Semester Update

(PhD research update,
my research orientation)

Advisor: Venkatesh Choppella, PhD

Sai Gollapudi

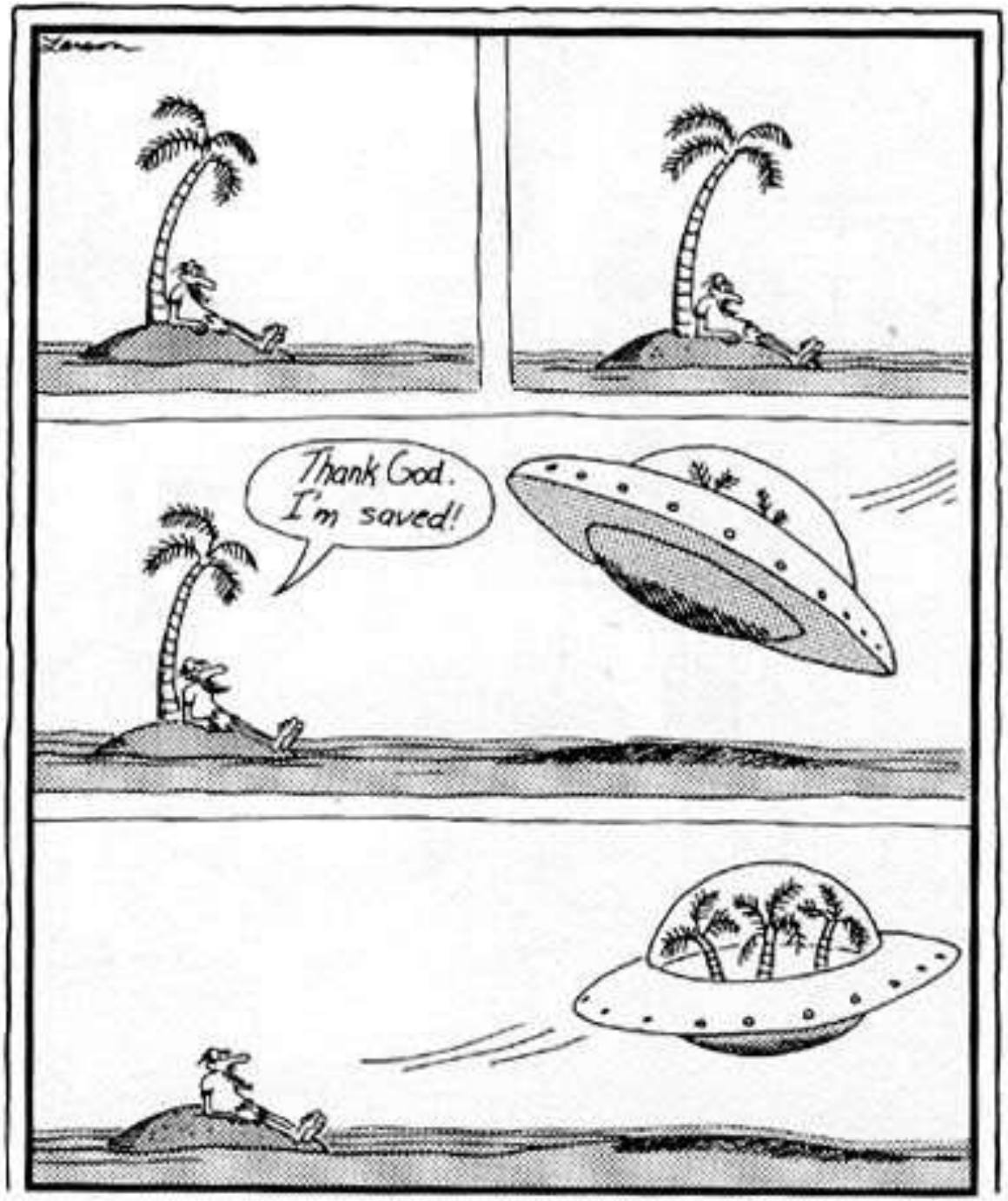
Fall, 2014

Imagine…

- Cartoon
- Website explaining something…

Thank god. I am saved!


