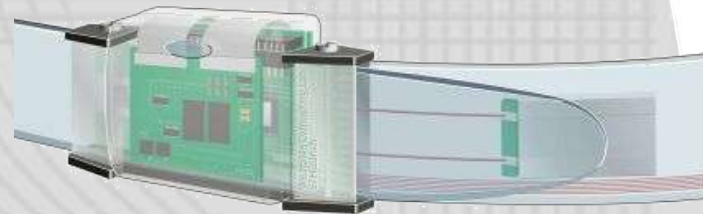
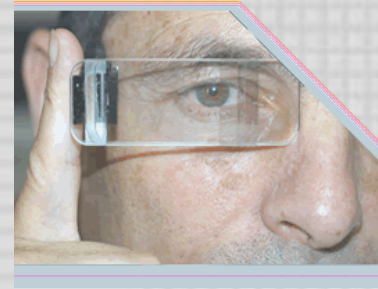
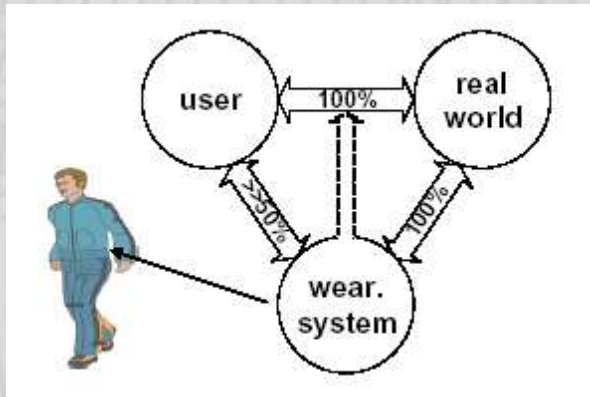
Test recordings & segmentation

- Real world recordings done at EADS Sogerma in Bordeaux



Usability issues: SNOW & wearIT@work

- 3D Multimodal Interactions : voice and gesture
- Content adaptation
- Device independence



- Unobtrusive, non-disruptive interactions
 - Work with the computer instead at the computer
- You author once and use anywhere, anytime with any device



SNOW Results (mid 2006)

- **Authoring environment**
 - Device-independent and modality-independent authoring environment
 - Semantic representation & multimodal XML Language
- **Environment for execution**
 - Common standard-based user interface integration format
 - Multiple access devices supporting different interaction modalities
 - Multimodal browser extensions
- **Robust Modalities**
 - Speech output according to the acoustic environment
 - Noise robustness of the Automatic Speech Recognition
 - Dynamic hand-gesture recognition with 1 camera in rough context
- **Feedback on mobile usability**



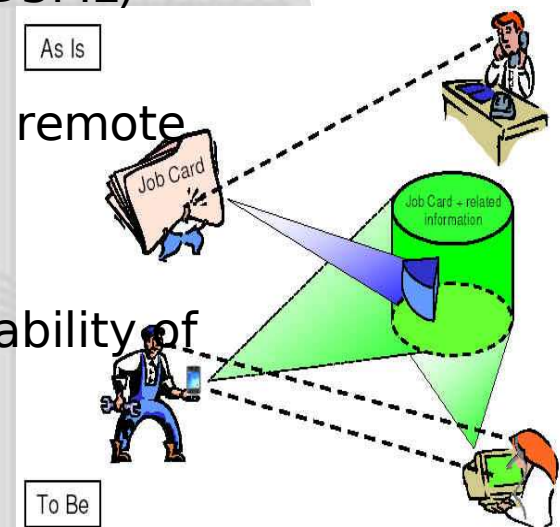
Key innovative contributions of SNOW

• Technical breakthroughs

- Voice technology for industrial environment
- Signal enhancement using non-linear adaptive algorithms
- Gesture recognition for industrial environment using information fusion
- Unification of Device Independence and Multimodality (D3ML)
- Combined authoring and data management based on standards and separation of concerns (XTM-P & D3ML)

• Innovations for maintenance services

- Easy access to large amounts of information and remote expertise
- Seamless information flow (no media breaks)
- Improved safety and reproducibility due to traceability of operations
- Enhanced usability and work conditions
- Improved productivity, flexibility and agility



Thank you for your attention

Contacts :

Pascale.hugues@airbus.com

David.roussel@eads.net

Nicolas.chevassus@eads.net

