Current Research in the UW Structural Informatics Group

January 2011

James F. Brinkley

Shared Vision

- Worldwide network of integrated data and knowledge
- Brings relevant information to point of need
- Fosters translation

Long-term Goals

- Contribute to shared vision
- A structural information framework for biomedicine
- Use as a basis for organizing biomedical information

Rationale

- Structure of the body is the most useful means for organizing biomedical information
- Most manifestations of health and disease are properties of anatomical entities

Primary Informatics Research Questions

- How to create a structural information framework
- How to use this framework as a basis for organizing information

Approach

- Classify types of information
- Represent and implement as different modules
- Integrate within a distributed framework
- Develop applications opportunistically
- Applications drive methods

Current Driving Biological Applications

- Anatomy Education
- Brain mapping
- Proteomics
- Clinical trials
- Craniofacial Malformations

Research Themes

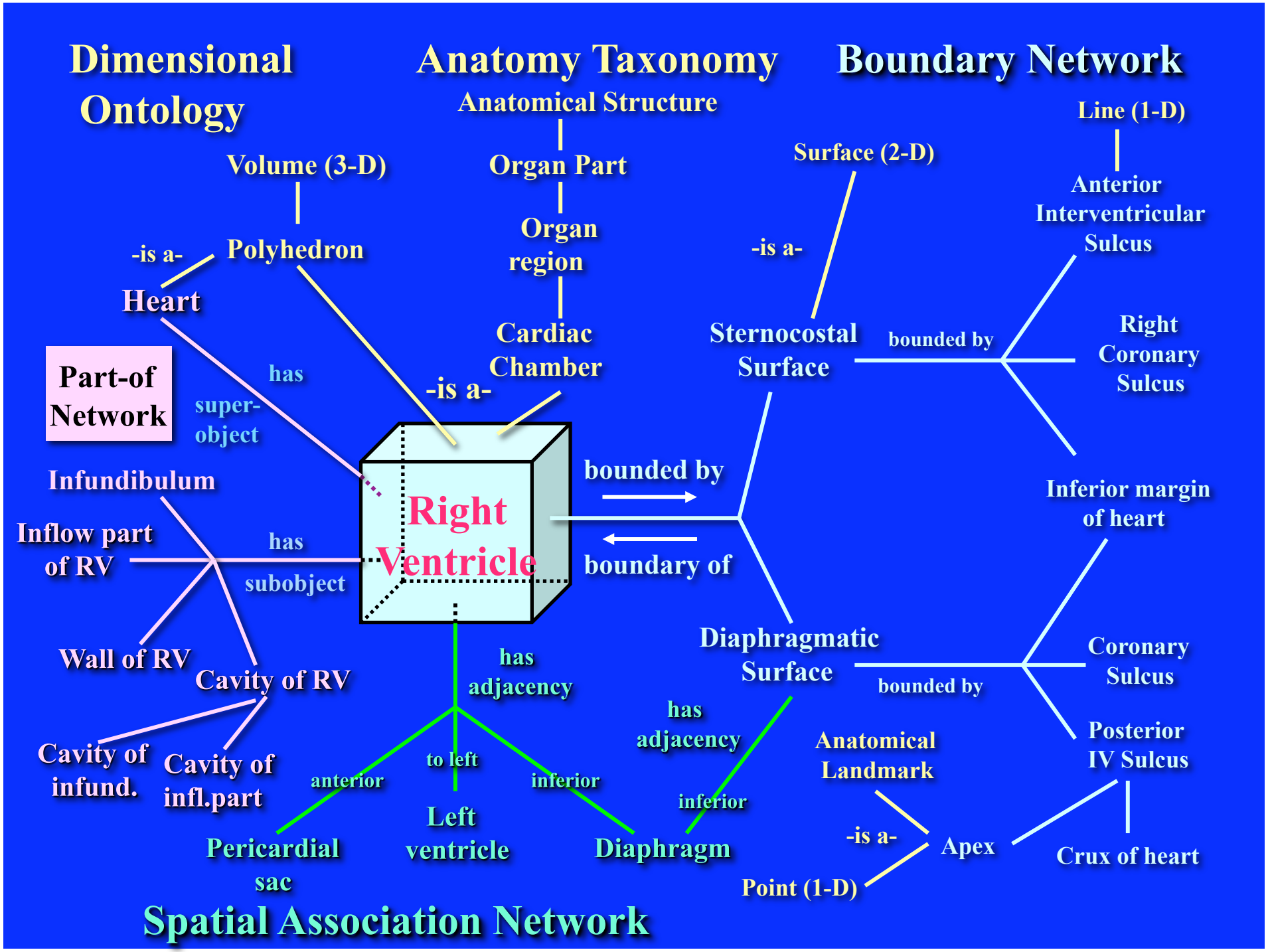
- Structural framework
- Data management
- Data integration
- Data visualization

Projects

- Foundational Model of Anatomy
- Ontology Views
- Data management
- Data integration
- Visualization

Foundational Model of Anatomy

Onard Mejino, Nolan Nichols, Todd Detwiler, Cornelius Rosse


Dimensional Ontology
Anatomy Taxonomy
Boundary Network
Anatomical Structure
Organ Part
Organ region
Cardiac Chamber
Volume (3-D)
Polyhedron
-is a-
Heart
Part-of Network
has
super-object
-is a-
Right Ventricle
bounded by
boundary of
Surface (2-D)
-is a-
Sternocostal Surface
bounded by
Line (1-D)
Anterior Interventricular Sulcus
Right Coronary Sulcus
Inferior margin of heart
Infundibulum
Inflow part of RV
has
subobject
Wall of RV
Cavity of RV
Cavity of infund.
Cavity of infl.part
has adjacency
to left
anterior
inferior
Pericardial sac
Left ventricle
Diaphragm
has adjacency
inferior
Diaphragmatic Surface
bounded by
Coronary Sulcus
Posterior IV Sulcus
Crux of heart
Anatomical Landmark
-is a-
Apex
Point (1-D)
Spatial Association Network

Protege

- Anatomical entity
  - Physical anatomical entity
    - Material physical anatomical entity
      - Anatomical structure
        - Body
        - Body part
        - Organ system
        - Organ
          - Solid organ
          - Cavitated organ
            - Organ with organ cavity
              - Esophagus
              - Stomach
              - Small intestine
              - Large intestine
              - Appendix
              - Anal canal (viewed anatomically)
              - Gallbladder
              - Uterus
              - Uterine tube
              - Urinary bladder
              - Ureter
              - Vagina
              - Urethra
              - Lacrimal duct
              - Pharyngotympanic tube
              - Male genital duct
              - Hollow tree organ
              - Eyeball
              - Neuraxis
              - Serous sac
              - Tooth
              - Membranous labyrinth
            - Organ with cavitated organ parts
        - Organ part
        - Tissue
        - Tissue subdivision
        - Cell
        - Cell part

Preferred Name V: Esophagus

UWDAID: 7131

Synonyms V C X S: Gullet

Non-English Equivalents V C X S: Oesophagus

Definition

Organ with organ cavity which is continuous proximally with the pharynx and distally with the stom
Examples: There is only one esophagus.

Member Of V + −: Set of viscera

Part V: Wall of esophagus, Lumen of esophagus, Cervical part of esophagus, Thoracic part of esophagus, Abdominal part of esophagus, Broncho-esophageus

Part Of V + −: Upper gastrointestinal tract, Foregut

Attributed Part V C

related part	anatomical/arbitrary	shared/unshared	partitio
Wall of esophagus	Anatomical	Unshared	Partition 1
Lumen of esophagus	Anatomical	Unshared	Partition 1
Cervical part of esophagus	Arbitrary	Unshared	Partition 2
Thoracic part of esophagus	Arbitrary	Unshared	Partition 2

Attributed Continuous With V C

related object	coordinate	laterality
Pharynx	Superior	
Stomach	Inferior	

Orientation V C

related object	coordinate	laterality
Plane of pharyngoesophageal ju...	Superior	
Plane of esophagogastric junction	Inferior	

Influence of the FMA

Web Images Videos Maps News Shopping Gmail more ▾

Google scholar "Biomedical Informatics" Search Advanced Scholar Search

Scholar Articles and patents anytime include citations Create email alert

A reference ontology for **biomedical informatics**: the Foundational Model of Anatomy
C Rosse, JLV Mejino - Journal of **biomedical informatics**, 2003 - Elsevier
The Foundational Model of Anatomy (FMA), initially developed as an enhancement of the anatomical content of UMLS, is a domain ontology of the concepts and relationships that pertain to the structural organization of the human body. It encompasses the material objects from the molecular ...
Cited by 519 - Related articles - All 13 versions

[BOOK] **Biomedical informatics**: computer applications in health care and biomedicine
EH Shortliffe… - 2006 - books.google.com
Edward H. Shortliffe, MD, PhD, MACP Department of **Biomedical Informatics** Columbia University Medical Center New York, NY 10032-3720 USA Series Editors: Kathryn J. Hannah, PhD, RN Adjunct Professor Department of Community Health Science Faculty of Medicine The ...
Cited by 531 - Related articles - Library Search - All 14 versions

[BOOK] Evaluation methods in **biomedical informatics**
CP Friedman… - 2006 - books.google.com
Charles P. Friedman Jeremy C. Wyatt Evaluation Methods in **Biomedical Informatics** Second Edition HEALTH INFORMATICS SERIES ... Health Informatics (formerly Computers in Health Care) Kathryn J. Hannah Marion J. Ball Series Editors This Or A3E5-86W-RPDH
Cited by 475 - Related articles - Library Search - All 14 versions

Biomedical informatics for proteomics
MS Boguski… - Nature, 2003 - nature.com
Success in proteomics depends upon careful study design and high-quality biological samples. Advanced information technologies, and also an ability to use existing knowledge to the full, will be crucial in making sense of the data. Despite its genome-scale potential, ...
Cited by 237 - Related articles - BL Direct - All 11 versions

The use of receiver operating characteristic curves in **biomedical informatics**
TA Lasko, JG Bhagwat, KH Zou… - … of **Biomedical Informatics**, 2005 - Elsevier
Receiver operating characteristic (ROC) curves are frequently used in **biomedical informatics** research to evaluate classification and prediction models for decision support, diagnosis, and prognosis. ROC analysis investigates the accuracy of a model's ability to separate ...
Cited by 175 - Related articles - All 15 versions

Current work on the FMA

- Reorganizing the neuroantomy axis of RadLex
- Incorporating Lymphatics
- Representing neural connectivity

Re-organizing the neuroanatomy axis of RadLex

Enabling RadLex with the Foundational Model of Anatomy Ontology to Organize and Integrate Neuro-imaging Data

Jose Leonardo V. Mejino Jr. MD[1], Landon T. Detwiler MS[1], Jessica A. Turner PhD[3], Maryann E. Martone PhD[4], Daniel L. Rubin MD, MS[5] and James F. Brinkley MD, PhD[1,2].
[1]Structural Informatics Group, [2]Biomedical and Health Informatics, University of Washington, [3]MIND Research Network, [4]Department of Neurosciences, University of California San Diego, [5]Stanford Center for Bioinformatics Research, Stanford University

Abstract

In this study we empowered RadLex with a robust ontological framework and additional neuroanatomical content derived from a reference ontology, the Foundational Model of Anatomy Ontology (FMA)[1], with the intent of providing RadLex the facility to correlate the different standards used in annotating neuro-radiological image data. It is the objective of this work to promote data sharing, data harmonization and interoperability between disparate neuro-radiological labeling systems.