Thank God.
I'm saved!


Thank God. I'm saved!

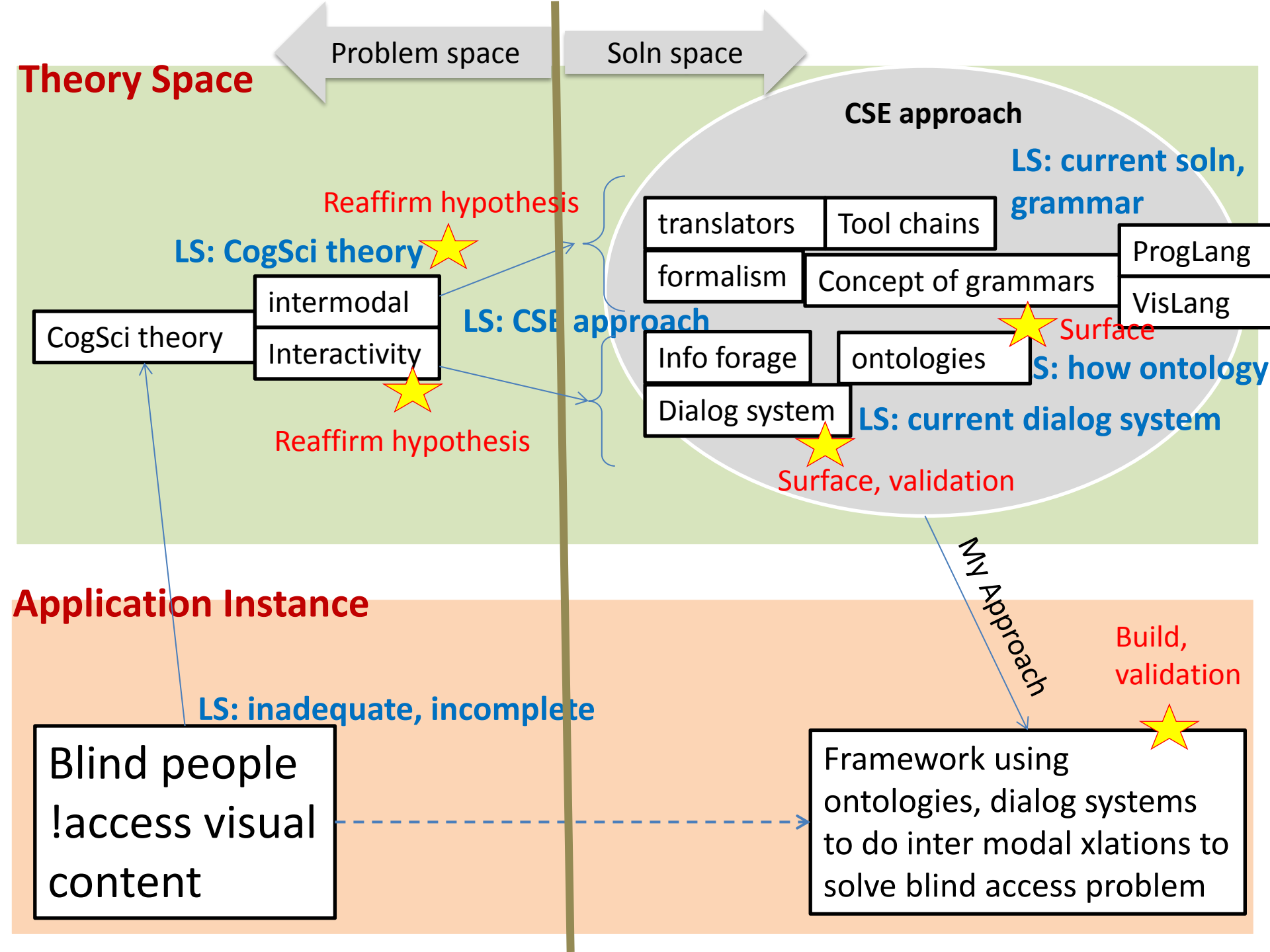
Application Space

- Web Accessibility / Visual renarration
  - Help visually challenged people gain access to informative visual content
- Research goal:
  - Framework to aid visually challenged users get access to structured visual content
- Current systems are inadequate, incomplete
  - Tactile
  - Sonification
  - <Alt-text>

