



WebGALAXY – Integrating Spoken Language and Hypertext Navigation

**Raymond Lau, Giovanni Flammia,
Christine Pao, and Victor Zue**

**<http://www.sls.lcs.mit.edu/raylau/webgalaxy>
webgalaxy@sls.lcs.mit.edu**

**Spoken Language Systems,
MIT Laboratory for Computer Science
Cambridge, Massachusetts
United States of America**



Outline

- **Introduction**
- **Conversational Systems & GALAXY**
- **WebGALAXY**
- **Video tape**
- **Conclusion**



Information Nirvana?



Illustration by Sally Lee

- Vast amounts of information and services available online
- Huge market opportunity for providers
- Nirvana for users?



Or Information Overload...



Illustration by Sally Lee

- **User interfaces have remained primitive**
 - Point, click, type
 - Keyword search engines
- **Valuable cognitive capacities wasted in tracking links and web space geography**



Conversational Access

- **Virtues: natural, flexible, efficient, economical**
- **Beyond “speakeable links”**
 - **Extend the user interface paradigm with systems that understand**
- **Particularly useful when**
 - **Information space is broad and complex**
 - **Concepts are complicated**
 - **Users are technically naive**

- Flexible: hands and eyes busy



GALAXY (ICSLP '94)

- **Architecture for work in conversational systems**
- **Understood three application domains: city guide, air travel, and weather**
- **Framework for research in: speech recognition, NL understanding, NL generation, etc.**
- **Two tier client-server architecture: servers handle compute and knowledge intensive functions**
- **Extensible: can add new servers incrementally**



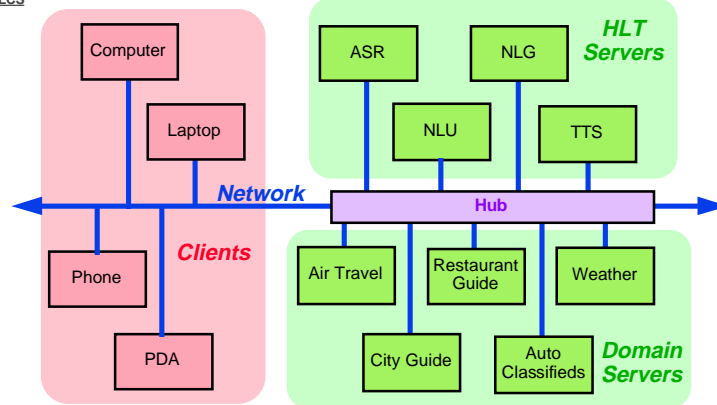
Universal Access with WebGALAXY

- **GALAXY**
 - Employed specialized protocols and technology, impeding mass availability
- **Web browsers**
 - Ubiquitous
 - Familiar interface
 - Widespread availability of web centric technologies
- **WebGALAXY: Make conversational technology available to a universal audience**
 - Move to feather-weight clients (fewer resources)
 - Adopt web standards to leverage browser penetration and technologies' capabilities

- Leveraging the technology: e.g., Unicode for multilingual displays



Three-Tier WebGALAXY Architecture



Client: Java applet in browser with telephone for speech

Hub: Mediates communications, maintains discourse

Servers: Compute and knowledge intensive functions

- **Client**
 - Java applet lives in browser
 - Uses telephone system for spoken interaction
- **Hub**
 - Mediates communications between multiple servers and clients
 - Maintains discourse state
 - Frees clients from having to deal with multiple servers, multiple protocols, etc.
- **Servers**
 - Handle compute and knowledge intensive functions



Client

- **Runs in any Java enabled browser**
 - Netscape Navigator, Microsoft Internet Explorer, etc.
- **Centered around a Java applet**
 - Communicates with hub
 - Allows hub to push new replies to client
- **Uses standard browser capabilities for graphical user interface**
- **Also have forms based version for more light-weight environments (e.g., PDAs)**



Video

- Video tape...



Conclusions

- **Trend towards enhancing point-and-click with richer, more flexible and intuitive models of user interaction**
- **Increasingly custom tailored presentations of information rather than static displays**
- **Conversation can be a major contributor**
- **WebGALAXY**
 - **Adoption of thin client and web standards increases audience of technology**
 - **Introduction of three-tier architecture makes thin client possible**



The Road Ahead

- **More universal access: Goal of any time, anyone, anywhere**
 - Internet telephony (one wire): Initial proof-of-concept experiments promising
 - Displayless systems (only a telephone required):
 - * e.g., Jupiter (which you can call here at EuroSpeech '97)
- **More information accessible via conversational systems**
 - Rapid application domain prototyping
 - Information extraction tools and new content organizations, e.g., metacontent