Introduction

Huge amounts of neuro-imaging data are being produced by different groups and they are recorded based on different, often study specific, brain parcellation and naming schemes. Using disparate naming conventions produces incompatible terms thereby making the correlation of data difficult to achieve. Current terminologies for neuro-imaging lack the semantic framework to explicitly declare the precise meanings of the terms and therefore neuro-imaging data and information represented by the terms cannot be readily associated and applied across different studies. RadLex[2] (Radiology Lexicon from RSNA) is a controlled terminology for radiology and seeks to provide the needed semantics for correlating the diverse terminologies used for annotating neuro-imaging data. In this work we leveraged the Foundational Model of Anatomy Ontology (FMA) to re-structure and reinforce the anatomical domain of RadLex so it can incorporate, accommodate and correlate the different annotation terminologies.

The approach we describe here serves two practical purposes:

1) reference ontologies provide a principled and robust framework on which to build applications specific to a particular field
2) underlying ontological framework facilitates and promotes integration, interoperability and reuse of knowledge among application ontologies.

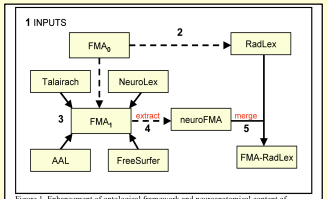


Figure 1. Enhancement of ontological framework and neuroanatomical content of RadLex.

Materials and Methods

Enhancing the neuroanatomy ontology of RadLex involved five major steps (shown in Figure 1):

1. selecting a reference ontology, the FMA ontology, a target application ontology, RadLex, three neuro-imaging annotation terminologies, Talairach Daemon Atlas[3], FreeSurfer atlas[4], Anatomical Automatic Labeling atlas (AAL)[5] and a neuroanatomical application ontology, NeuroLex[6], as inputs to the system;
2. application of the high level class taxonomy of the FMA to re-organize the anatomy hierarchy of RadLex[7];
3. enhancement of the neuroanatomy content of the FMA to capture the representations that different terminologies intend to use for annotating neuroimaging data as well as other forms of neuroscientific data;
4. extraction of the enhanced neuroanatomy component of the FMA, the NeuroFMA, as an ontology "view" for incorporation into RadLex;
5. merging of the extracted NeuroFMA with the ontologically re-organized RadLex .

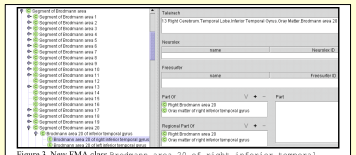


Figure 3. New FMA class `Brodmann area 20 of right inferior temporal gyrus` has been entered to accommodate the Talairach term.

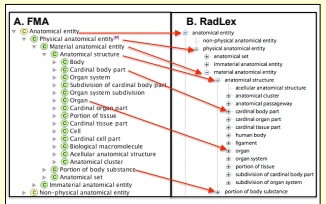


Figure 2. Class taxonomy of FMA applied to RadLex with strict adherence to IS_A only relationships between anatomical entities.

Results

Enhancement of Anatomy Taxonomy of RadLex. Adoption of the ontological framework of the FMA assures a consistent Aristotelian-type inheritance taxonomy for RadLex (Figure 2). The derived ontology provides explicit semantics for RadLex terms .

Enhancement of Neuroanatomy content of FMA. Classes and spatio-structural relations were added in the FMA to accommodate and represent the entities referenced by the different annotation terminologies (Figures 3 and 4). Explicit ontological representation therefore allowed for the correlation of the different terms by using FMA properties such as IS_A and PART_OF (Figure 5).

Extraction of neuroanatomical "view", NeuroFMA, for incoporation into RadLex. View extraction is performed via a procedural program that is written in JAVA, utilizing the Protégé ontology API. Rather than creating a view by starting from an empty ontology and then adding classes, the process starts with a complete copy of the FMA and then eliminates everything not required in the NeuroFMA (Figure 6).

Merging of NeuroFMA into RadLex. Technical details for this step are beyond the scope of this presentation. However we found that we could coalesce classes from the two ontologies in RadLex. The merging produced "FMA-like" structure to RadLex. A total of 12, 579 classes and 33,361 property values were imported into RadLex from the NeuroFMA. It would have been very difficult and time-consuming to implement these changes manually.

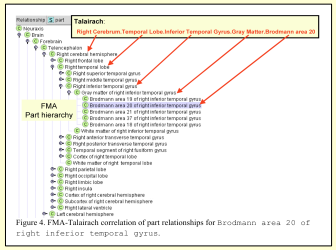


Figure 4. FMA-Talairach correlation of part relationships for `Brodmann area 20 of right inferior temporal gyrus`.

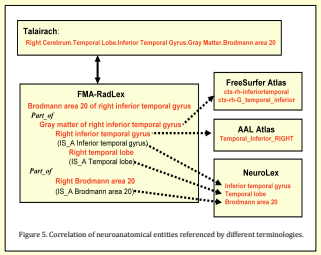


Figure 5. Correlation of neuroanatomical entities referenced by different terminologies.

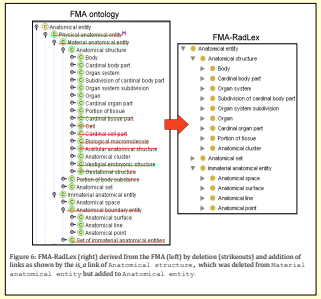


Figure 6: FMA-RadLex (right) derived from the FMA (left) by deletion (strikeouts) and addition of links as shown by the it_a link of `Anatomical structure`, which was deleted from `Material anatomical entity` but added to `Anatomical entity`.

Conclusion

We have shown how the ontological framework of the FMA explicitly defined the entities represented by the different parcellation and naming schemes and by doing so it becomes possible to ascertain the relationships which correlate these terms, a prerequisite step for sharing and harmonizing data. We have started using the ontology to annotate fMRI datasets and derive inferences about relationships between the datasets[8].

Acknowledgment
Supported by NHLBI grant HL08770 and NIBIB contract # HHSN268200800020C.

References

1.Rosse C, Mejino JLV. 2007. The Foundational Model of Anatomy Ontology, In: Burger A, Davidson D, Baldock R (eds). Anatomy Ontologies for Bioinformatics: Principles and Practice, London: Springer, pp 59-117.
2.Langlotz CP. RadLex 2006. a new method for indexing online educational materials. Radiographics 26:1595–1597.
3.Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. Thieme Medical Publishers, New York.
4.Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. Automated Talairach Atlas Labels For Functional Brain Mapping. Human Brain Mapping 10:120-131(2000).
5.N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, Bernard Mazoyer and M. Joliot (January 2002). "Automated Anatomical Labeling of activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI single-subject brain". NeuroImage **15** (1): 273–289.
6.http://neurolex.org/wiki/Main_Page
7.Mejino JLV, Rubin DL, Brinkley JF. FMA-RadLex: An application Ontology of Radiological Anatomy derived from the Foundational Model of Anatomy Reference Ontology. Proc. AMIA Symp 2008:465-469.
8.Turner JA, Mejino JLV, Brinkley JF, Detwiler LT, Lee HJ, Martone ME and Rubin DL (2010) Application of neuroanatomical ontologies for neuroimaging data annotation. *Frontiers in Neuroinformatics* 4(10):pp. 1-12.

Incorporating Lymphatics

FMA class taxonomy

- Lymph node
  - Lymph node of head
  - Lymph node of body proper
    - Lymph node of neck
    - Lymph node of trunk
      - Lymph node of thorax
      - Lymph node of abdomen
        - Lymph node of abdomen proper
          - Parietal lymph node of abdomen proper
          - Visceral lymph node of abdomen proper
            - Intestinal lymph node
              - Pre-aortic lymph node
                - Celiac lymph node
                - Superior mesenteric lymph node
                - Inferior mesenteric lymph node
                - Pre-aortic proper lymph node
              - Gastric lymph node
              - Hepatic lymph node
              - Gastrohepatic ligament node
              - Pyloric lymph node
              - Pancreatic lymph node
              - Pancreaticoduodenal lymph node
              - Colic lymph node
              - Ileocolic lymph node
              - Pararectal lymph node
              - Appendicular lymph node
              - Sigmoid lymph node
              - Superior rectal lymph node
              - Mesenteric lymph node
            - Anorectal lymph node
            - Inferior rectal lymph node
            - Cardiophrenic angle lymph node
            - Retrocrural lymph node
            - Node of posterior iliac crest
        - Lymph node of pelvis
    - Lymph node of upper limb

Celiac lymph node (type=Pre-aortic lymph node)

Preferred Name	FMAID
Celiac lymph node	12792

Synonym	Non-English Equivalent
Celiac node	Nodus lymphaticus coeliacus
No. 9 celiac lymph node	
JPS no. 9 node	
AICC level 20 node	

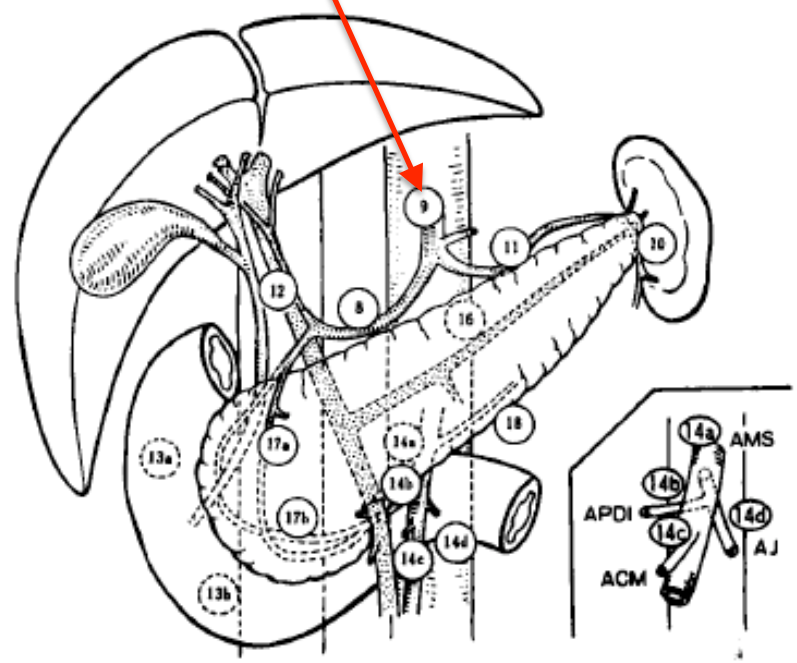


Figure 1. Classification of regional lymph nodes of the pancreas. Inset, subdivision of area 14. AMS: the superior mesenteric artery; AJ: the jejunal artery; APDI: the inferior pancreaticoduodenal artery; ACM: the medial colic artery.

Nakagawa et. al. CANCER 1994, Vol. 73, No. 4, 1155-1162.

Representing Neural Connectivity

Representing Neural Connectivity in the Foundational Model of Anatomy Ontology

Nolan Nichols[1], Aaron Perlmutter[1], Jose L. V. Mejino Jr. [2], Daniel L. Rubin[3], and James F. Brinkley[1,2,4]
Departments of Medical Education and Biomedical Informatics[1], Biological Structure[2], Computer Science and Engineering[4]
University of Washington, Seattle, WA, Stanford Center for Bioinformatics Research, Stanford University[3]

1. Summary

• Our current effort focuses on representing neural connectivity relationships between gray and white matter structures in the Foundational Model of Anatomy Ontology (FMA)[1].