My aspiration (Solution Space)

- Using **cognitive science** principles and current **computer science techniques** and technologies, I wish to come up with an approach & framework to address this accessibility problem

- Solution can be of use to sighted users also in situational blindness cases… or in better understanding representation / manipulation of diagrams


Problem space
Soln space
Theory Space
CSE approach
LS: current soln, grammar
Reaffirm hypothesis
LS: CogSci theory
translators
Tool chains
formalism
Concept of grammars
ProgLang
VisLang
intermodal
CogSci theory
Interactivity
LS: CSE approach
Surface
Info forage
ontologies
S: how ontology
Dialog system
LS: current dialog system
Reaffirm hypothesis
Surface, validation
My Approach
Application Instance
Build, validation
LS: inadequate, incomplete
Blind people !access visual content
Framework using ontologies, dialog systems to do inter modal xlations to solve blind access problem

My proposed (primary) contributions

1. (Developing) an Approach that uses existing CSE components to address CogSci inter-modality and interactivity problem
   - CSE components utilize: formalism, grammar, ontologies, translation systems etc.
2. (Developing) a Framework that uses above approach for providing visual accessibility
   1. Provides the social value of blind-user accessibility
   2. Also an alternate approach for sighted-users

My work this semester

Exploring & crystalizing my problem
Articulating my approach

Clarity

Beginning my literature surveys

Confidence

Targeting publications (W4A) – but which experiment?

Credibility

My Findings…

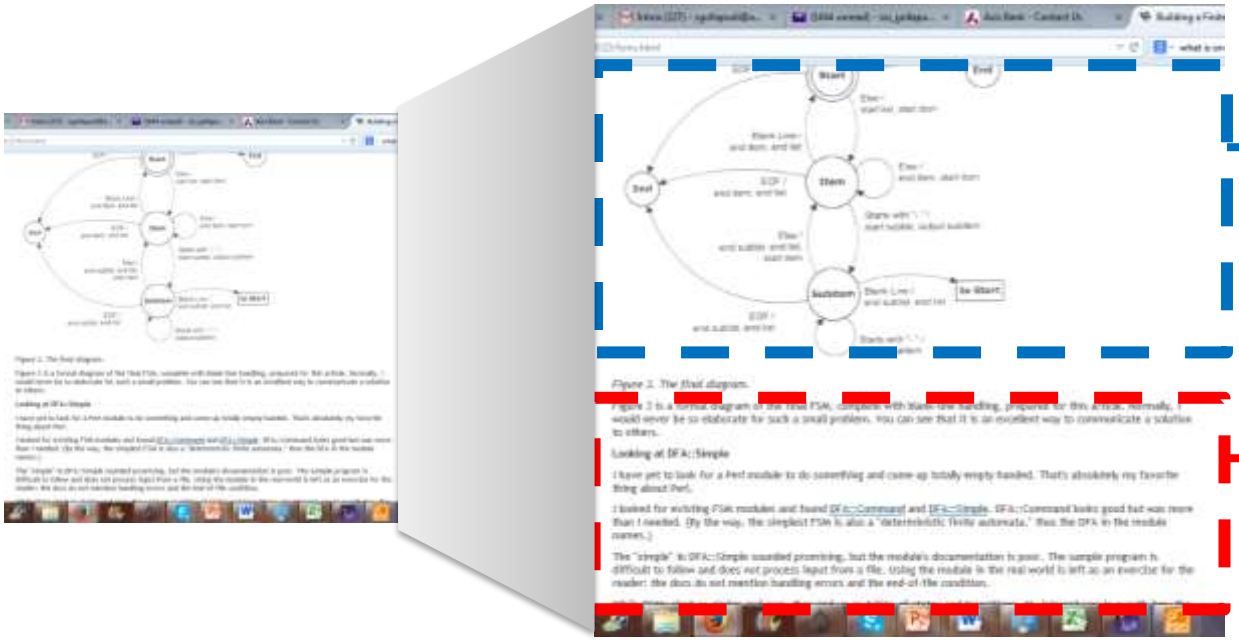


Figure may be presented as brief narrative (given by the <alt text> tag)

