CHALLENGES	APPROACHES/SOLUTIONS
ONTOLOGY Challenges • Ontology Creation/Population • Tasks <-> Tools - reusable across domains • Understand a process model (and human's role in this) • Semantic Web • User-centered process view • Convert the (HCI) disbelievers … and keep them practicing • "top" or core ontology (use this to bootstrap new domains), Ontology integration • Rapid customization (to specific domains) • Use domain specific ontologies to organize massive documents • Find, learn, collaboration with domain ontology creators • Integration of shallow/deep methods **ONTOLOGY Problems** - Ontology quality - Access to info, knowledge visualizations - Understanding - Ambiguity	**ONTOLOGY Approaches** • Relation of HLT to ontological tasks • KR, linguisits, & ontologies to jointly address … • Component –based methods for • Life cycle • Re-use • Decomposition • Use HLT to support knowledge audits –> Identify IP -> innovation • Context capture • Controlled, language management **ONTOLOGY Solutions** • Plug-in (for IE) • Semantic Web • Tools to leverage small ontologies -> large ontologies
SUMMARIZATION Challenges: • level/depth of analysis/representation (E.g., Speech acts, RST, semantic rels) • Sumarization presentation/visualization • Speech (not good for long texts) • Indicative vs. inforamtive, concepts vs. ideas • Action-oriented summaries (e.g., executive/management summaries)	**SUMMARIZATION Solutions** - Analysis -> transformation -> presentation
MULTILINGUAL Problems • Relational between cultures, languages, lexical resources, ontologies • Domain knowledge • Fine-grained linguistic knowledge (e.g., stylistic details) • Size, complexity 200 languages -> 39k language pairs • Language invisibility large-scale, robust NLP • Adaptation/integration of semantic resources	**MULTILINGUAL Solutions** • resources: wordnet, euronet, application database, text resources • Interlingua approach • Statistical -> deeply annotated data + machine learning • Translation memories + ML • Multimodal/multimedia sols • Multiple ontologies tailored to users, tasks

• Content-driven hypertextual authoring • Cross-lingual news linking • Advanced software technologies/platform • Communication/transaction success	

MULTIMEDIA Challenges
 	
- Processing – centralized/mobile
 	
- Privacy, security, scaleability
 	
    • Remembering + forgeting
    • multilingual and multisource IE – incremental information building
    • cross-document co-ref resolution
	MULTIMEDIA Solutions
	
	    • Location-based services
	    • "forgetting"

Input to a Technology Road Map:

Enabling Technologies/Infrastructure

- Mobile communications
- Push service failures (e.g., pointcast)
- Satellite communication bankruptcy
- Fibre explosion

Services

- video on demand failure – need for content based access

Resources

- RDF, DAML, OIL?
- Ease of integration
- IE, NE

Fundamental/Hard Problems

- Noisy Speech Recognition
- Non-literal language
- Semantic web (e.g., who is going to populate it)

Ontologies

- Auto Web Taxonomy Generation
- High Quality MT

/\

||

- Tools for ontology generation, merging

- Free CYC?

Summarization

"conceptual" or "content" level diff (email, documents, patents)

Query dependent, Multiple perspective Summarization (representation and output)

∧ ∧ ∧

‖ ‖ ‖

entity discourse co-ref

Multilingual

- interlingua
- deeply annotated data + ML
- user appropriate translations
- English Interlingua

Multimedia

- personalized content based news
- multimedia I/O (maps, gesture)

/\

‖

- multimedia data and annotation (images, maps, video, medical)

Standards

- Process – Reusable interchangeable modules (e.g., POS, NE)
- Data (XML, text encoding, W3C)

NLP

Robust, deep language processing (e.g. LFG parsing which is fast but inaccurate still)

KM/Information Integration

Integrated mining, query of mail, DB, process knowledge

CORE ENABLING RESOURCES

- (intelligent) text annotation (feeds all areas)

- large annotated corpora