• The FMA contains a number of terms that imply either structural or functional connectivity, such as sends_output_to or receives_input_from, but the semantics of their structural connectivity relationships were not yet made formally explicit.

• To formalize structural connectivity relationships in the FMA we developed a set of definitions to disambiguate and clarify the terminologies describing the types of connectivity relationships that exist between gray and white matter structures at different levels of granularity.

• Scales of connectivity relations range from long-range association, commissural, and projection fibers at the mesoscopic scale to synaptic junctions at the microscopic scale.

• This work focused on generating a representation of connectivity at the mesoscopic scale, which aims to facilitate the integration of annotated open-access neuroimaging datasets - including structural MRI (sMRI), functional MRI (fMRI) and diffusion tensor imaging (DTI).

2. Methodology

• The FMA is a reference ontology for the domain of anatomy that symbolically represents the phenotypic organization of the human body at all levels of granularity.

• In this study we applied FMA principles to the represent structural connectivity properties of gray matter and white matter neural structures using the principle of Anatomical Structural Abstraction (ASA).

• *Connectivity*, in addition to *Location* and *Orientation*, is one of the three components of the ASA Spatial Association Network (SAn). We focused on explicitly representing connectivity properties between white matter and gray matter neural entities at the mesoscopic scale.

• **Gray Matter Properties**
- *has_projection* and *projects_from* are inverse properties between "structure of origin" and the fibers it sends out.
  - 'Brodmann area 39 of inferior parietal lobule' has_projection 'Superior longitudinal fasciculus proper'
  - 'Superior longitudinal fasciculus proper' *projects_from* 'Brodmann area 6 of inferior parietal lobule'

•**White Matter Properties**
- *receives_projection* and *projects_to* are inverse properties between "structure of termination (target)" and the fibers it receives.
  - 'Dorsal segment of superior longitudinal fasciculus' *projects_to* 'Brodmann area 6 of superior frontal gyrus'
  - 'Brodmann area 6 of superior frontal gyrus' *receives_projection* from 'Superior longitudinal fasciculus proper'

3. Example of a Connectivity Relation

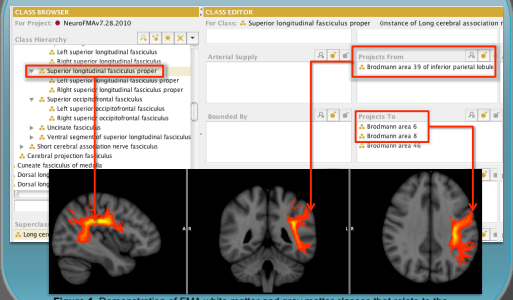


Figure 1. Demonstration of FMA white matter and gray matter classes that relate to the JHU DTI-81 atlas probability distribution of the left superior longitudinal fasciculus.

4. Relate Tract and Region Based Atlases

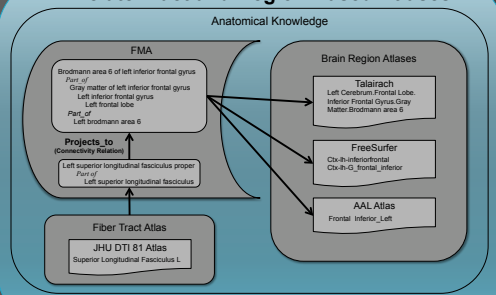


Figure 2. Labels are used to annotate white matter structures in the JHU-DTI-81 fiber tract atlas and grey matter structures in atlases based on brain regions. The FMA can be used to reconcile the different parcellation schemas..

5. Sub-property Hierarchy and Definitions

Sub-property Hierarchy

Connectivity Hierarchy
- connected_to
  - attached_to*
    - innervates
    - projects_to*
  - receives_attachment_from*
    - innervated_by
    - receives_projection _from*
  - continuous_with*
    - has_projection*
    - projects_from*
  - synapses_with*
  - has pathway
    - has efferent pathway
      - sends_output_to*
    - has afferent pathway
      - receives_input_from*

Term	Definition
Connected_to	Structural anatomical property which holds between each anatomical structure of type A and some anatomical structure of type B such that each structure shares some part of its bona fide anatomical surface with that of the other.
Continuous_with	Connected_to property which holds between each anatomical entity of type A and some anatomical entity of type B such that there is no bona fide boundary between their contiguous constitutional parts.
Synapse_with	Connected_to property where there is apposition between the presynaptic membrane of a neurite of one neuron and the postsynaptic membrane of one or more neurites of another neuron or a region of a muscle cell or a gland cell and some form of neurotransmission is evident between them.
Projects_to*	Attached_to property where individual axons comprising a fiber tract originating from one or more brain regions synapse_with neurites or somas of a collection of neurons located in one or more other brain regions. This relation may be synonymous with '*terminates_in*'.
Projects_from*	Continuous_with property where individual axons comprising a fiber tract are parts of a collection of neurons located in one or more brain regions. This relation may be synonymous with '*originates_from*'.
Sends_output_to*	Efferent pathway property consisting of relations where A has_projection B and B projects_to C, and where neuro-mission is sent from A to C.
Receives_input_from*	Afferent pathway property consisting of relations where A receives_projection_from B and B projects_from C, and where neurotransmission is received by A from C.

Figure 3. Connectivity properties arranged in a hierarchy in the Spatial Association Network (SAn) of the FMA (above). Definitions for properties implemented in the FMA* (right).

6. Future Work

• Continue developing the FMA representation of connectivity at the mesoscopic scale and begin implementation of additional connectivity relations at finer levels of granularity.

• Develop a more rich representation of the Brodmann Area parcellation schema by explicitly defining cytoarchitectonic, regional, and long-range connectivity properties.

• Demonstrate the utility of connectivity relations in the FMA for knowledge discovery by integrating region-based annotations from sMRI and fMRI datasets with tract-based annotations from DTI datasets.

• Determine how new knowledge about neural connectivity from the human connectome project and related structural and functional connectivity research can can be incorporated into the FMA.

Acknowledgement
Supported by NHLBI grant HL08770 and NIBIB contract # HHSN268200800020C.

References

1. Rosse C and Mejino JLV. (2007) The Foundational Model of Anatomy Ontology, in Burger A, Davidson D and Baldock R. Eds. Anatomy Ontologies for Bioinformatics. Principles and Practice, pages pp. 59-117. Springer.
2. Detwiler LT, Suciu, D, Franklin JD, Moore EB, Poliakov AV, Lee ES, Corina DP, Ojemann GA and Brinkley, JF. (2009) Distributed XQuery-based integration and visualization of multimodality data: Application to brain mapping. Frontiers in Neuroinformatics.
3. Turner JA, Mejino JLV, Brinkley JF, Detwiler LT, Lee HJ, Martone ME, and Rubin DL. (2010) Application of neuroanatomical ontologies for neuroimaging data annotation. Frontiers in Neuroinformatics.
4. Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PCM, Mori S. (2004) Fiber tract-based atlas of human white matter anatomy. Radiology.

AMIA 2010 Annual Symposium Washington, DC November 13-17, 2010

Planned work on the FMA

- Convert to OWL 2
- Use as a basis for application ontologies
- Use as a basis for reorganizing other ontologies

Projects

- Foundational Model of Anatomy
- <u>Ontology Views</u>
- Data management
- Data integration
- Visualization

Ontology Views over the Semantic Web

Todd Detwiler, Marianne Shaw, Onard Mejino,
Dan Suciu, Linda Shapiro, John Gennari, Dan Cook,
Nicola Dell, Natasha Noy, Daniel Rubin,
Mark Musen

Motivation

- Reference ontologies like the FMA are (or will be) too large for practical use
- How can reference ontologies be made practical for applications?

Approach

- Application ontologies as views over one or more reference ontologies
- A view is a query that defines a formal transformation from one or more source ontologies to a target application ontology

SparQL Extensions: vSparQL

Marianne Shaw, Todd Detwiler and Dan Suciu

- Gleen Regular path library
- Subqueries
- Recursive Queries
- Skolem functions

vSparQL Subquery

Subquery

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX gleen:<java:edu.washington.sig.gleen.>
PREFIX fma:<http://sig.biostr.washington.edu/fma3.0#>

SELECT ?subj ?prop ?obj
FROM NAMED <liver> [
        CONSTRUCT {fma:Liver fma:regional_part ?z}
        FROM <http://sig.biostr.washington.edu/fma3.0>
        WHERE {fma:Liver fma:regional_part ?z .}
]
WHERE { GRAPH <liver> {?subj ?prop ?obj .} }
```

subj	prop	obj
fma:Liver	fma:regional_part	fma:Right_lobe_of_liver
fma:Liver	fma:regional_part	fma:Left_lobe_of_liver

vSparQL Recursive Query

Base case

Recursive step

```
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX gleen:<java:edu.washington.sig.gleen.>
PREFIX fma:<http://sig.biostr.washington.edu/fma3.0#>

SELECT ?subj ?prop ?obj
FROM NAMED <liver> [
        CONSTRUCT {fma:Liver fma:regional_part ?z}
        FROM <http://sig.biostr.washington.edu/fma3.0>
        WHERE {fma:Liver fma:regional_part ?z .}