Text can be presented as input to either a Braille or to a text-to-speech system

Raw online illustrated input

Current options only address the text portion

Output is provided as text only

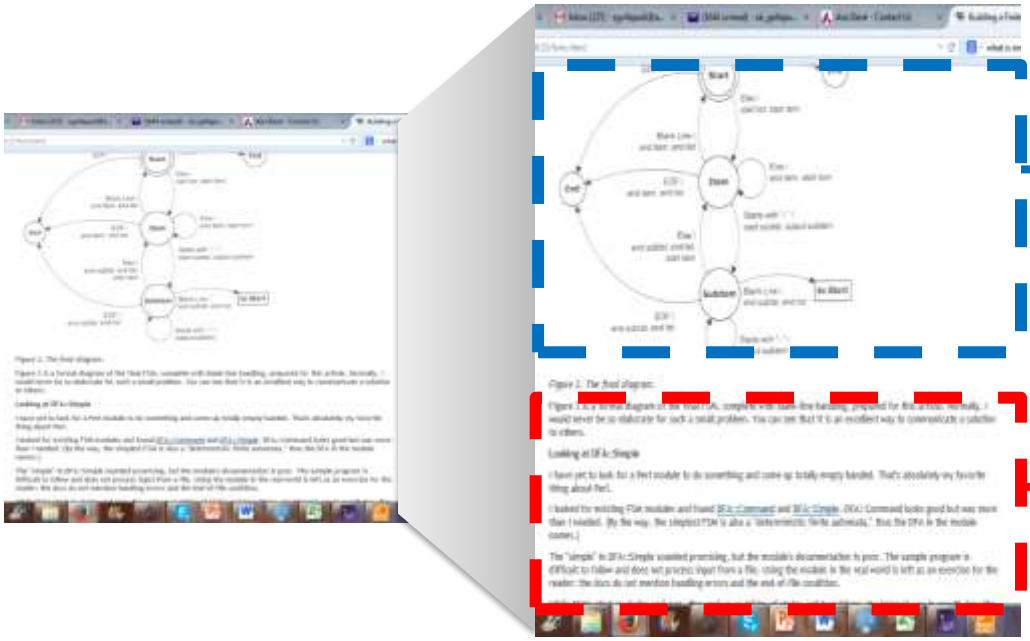
CogSci

- Gaze, Scanpath, Visual Exploration, info foraging
- Partial Hypothesis, observation, validation, improved hypothesis
  - **Interactivity**

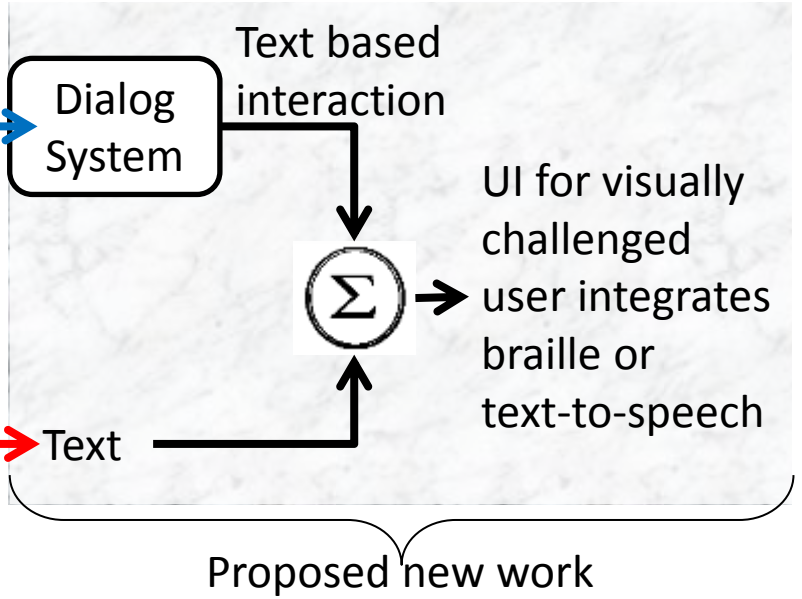
- Brain + Perceptual systems
- Will functions of organ X be replaced by Y in disabled cases? Will the same intent / tendencies of the brain (or neural processes) persist even in the new organ? How does it impact modal translations?
  - **Inter Modal + Translations**