        UNION
        # Get all of the regional_parts of the liver recursively
        CONSTRUCT {?c fma:regional_part ?d}
        FROM NAMED <liver>
        FROM NAMED <http://sig.biostr.washington.edu/fma3.0>
        WHERE {
            GRAPH <liver>
                  { ?a ?b ?c . } .
            GRAPH <http://sig.biostr.washington.edu/fma3.0>
                  { ?c fma:regional_part ?d . } .
        }
]
WHERE { GRAPH <liver> {?subj ?prop ?obj .} }
```

vSparQL Recursive Query Results

subj	prop	obj
fma:Left_lobe_proper_of_liver	fma:regional_part	fma:Medial_segment_of_left_lobe_of_liver
fma:Left_lobe_proper_of_liver	fma:regional_part	fma:Lateral_segment_of_left_lobe_of_liver
fma:Lateral_segment_of_left_lobe_of_liver	fma:regional_part	fma:Lateral_superior_area_of_lateral_segment_of_left_lobe_of_liver
fma:Lateral_segment_of_left_lobe_of_liver	fma:regional_part	fma:Lateral_inferior_area_of_lateral_segment_of_left_lobe_of_liver
fma:Left_lobe_of_liver	fma:regional_part	fma:Left_lobe_proper_of_liver
fma:Left_lobe_of_liver	fma:regional_part	fma:Caudate_lobe_of_liver
fma:Left_lobe_of_liver	fma:regional_part	fma:Quadrate_lobe_of_liver
fma:Right_lobe_of_liver	fma:regional_part	fma:Anterior_segment_of_right_lobe_of_liver
fma:Right_lobe_of_liver	fma:regional_part	fma:Posterior_segment_of_right_lobe_of_liver
fma:Anterior_segment_of_right_lobe_of_liver	fma:regional_part	fma:Anterior_inferior_area_of_anterior_segment_of_right_lobe_of_liver
fma:Anterior_segment_of_right_lobe_of_liver	fma:regional_part	fma:Anterior_superior_area_of_anterior_segment_of_right_lobe_of_liver
fma:Liver	fma:regional_part	fma:Left_lobe_of_liver
fma:Liver	fma:regional_part	fma:Right_lobe_of_liver
fma:Caudate_lobe_of_liver	fma:regional_part	fma:Right_segment_of_caudate_lobe_of_liver
fma:Caudate_lobe_of_liver	fma:regional_part	fma:Papillary_process_of_caudate_lobe_of_liver
fma:Caudate_lobe_of_liver	fma:regional_part	fma:Left_segment_of_caudate_lobe_of_liver
fma:Caudate_lobe_of_liver	fma:regional_part	fma:Caudate_process_of_caudate_lobe_of_liver
fma:Medial_segment_of_left_lobe_of_liver	fma:regional_part	fma:Medial_inferior_area_of_medial_segment_of_left_lobe_of_liver
fma:Medial_segment_of_left_lobe_of_liver	fma:regional_part	fma:Medial_superior_area_of_medial_segment_of_left_lobe_of_liver
fma:Posterior_segment_of_right_lobe_of_liver	fma:regional_part	fma:Posterior_superior_area_of_posterior_segment_of_right_lobe_of_liver
fma:Posterior_segment_of_right_lobe_of_liver	fma:regional_part	fma:Posterior_inferior_area_of_posterior_segment_of_right_lobe_of_liver

YJBIN 1697
28 August 2010
ARTICLE IN PRESS
No. of Pages 18, Model 5G

Journal of Biomedical Informatics xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

vSPARQL: A view definition language for the semantic web

Marianne Shaw [a,*], Landon T. Detwiler [b], Natalya Noy [c], James Brinkley [a,b], Dan Suciu [a]

[a] University of Washington, Computer Science and Engineering, Box 352350, Seattle, WA 98195, USA
[b] University of Washington, Structural Informatics Group, Biological Structure, Seattle, WA 98195, USA
[c] Stanford University, Center for Biomedical Informatics Research, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 1 December 2009
Available online xxxx

Keywords:
Translational medicine applications
RDF views
Ontologies
Vocabularies

ABSTRACT

Translational medicine applications would like to leverage the biological and biomedical ontologies, vocabularies, and data sets available on the semantic web. We present a general solution for RDF information set reuse inspired by database views. Our view definition language, vSPARQL, allows applications to specify the exact content that they are interested in and how that content should be restructured or modified. Applications can access relevant content by querying against these view definitions. We evaluate the expressivity of our approach by defining views for practical use cases and comparing our view definition language to existing query languages.

1. Introduction

The semantic web seeks to enable computers to automatically associate and process well-defined information on the web. To that end, a number of biological and biomedical information sets have been developed for or converted to semantic web formats. These information sets include vocabularies, ontologies, and data sets; they may be available in basic RDF [1] or languages with higher-level semantics, such as OWL [2].

Translational medicine applications want to leverage the biomedical information sets available on the semantic web. For example, radiologists may want an application for viewing and annotating medical images. The application can aid annotation by offering suggestions for visible anatomical parts from the Foundational Model of Anatomy (FMA) [10]. Additionally, if suspicious nodes are identified on an image, the application can offer suggestions from NCI Thesaurus [21] for annotating the nodes.

This paper asks the question: *how can we make it easier for applications to leverage the information sets available on the semantic web?* In leveraging these information sets, applications must contend with their size, content depth and breadth, and rate of change.

Many biomedical information sets are of significant size. For example, Reactome [22] as distributed through NeuroCommons contains 3.6M triples, the FMA 3.0 in OWL contains 1.7M triples, and UniProt [5] v13.4 contains 1.7B triples. Additionally, medical applications may require information from multiple information sets, such as the FMA/NCI example above. The combined information sets may be too large for applications to easily manipulate.

Applications may only need a portion of the content contained in an information set, which are likely to have a wider scope and greater depth than that needed by any particular application. Applications determine the set of relevant content needed from an information set; the set could be a list of concepts, a subgraph of related terms, a sub-ontology, or any other subset of information. Techniques ([9,41,39,43]) have been developed for some of these scenarios and are discussed in related work.

While leveraging biomedical information sets, applications may wish to augment, modify, or restructure selected content. For example, an application may seek to define a new relationship between selected concepts. Alternatively, an application may disagree with a portion of the information set (e.g. the content is too coarsely specified) and need to "fix" it.

In addition to specifying relevant content, applications must accommodate updates to information sets. If an application includes a materialized subset, new materialized subsets may need to be derived and distributed every time an update occurs. An application that queries the information set accesses the most up-to-date content; however, the queries must be written to restrict responses to the application-specified content.

We propose a general solution for information set reuse. We validate our approach with eight different use cases over one or more of four biomedical data sets. Our approach is inspired by database views. A database view identifies a subset of a dataset that can be accessed through the view; queries on the view can only access the specified subset. However, a database view may export a restructured or modified version of the data; the exported view is not limited to faithfully replicating a subset of the underlying information.

* Corresponding author.
E-mail addresses: mar@cs.washington.edu (M. Shaw), det@u.washington.edu (L.T. Detwiler), noy@stanford.edu (N. Noy), brinkley@u.washington.edu (J. Brinkley), suciu@cs.washington.edu (D. Suciu).

1532-0464/$ - see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.jbi.2010.08.008

Graphical User Interface over vSPARQL

An Application for Creating and Debugging vSPARQL View Definitions

Marianne Shaw, MS[1], Landon T. Detwiler, MS[2], James F. Brinkley, MD, PhD[2], and Dan Suciu, PhD[1]
[1]Computer Science & Engineering, [2]Structural Informatics Group, University of Washington, Seattle, WA

Background

•Biomedical information sets (ontologies, thesauri, data sets) are being made available for the semantic web in RDF-compatible formats.

- Foundational Model of Anatomy
- Reactome
- NCI Thesaurus

•Medical and biomedical applications want to leverage these information sets. However, these information sets are typically very large and comprehensive; they contain more information than is needed by an application.

•vSPARQL extends SPARQL (W3C query language for RDF) to enable researchers to create view definitions that extract, modify, combine, and augment the relevant information from these RDF graphs.

•vSPARQL allows subqueries and recursive subqueries to be defined and combined to create view definitions. The subqueries allow view definitions to be broken down into small units. Recursive subqueries allow data to be extracted by iteratively traversing paths of arbitrary length.

vSPARQL GUI

•The vSPARQL GUI was designed to help users develop and debug view definitions.

• The GUI enables users to visualize a view definition as a dataflow graph connecting data sources, (recursive) subqueries, and SPARQL queries. Each of these language constructs are component blocks in the leftmost panel.

•Selecting a component in the leftmost panel adds an instance of it to the overall view. Within a block, users fill in: SPARQL graph patterns and results, data sources, and a name for intermediate result graphs.

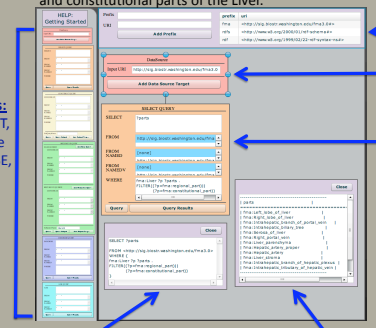
•Users connect the output of one component to the input of another; the output of these blocks can be accessed via pull-down lists. These connections create a graph through which intermediate results flow.

• The vSPARQL GUI helps researchers debug view definitions as they are being created. All query form and (recursive) subquery component blocks have a "Query output" button. Pressing the button causes the view definition *up to that point in the dataflow graph* to be executed. The results are displayed in two formats: RDF/XML and a list of triples.

•The "Query output" button allows researchers to "probe" specific points in the dataflow graph to determine if the results are correct before combining them with other results.

•Funded by NIH HL087706.

SPARQL SELECT query getting all of the regional and constitutional parts of the Liver.



Component Blocks: data source, SELECT, subquery, recursive subquery, DESCRIBE, ASK

Prefixes for URI namespaces

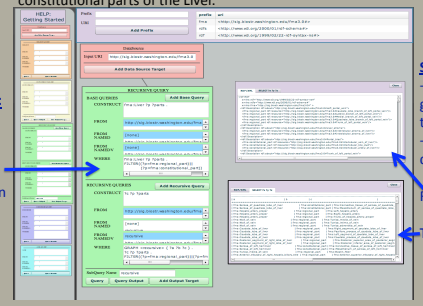
Data source: Specify URI of RDF graph

SELECT Query Block
- Specify variables to project.
- Indicate which incoming data source to query using FROM pull-down lists.
- Specify desired graph pattern in WHERE textbox.

Query Output: Displays vSPARQL query at this point in the dataflow graph.

SELECT Query Result: Displays the bindings for the projected variables.

Recursive subquery getting, transitively, all of the regional and constitutional parts of the Liver.



Recursive Query Block:
-Specify base case to seed result graph.
-Specify recursive case; this case is repeatedly executed, building upon the contents of the result graph, until no new data is added to the result graph.

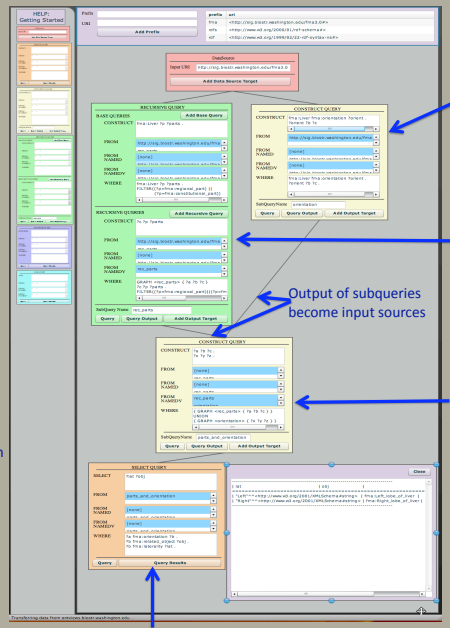
Subquery Result:
-Displays the output graph at this point in the dataflow graph.
-Two tabs for two different formats.

RDF/XML

Triples

Demo: http://ontviews.biostr.washington.edu:8080/VSparQL_Service/GUI/vsparql_oh_degrafa.html

vSPARQL view definition combining all (transitive) regional and constitutional parts of the liver with information about the liver's orientation. The SELECT query projects the anatomical objects related to the liver and their relative laterality.



Subquery Block:
- Specify pattern to add to output graph when WHERE clause pattern found.
- Provide name (orientation) for output graph.

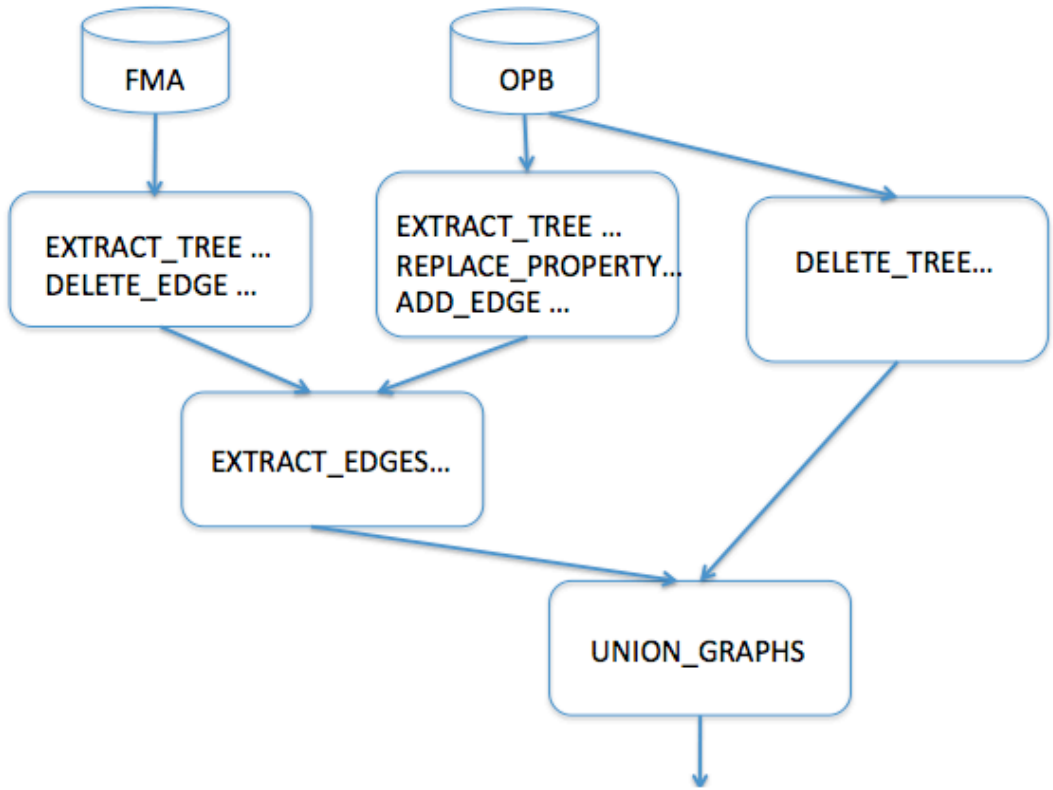
Recursive Subquery Block:
-Transitively find regional & constitutional parts of liver.
-Name output graph (rec_parts).

Subquery Block:
-Combine input sources into a single graph.
-Using FROM NAMEDV allows blank nodes to be comparable.

Output of subqueries become input sources

SELECT Query Block:
-Query over view definition results.

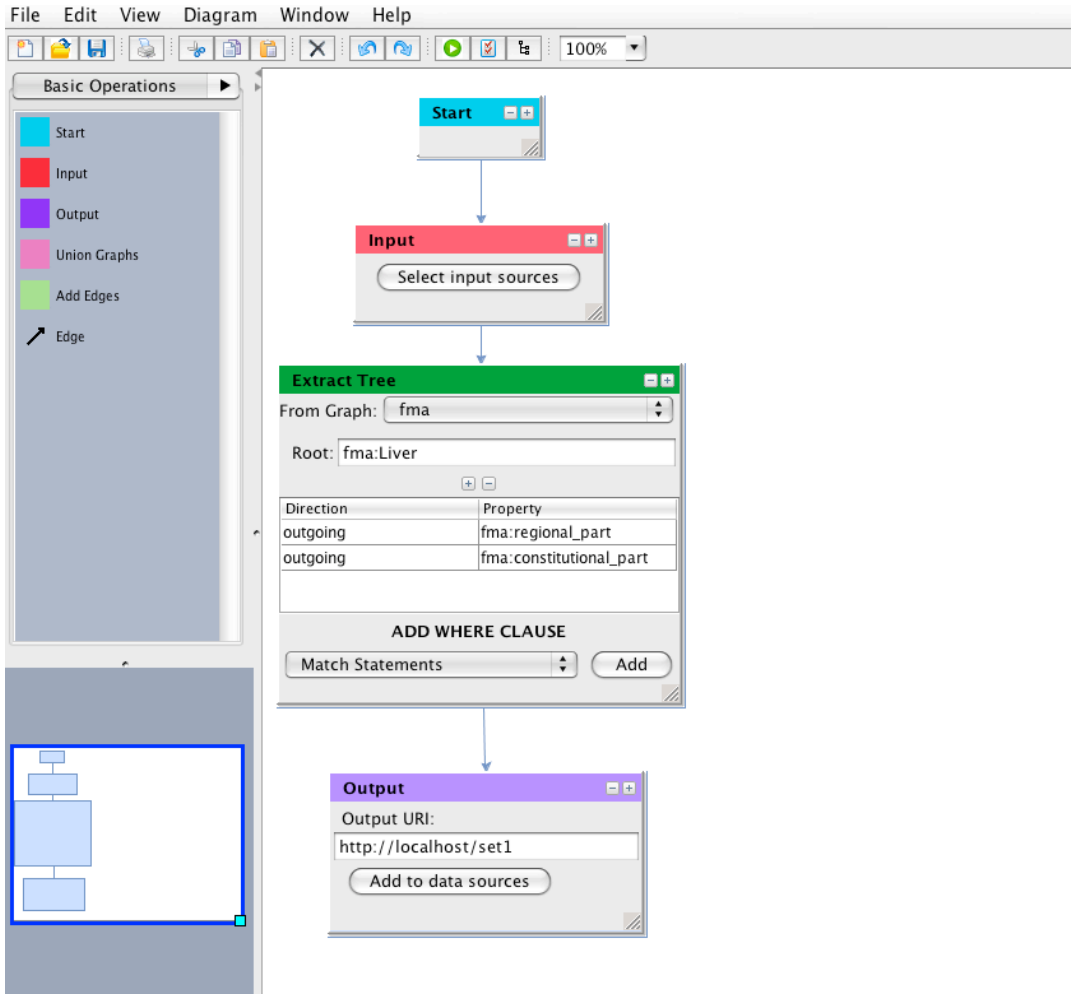
Intermediate Language (IML)



Recursive Liver Parts in IML

```
PREFIX fma: <http://sig.biostr.washington.edu/fma3.0#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX holder: <http://localhost/holder#>
PREFIX VSparQL: <java:vsparql.ext.#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

INPUT
<http://sig.biostr.washington.edu/fma3.0>{
EXTRACT_TREE { fma:Liver [ forward (fma:regional_part) , forward (fma:constitutional_part)  ]  }
GRAPH  <http://sig.biostr.washington.edu/fma3.0>
}
OUTPUT <http://localhost/set1>
```

Graphical User Interface over IML



Generated IML and Results

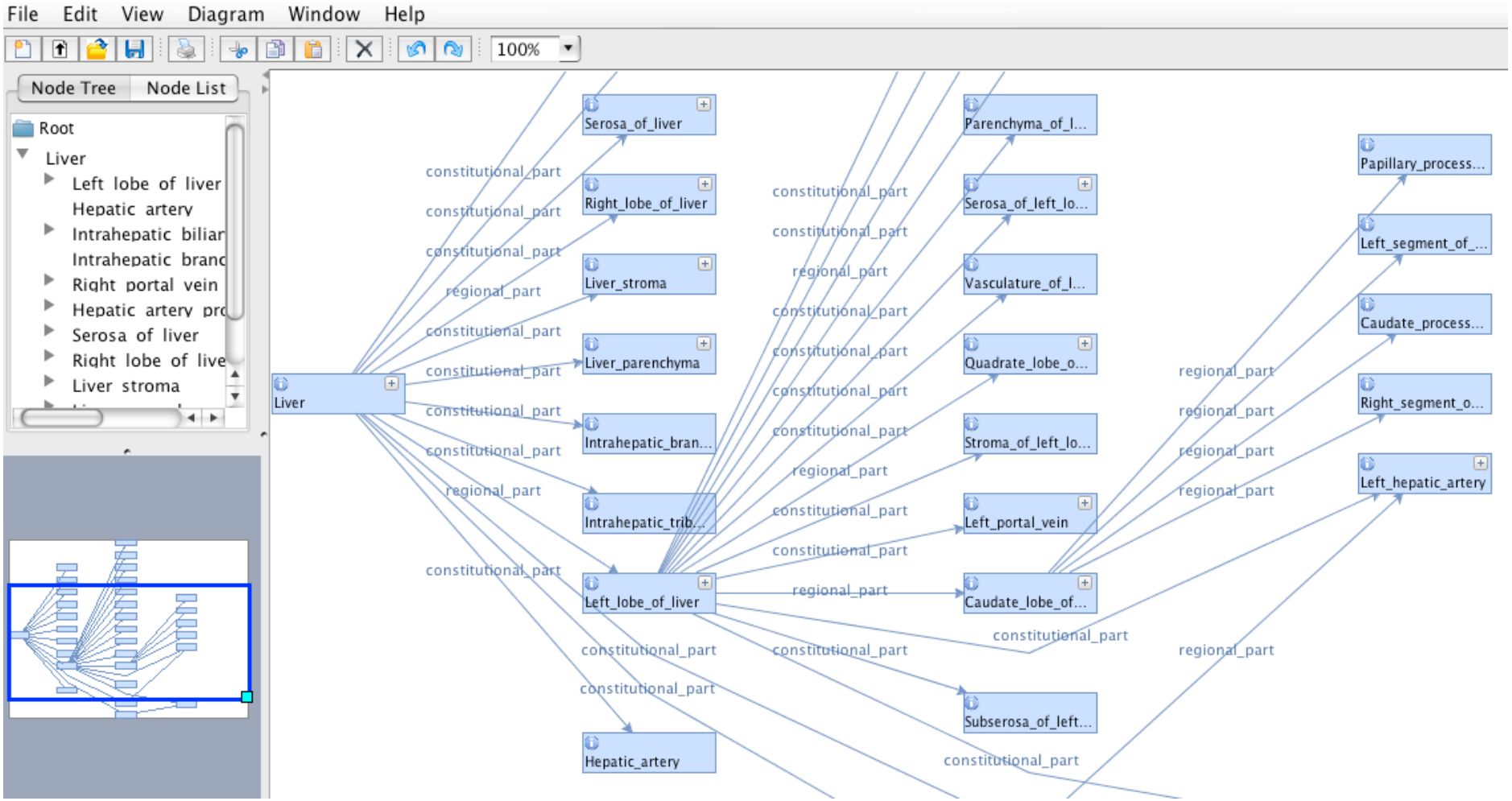
Generated Query

Save query | Save results | Run query | Visualize results

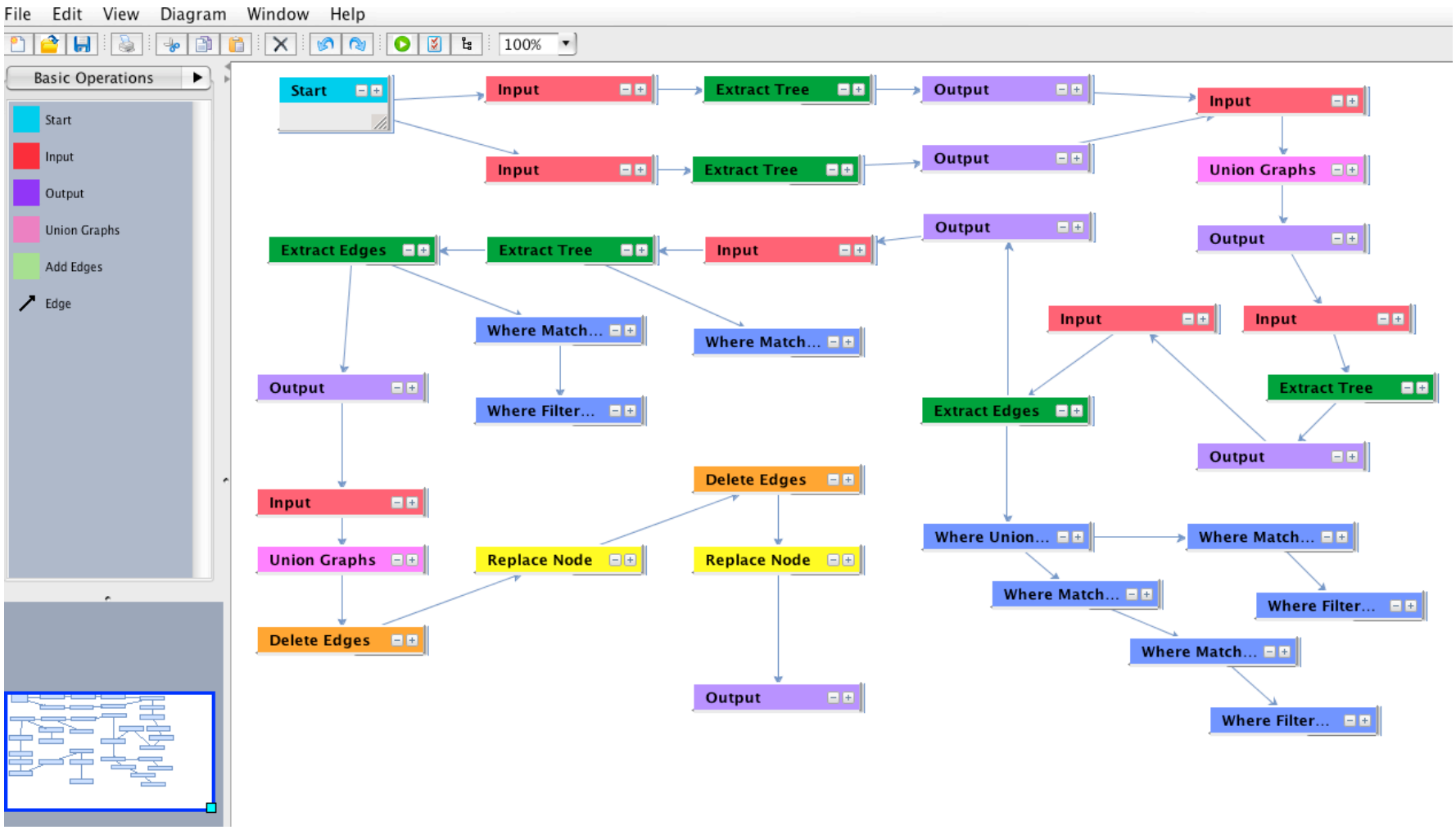
```
PREFIX fma: <http://sig.biostr.washington.edu/fma3.0#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX holder: <http://localhost/holder#>
PREFIX VSparQL: <java:vsparql.ext.#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