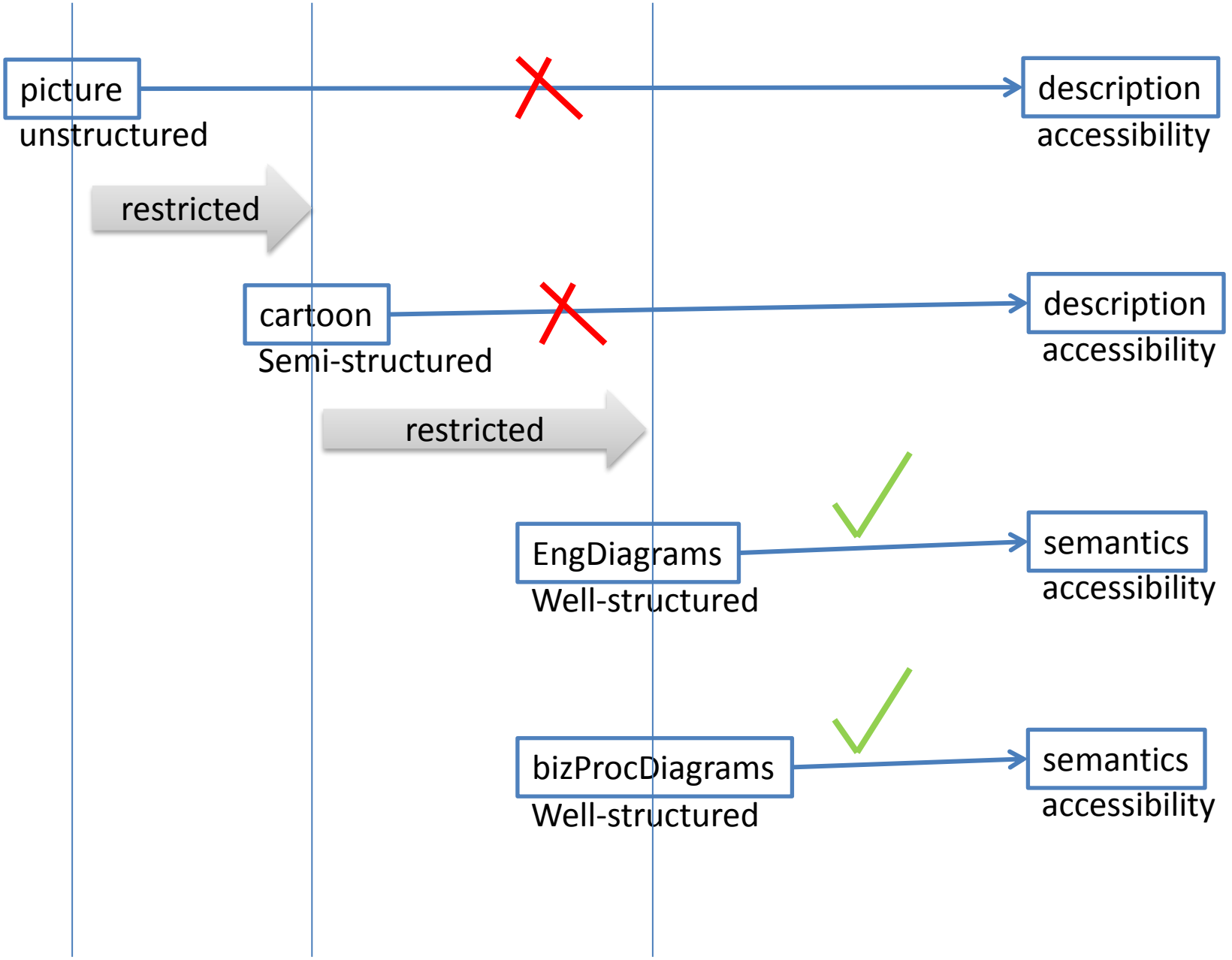
online illustrated input

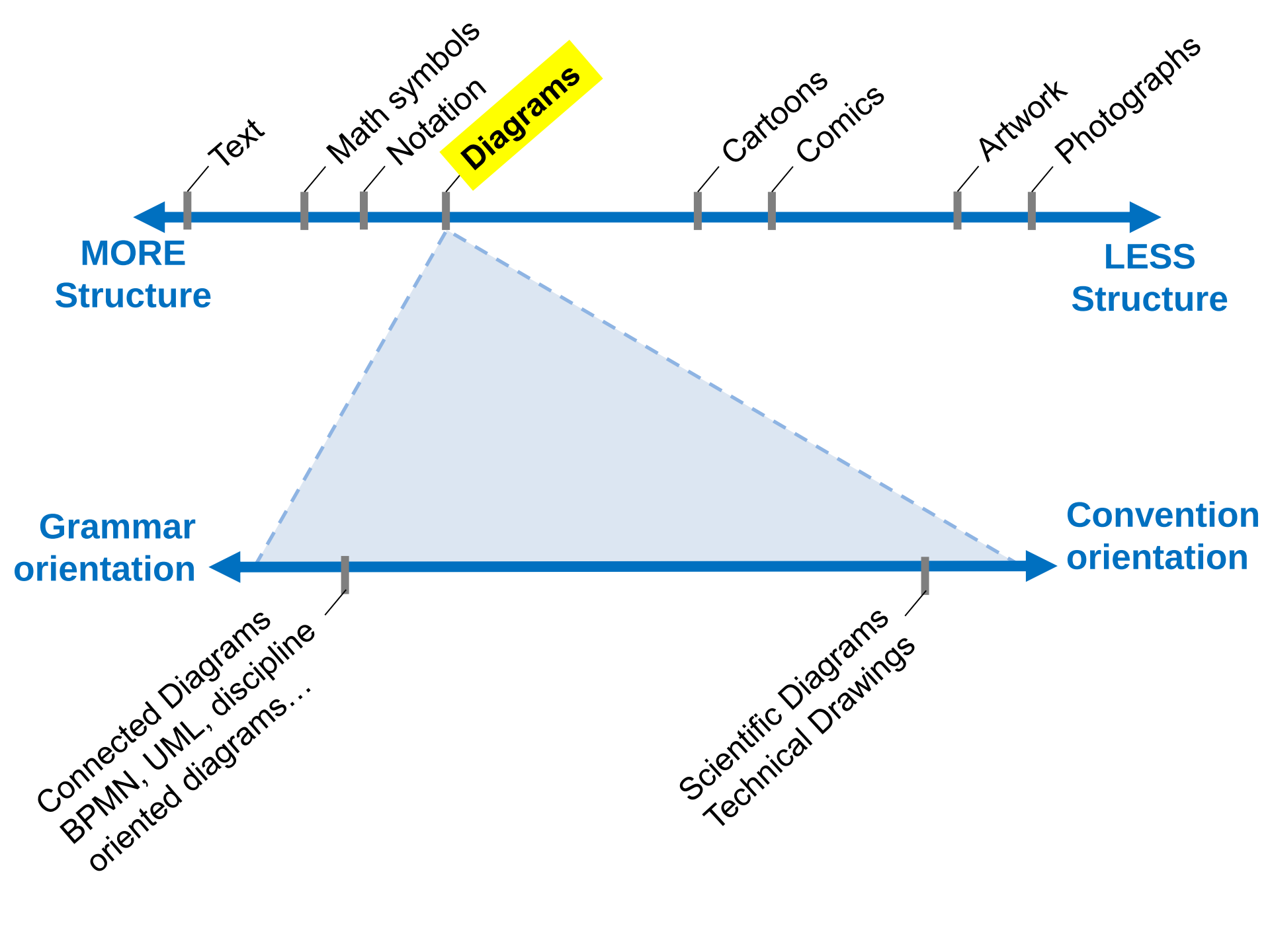
Our system processes the text + diagram

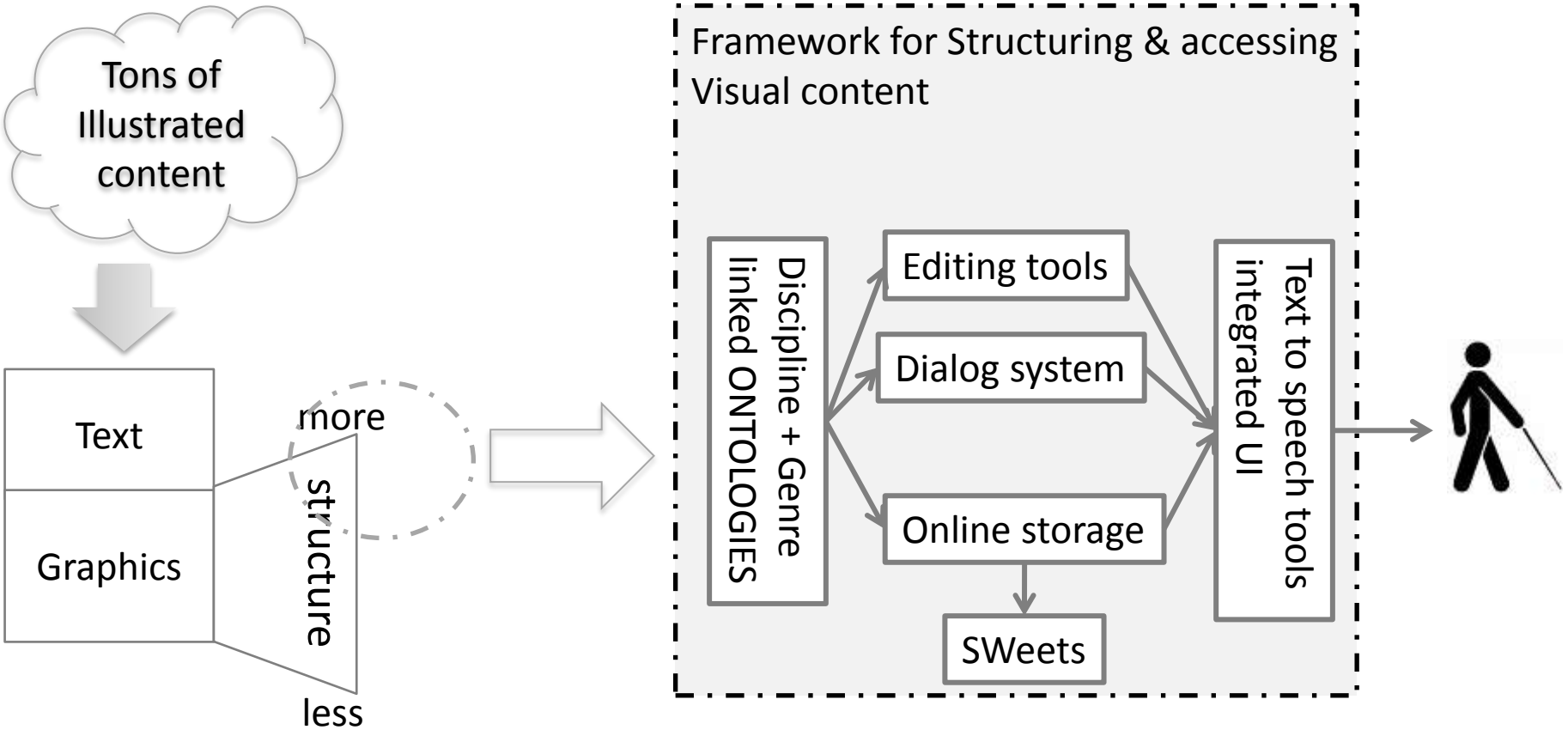


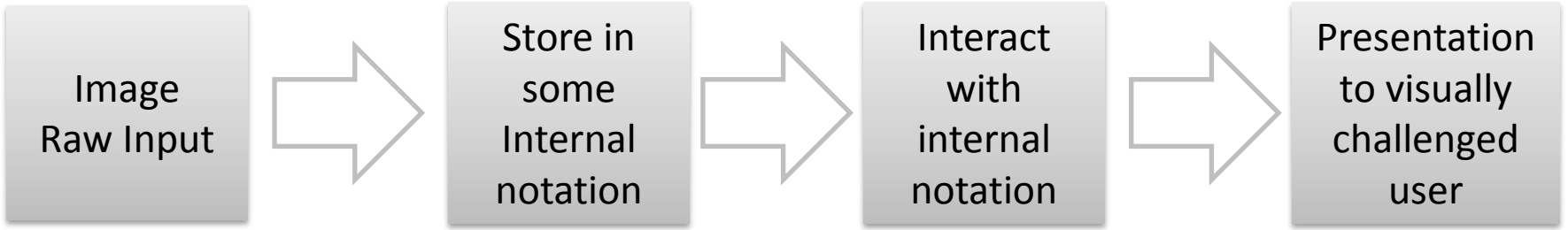
Dialog system to interact with visuals. Plus normal text as inputs to UI

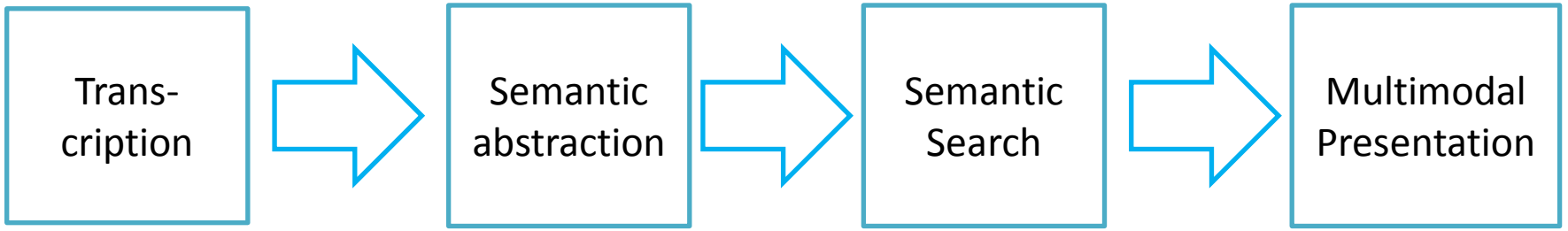



picture
unstructured
description
accessibility
restricted
cartoon
Semi-structured
description
accessibility
restricted
EngDiagrams
Well-structured
semantics
accessibility
bizProcDiagrams
Well-structured
semantics
accessibility


Text
Math symbols
Notation
Diagrams
Cartoons
Comics
Artwork
Photographs
MORE Structure
LESS Structure
Grammar orientation
Convention orientation
Connected Diagrams BPMN, UML, discipline oriented diagrams…
Scientific Diagrams Technical Drawings


Tons of Illustrated content
Text
Graphics
more
structure
less
Framework for Structuring & accessing Visual content
Discipline + Genre linked ONTOLOGIES
Editing tools
Dialog system
Online storage
SWeets
Text to speech tools integrated UI





Techniques:

1. Orig Author AltText
2. Volunteer User Renarration
3. Crowd sourced Img annotation
4. Computer mediated Image Processing
5. Using CAD techn (custom editor)

Techniques:

1. Existing notations: XPDL, DOT
2. Proprietary grammar
3. Ontologies design, use & validation

Techniques:

1. Query engine
2. Stored vs real-time eval
3. Dialogue systems

Techniques:

1. Text
2. Text + audio
3. Multi-track audio
4. Grammar for sound effects track

Text to speech convertor

UI

Natural language Text to Query lang translator

Ontology4 for text queries

Query Engine

Dialog system

Ontology3 for queries

Semantic model

Ontology2 for representation

Internal model

Descriptive model

Ontology1 for representation

XPDL

Next Steps

W4A <- experiments

TACCESS <- clarity in story + significant research findings