INPUT
<http://sig.biostr.washington.edu/fma3.0>{
EXTRACT_TREE { fma:Liver [ forward (fma:regional_part) , forward (fma:constitu
GRAPH <http://sig.biostr.washington.edu/fma3.0>
}
OUTPUT <http://localhost/set1>
```

```
<rdf:RDF
  xmlns:VSparQL="java:vsparql.ext.#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:holder="http://localhost/holder#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:fma="http://sig.biostr.washington.edu/fma3.0#" >
 <rdf:Description rdf:about="http://sig.biostr.washington.edu/fma3.0#Left_p
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.biostr.washington.edu/fma3.0#Liver_s
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.biostr.washington.edu/fma3.0#Portal_
  <fma:constitutional_part rdf:resource="http://sig.biostr.washington.edu/fm
  <fma:constitutional_part rdf:resource="http://sig.biostr.washington.edu/fm
  <fma:constitutional_part rdf:resource="http://sig.biostr.washington.edu/fm
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.biostr.washington.edu/fma3.0#Trunk_
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.biostr.washington.edu/fma3.0#Anteri
  <fma:regional_part rdf:resource="http://sig.biostr.washington.edu/fma3.0#
```

Visualize Results



More Complex Views



Views current work

- Optimization of query execution
- Evaluate on use cases
- Examine other query languages
- View Query Manager

Projects

- Foundational Model of Anatomy
- Ontology Views
- <u>Data management</u>
- Data integration
- Visualization

Data Management

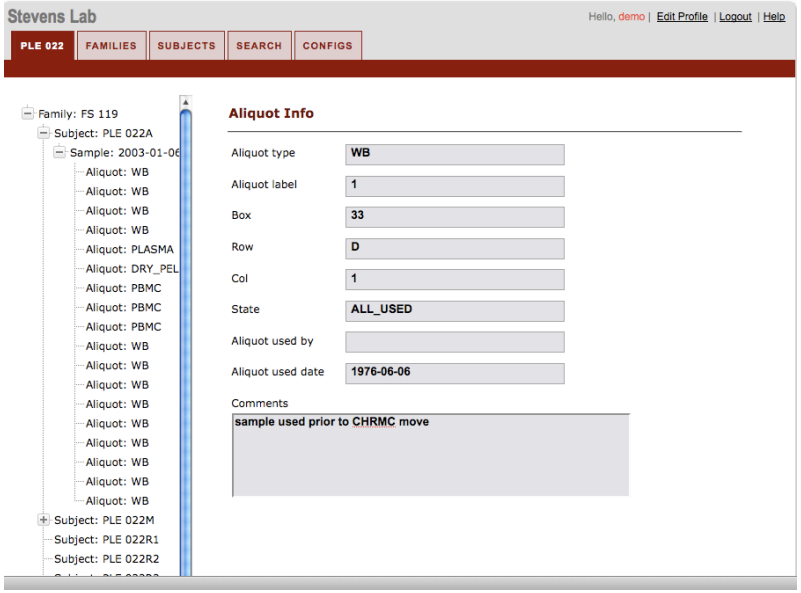
Joshua Franklin, Xenia Hertzenberg, Ron Shaker,
Nolan Nichols

Driving applications

- Institute of Translational Health Sciences (ITHS)
- UW Neuroproteomics Center
- UW Interdisciplinary Brain Imaging Center (IBIC)

ITHS

RedCap Electronic Data Capture



Rlab Freezer Management

RedCap/Rlab “Baby Shower” with Stevens Lab

caTissue at ITHS

caTissue Suite v 1.1

Report Problems | Contact Us | Summary | Logout

caTissue Suite

Home | Administrative Data | Biospecimen Data | Search | Help

Collection Protocol:

CP Prostate Cancer

Participant (Protocol ID): Register New

- Butterworth , Job (N/A)
- Doe , John (N/A)
- Fruit , JuJu (N/A)

Specimen Details for Doe , John (N/A)

- T3.0: CP Prostate Cancer Label: 10-21-
  - 1

Edit Specimen Collection Group | View Surgical Pathology Report | View Annotation | Consents

* Collection Protocol Title	CP Prostate Cancer	* Participant Name (Protocol ID)	Doe, John(N/A)
* Specimen Group Name	CP Prostate Cancer_1_1	Barcode	
* Study Calendar Event Point	3.0,CP Prostate Cancer L Days	* Collection Site	
Offset	0		
* Clinical Diagnosis	Not Specified	* Clinical Status	Not Specified
Medical Record Number	-- Select --	Surgical Pathology Number	
* Activity Status	Active	* Collection Status	Pending
Comments			

Events — Hide Details

Collected		Received	
* Collector	Admin, Admin	* Receiver	Admin, Admin
* Date	10-21-2010 [MM-DD-YYYY]	* Date	10-21-2010 [MM-DD-YYYY]
* Time	18 Hr. 33 Min.	* Time	18 Hr. 33 Min.
* Procedure	Use CP Defaults	* Quality	Use CP Defaults
* Container	Use CP Defaults	Comments	
Comments			

Add Multiple Specimens

Number of Specimen(s) 0

☑ Specimen entry based on collection protocol

☐ Print Labels

Submit | Add Specimen | Add Multiple Specimens | Delete

UW Neuroproteomics Center

Neuroproteomics LIMS

Home | Account | Admin

Welcome Jim Brinkley! Logout Return

New Project | Upload Data | List Uploads

Recent pages: View Project -> Change Information -> Internal Front Page -> View Project -> **Sequest Results**

Sequest Results

Project ID: 5 **Experiment ID:** 38 **Program:** SEQUEST

Min. Scan		Max. Scan		Min. Charge		Max. Charge	
Min. RT		Max. RT		Min. Obs. Mass		Max. Obs. Mass	
Max. XCorr Rank	1			Min. DeltaCN		Min. Sp	
Min. XCorr (+1)		Min. XCorr (+2)		Min. XCorr (+3)		Min. XCorr (>3)	
Peptide				Modified peptides ☑		Unmodified peptides ☑	

Exact: ☑

File(s)

Enter comma-separated file names

Update

 # Results: 12685 **# Results (filtered):**12685

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 >> Last Page 1 of 254

File	Scan	Charge	Obs. Mass	RT	XCr	XCorr	dCN	e-value	Peptide	Protein
2010_Feb_08_cost_plus_34	3395	2	1419.91	1311.3	1	0.94	0.13500	2.9	L.KNIVKVIRPQPK.Q	IPI00375174 (2) IPI00889592
2010_Feb_08_cost_plus_34	7991	3	1820.13	2916.95	1	1.39	0.08900	0.3	R.TLYTDIAIGLLKKIQK.K	IPI00030288 (3) IPI00513682 IPI00642913
2010_Feb_08_cost_plus_34	12430	1	898.61	4488.56	1	0.83	0.32000	0.3	K.SLVLVRVL.V	IPI00377045 (3)
2010_Feb_08_cost_plus_34	3411	2	1756.87	1316.62	1	2.07	0.07700	0.3	K.IEVLQQHENEDIYK.L	IPI00299033 (2)
2010_Feb_08_cost_plus_34	8006	4	1963.27	2921.9	1	1.51	0.28400	0.3	M.KLLGALLALAALLQGAVSLK.I	IPI00031065
2010_Feb_08_cost_plus_34	12449	1	810.56	4495.02	1	0.47	0.40800	3.7	A.PLVIVLR.S	IPI00065537
2010_Feb_08_cost_plus_34	3431	2	2251.14	1323.47	1	3.8	0.12200	0.2	K.VQLHLIESEILDAQEERET.R	IPI00396243 (3)
2010_Feb_08_cost_plus_34	8024	3	2051.33	2928.57	1	0.84	0.05500	0.6	P.ILNVLVVLGVVLLGQFVVR.R	IPI00386229 (5)
2010_Feb_08_cost_plus_34	12471	2	1554.05	4503.19	1	0.77	0.25300	0.1	R.VPLLLPLLLLLALGP.G	IPI00395488
2010_Feb_08_cost_plus_34	3450	2	1829.89	1329.76	1	2.01	0.06100	2.1	L.EICPCNKLKSSDAYK.K	IPI00107625

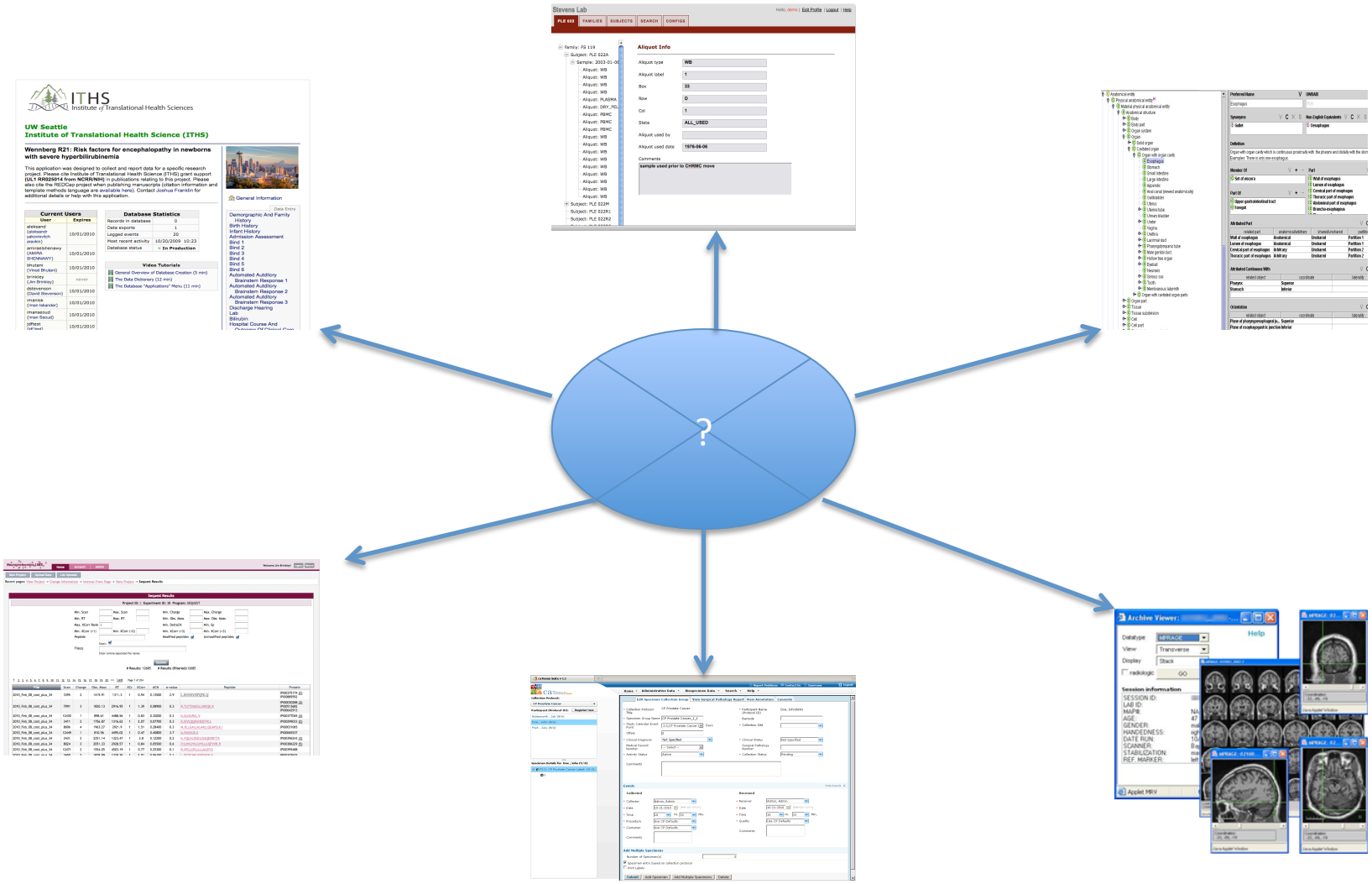
XNAT at IBIC


Archive Viewer:
Help
Datatype
MPRAGE
View
Transverse
Display
Stack
radiologic
GO
Session information
SESSION ID:
LAB ID:
AGE:
47
GENDER:
HANDEDNESS:
DATE RUN:
SCANNER:
STABILIZATION:
REF. MARKER:
Applet MRV
MPRAGE: 02100...
MPRAGE: 02...
MPRAGE: 02...

Data management challenges/future work

- Build versus buy
- Incorporating common terminology for later data integration
- Reducing the time to develop custom applications

How to link together



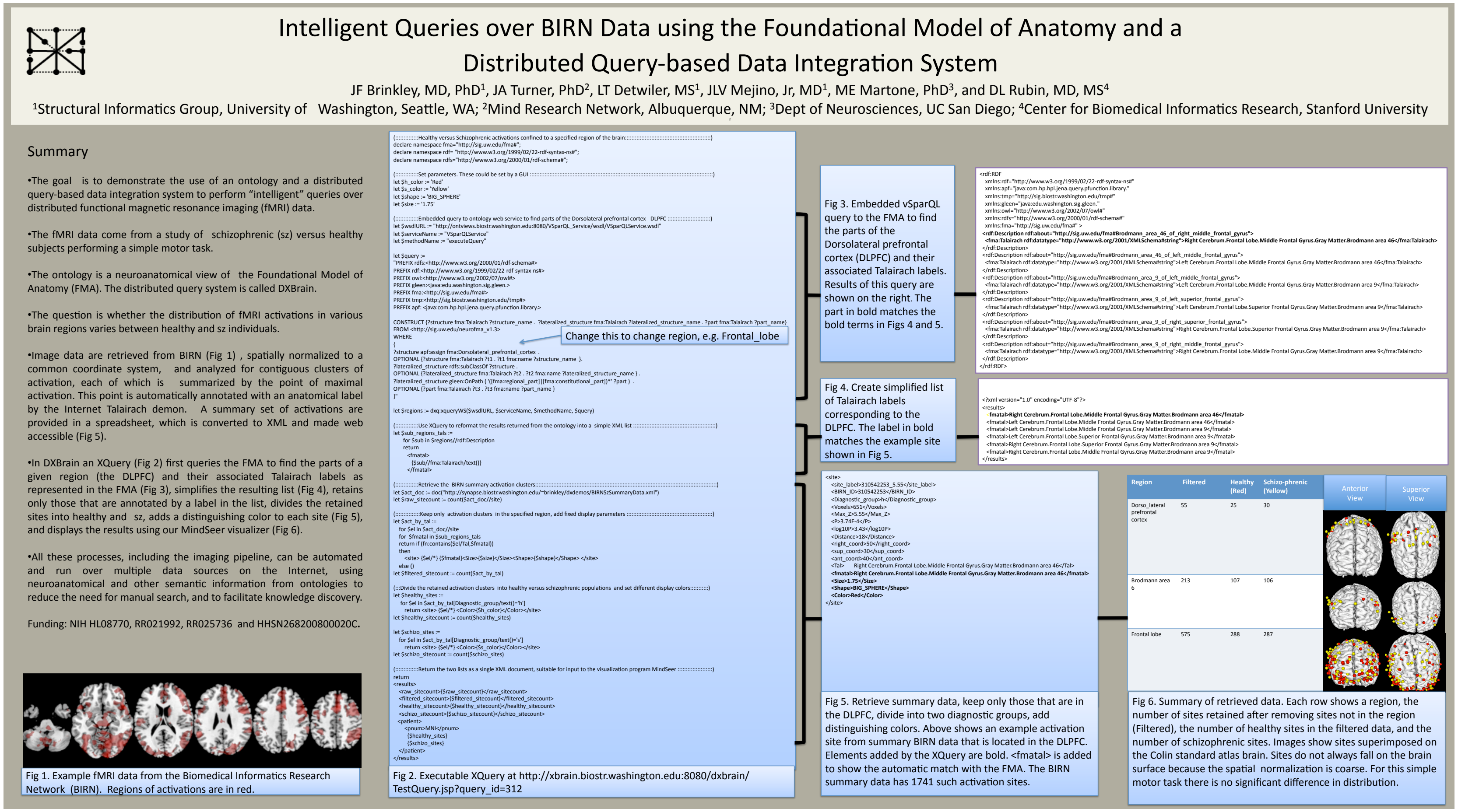
Projects

- Foundational Model of Anatomy
- Ontology Views
- Data management
- <u>Data integration</u>
- Visualization

Data Integration

Todd Detwiler, Joshua Franklin, Ron Shaker

Lightweight Distributed Queries



Intelligent Queries over BIRN Data using the Foundational Model of Anatomy and a Distributed Query-based Data Integration System

JF Brinkley, MD, PhD[1], JA Turner, PhD[2], LT Detwiler, MS[1], JLV Mejino, Jr, MD[1], ME Martone, PhD[3], and DL Rubin, MD, MS[4]
[1]Structural Informatics Group, University of Washington, Seattle, WA; [2]Mind Research Network, Albuquerque, NM; [3]Dept of Neurosciences, UC San Diego; [4]Center for Biomedical Informatics Research, Stanford University

Summary

•The goal is to demonstrate the use of an ontology and a distributed query-based data integration system to perform “intelligent” queries over distributed functional magnetic resonance imaging (fMRI) data.

•The fMRI data come from a study of schizophrenic (sz) versus healthy subjects performing a simple motor task.

•The ontology is a neuroanatomical view of the Foundational Model of Anatomy (FMA). The distributed query system is called DXBrain.

•The question is whether the distribution of fMRI activations in various brain regions varies between healthy and sz individuals.

•Image data are retrieved from BIRN (Fig 1) , spatially normalized to a common coordinate system, and analyzed for contiguous clusters of activation, each of which is summarized by the point of maximal activation. This point is automatically annotated with an anatomical label by the Internet Talairach demon. A summary set of activations are provided in a spreadsheet, which is converted to XML and made web accessible (Fig 5).

•In DXBrain an XQuery (Fig 2) first queries the FMA to find the parts of a given region (the DLPFC) and their associated Talairach labels as represented in the FMA (Fig 3), simplifies the resulting list (Fig 4), retains only those that are annotated by a label in the list, divides the retained sites into healthy and sz, adds a distinguishing color to each site (Fig 5), and displays the results using our MindSeer visualizer (Fig 6).

•All these processes, including the imaging pipeline, can be automated and run over multiple data sources on the Internet, using neuroanatomical and other semantic information from ontologies to reduce the need for manual search, and to facilitate knowledge discovery.

Funding: NIH HL08770, RR021992, RR025736 and HHSN268200800020C.

Fig 1. Example fMRI data from the Biomedical Informatics Research Network (BIRN). Regions of activations are in red.

```
(::::::::::::::Healthy versus Schizophrenic activations confined to a specified region of the brain::::::::::::::::::::::::::::::::::::::::::::::::::)
declare namespace fma="http://sig.uw.edu/fma#";
declare namespace rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#";
declare namespace rdfs="http://www.w3.org/2000/01/rdf-schema#";

(::::::::::::::Set parameters. These could be set by a GUI ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::)
let $h_color := 'Red'
let $s_color := 'Yellow'
let $shape := 'BIG_SPHERE'
let $size := '1.75'

(::::::::::::::Embedded query to ontology web service to find parts of the Dorsolateral prefrontal cortex - DLPFC ::::::::::::::::::::::::::)
let $wsdlURL := "http://ontviews.biostr.washington.edu:8080/VSparQL_Service/wsdl/VSparQLService.wsdl"
let $serviceName := "VSparQLService"
let $methodName := "executeQuery"

let $query :=
"PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX gleen:<java:edu.washington.sig.gleen.>
PREFIX fma:<http://sig.uw.edu/fma#>
PREFIX tmp:<http://sig.biostr.washington.edu/tmp#>
PREFIX apf: <java:com.hp.hpl.jena.query.pfunction.library.>

CONSTRUCT {?structure fma:Talairach ?structure_name .  ?lateralized_structure fma:Talairach ?lateralized_structure_name . ?part fma:Talairach ?part_name}
FROM <http://sig.uw.edu/neurofma_v1.3>
WHERE
{
?structure apf:assign fma:Dorsolateral_prefrontal_cortex .
OPTIONAL {?structure fma:Talairach ?t1 . ?t1 fma:name ?structure_name  }.
?lateralized_structure rdfs:subClassOf ?structure .
OPTIONAL {?lateralized_structure fma:Talairach ?t2 . ?t2 fma:name ?lateralized_structure_name } .
?lateralized_structure gleen:OnPath ( '([fma:regional_part]|[fma:constitutional_part])*' ?part )  .
OPTIONAL {?part fma:Talairach ?t3 . ?t3 fma:name ?part_name }
}"

let $regions := dxq:xqueryWS($wsdlURL, $serviceName, $methodName, $query)

(::::::::::::::Use XQuery to reformat the results returned from the ontology into a  simple XML list ::::::::::::::::::::::::::::::::::::::::::::::::)
let $sub_regions_tals :=
      for $sub in $regions//rdf:Description
      return
        <fmatal>
          {$sub//fma:Talairach/text()}
        </fmatal>

(::::::::::::::Retrieve the  BIRN summary activation clusters::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::)
let $act_doc := doc("http://synapse.biostr.washington.edu/~brinkley/dxdemos/BIRNSzSummaryData.xml")
let $raw_sitecount := count($act_doc//site)

(::::::::::::::Keep only  activation clusters  in the specified region, add fixed display parameters ::::::::::::::::::::::::::::::::::::::::::::::::)
let $act_by_tal :=
   for $el in $act_doc//site
   for  $fmatal in $sub_regions_tals
   return if (fn:contains($el/Tal,$fmatal))
   then
      <site> {$el/*} {$fmatal}<Size>{$size}</Size><Shape>{$shape}</Shape> </site>
   else ()
let $filtered_sitecount := count($act_by_tal)

(:::Divide the retained activation clusters  into healthy versus schizophrenic populations  and set different display colors::::::::::)
let $healthy_sites :=
    for $el in $act_by_tal[Diagnostic_group/text()='h']
      return <site> {$el/*} <Color>{$h_color}</Color></site>
let $healthy_sitecount := count($healthy_sites)

let $schizo_sites :=
   for $el in $act_by_tal[Diagnostic_group/text()='s']
      return <site> {$el/*} <Color>{$s_color}</Color></site>
let $schizo_sitecount := count($schizo_sites)

(::::::::::::::Return the two lists as a single XML document, suitable for input to the visualization program MindSeer ::::::::::::::::::::)
return
<results>
   <raw_sitecount>{$raw_sitecount}</raw_sitecount>
   <filtered_sitecount>{$filtered_sitecount}</filtered_sitecount>
   <healthy_sitecount>{$healthy_sitecount}</healthy_sitecount>
   <schizo_sitecount>{$schizo_sitecount}</schizo_sitecount>
  <patient>
      <pnum>MNI</pnum>
        {$healthy_sites}
        {$schizo_sites}
   </patient>
</results>
```

Change this to change region, e.g. Frontal_lobe

Fig 2. Executable XQuery at http://xbrain.biostr.washington.edu:8080/dxbrain/TestQuery.jsp?query_id=312

Fig 3. Embedded vSparQL query to the FMA to find the parts of the Dorsolateral prefrontal cortex (DLPFC) and their associated Talairach labels. Results of this query are shown on the right. The part in bold matches the bold terms in Figs 4 and 5.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:apf="java:com.hp.hpl.jena.query.pfunction.library."
  xmlns:tmp="http://sig.biostr.washington.edu/tmp#"
  xmlns:gleen="java:edu.washington.sig.gleen."
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:fma="http://sig.uw.edu/fma#" >
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_46_of_right_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_46_of_left_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_left_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_left_superior_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_right_superior_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_right_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
</rdf:RDF>
```

Fig 4. Create simplified list of Talairach labels corresponding to the DLPFC. The label in bold matches the example site shown in Fig 5.

```
<?xml version="1.0" encoding="UTF-8"?>
<results>
  <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Right Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
</results>
```

```
<site>
   <site_label>310542253_5.55</site_label>
   <BIRN_ID>310542253</BIRN_ID>
   <Diagnostic_group>h</Diagnostic_group>
   <Voxels>651</Voxels>
   <Max_Z>5.55</Max_Z>
   <P>3.74E-4</P>
   <log10P>3.43</log10P>
   <Distance>18</Distance>
   <right_coord>50</right_coord>
   <sup_coord>30</sup_coord>
   <ant_coord>40</ant_coord>
   <Tal>      Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</Tal>
   <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
   <Size>1.75</Size>
   <Shape>BIG_SPHERE</Shape>
   <Color>Red</Color>
</site>
```

Fig 5. Retrieve summary data, keep only those that are in the DLPFC, divide into two diagnostic groups, add distinguishing colors. Above shows an example activation site from summary BIRN data that is located in the DLPFC. Elements added by the XQuery are bold. <fmatal> is added to show the automatic match with the FMA. The BIRN summary data has 1741 such activation sites.

Region	Filtered	Healthy (Red)	Schizo-phrenic (Yellow)	Anterior View	Superior View
Dorso_lateral prefrontal cortex	55	25	30		
Brodmann area 6	213	107	106		
Frontal lobe	575	288	287		

Fig 6. Summary of retrieved data. Each row shows a region, the number of sites retained after removing sites not in the region (Filtered), the number of healthy sites in the filtered data, and the number of schizophrenic sites. Images show sites superimposed on the Colin standard atlas brain. Sites do not always fall on the brain surface because the spatial normalization is coarse. For this simple motor task there is no significant difference in distribution.

View Query Manager

Todd Detwiler

A tool for creating, managing and querying views over the web

- Many web representations
  - XML, RDF, OWL
- Many query languages and Interfaces
  - SparQL, vSparQL, XQuery, DL-Query …
- Common foundation is XML
- Allow user to choose the sources and query language

Distributed XQuery

Structural Informatics Group

Query Manager

Hello Jim Brinkley **Logout**

Manage **Search** **Edit** **User** **Help**

Title

AMIA 2010: BIRN data: Healthy vs schizophrenic activations confined to a specified region of the brain

Description

Determine parts of a given structure. Retrieve summary BIRN fmri activations, retain only those that have been annotated by the Talairach demon (TD) to be in the given structure. Divide into healthy vs schizophrenic, and set color of display spheres to be different for each type of population. Also shows automatic matching of matching of TD labels with FMA labels. This is

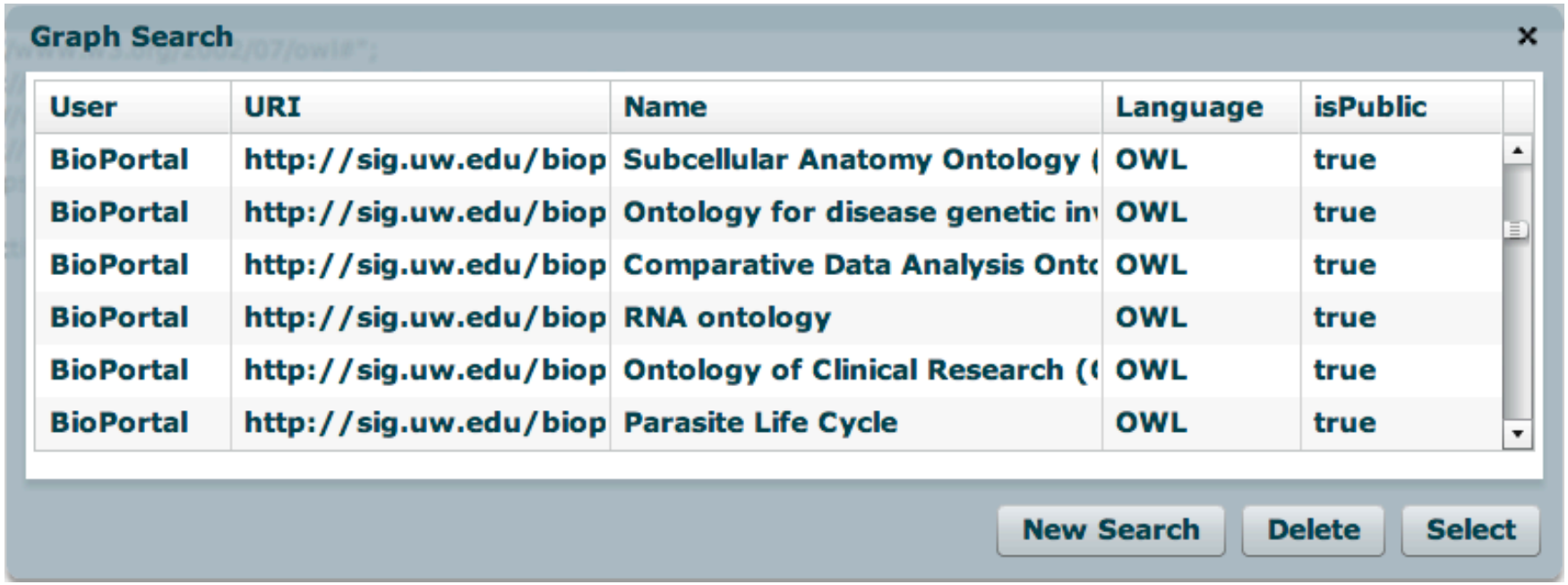
Query

```
(::::::::::::::Healthy versus Schizophrenic activations confined to a specified region of the brain::::::::::::::::::::::::::::::::::::::::::::::::::::::)
declare namespace fma="http://sig.uw.edu/fma#";
declare namespace rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#";
declare namespace rdfs="http://www.w3.org/2000/01/rdf-schema#";

(::::::::::::::Set parameters. These could be set by a GUI ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::)
let $h_color := 'Red'
let $s_color := 'Yellow'
let $shape := 'BIG_SPHERE'
let $size := '1.75'

(::::::::::::::Embedded query to ontology web service to find parts of the Dorsolateral prefrontal cortex - DLPFC ::::::::::::::::::::::::::::::)
let $wsdlURL := "http://ontviews.biostr.washington.edu:8080/VSparQL_Service/wsdl/VSparQLService.wsdl"
let $serviceName := "VSparQLService"
let $methodName := "executeQuery"

let $query :=
"PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX gleen:<java:edu.washington.sig.gleen.>
PREFIX fma:<http://sig.uw.edu/fma#>
PREFIX tmp:<http://sig.biostr.washington.edu/tmp#>
PREFIX apf: <java:com.hp.hpl.jena.query.pfunction.library.>
```

Language **DXQuery**

Execute

Search over any Bioportal Ontology



Graph Search

User	URI	Name	Language	isPublic
BioPortal	http://sig.uw.edu/biop	Subcellular Anatomy Ontology (	OWL	true
BioPortal	http://sig.uw.edu/biop	Ontology for disease genetic in	OWL	true
BioPortal	http://sig.uw.edu/biop	Comparative Data Analysis Onto	OWL	true
BioPortal	http://sig.uw.edu/biop	RNA ontology	OWL	true
BioPortal	http://sig.uw.edu/biop	Ontology of Clinical Research (	OWL	true
BioPortal	http://sig.uw.edu/biop	Parasite Life Cycle	OWL	true

New Search Delete Select

Value sets over OCRe

Structural Informatics Group

Query Manager

Hello Jim Brinkley **Logout**

Manage **Search** **Edit** **User** **Help**

Title

OCRe telecommunication value set over Bioportal version of OCRe

Description

A simple of example of a query generated value set from OCRe, in this case values are chosen as all instances of http://purl.org/net/OCRe/research.owl#Telecommunication_scheme

Query

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX research: <http://purl.org/net/OCRe/research.owl#>
PREFIX valset:<http://sig.uw.edu/valueset#>

CONSTRUCT {valset:set valset:hasMember ?x.}
FROM <http://rest.bioontology.org/bioportal/ontologies/download/44678>
WHERE
{
  ?x rdf:type/rdfs:subClassOf* research:Telecommunication_scheme
}
```

Language **VSparQL**

Execute

Value Set Service

http://xiphoid:8080/QueryService/ParameterizedQueryResults?qid=62

Most Visited Getting Started Latest Headlines

http://xiphoid....ryManager.html# | DXBrain Project - Query | http://xiphoid:...yResults?qid=62


```
<rdf:RDF
    xmlns:OCRe-Start-Here="http://purl.org/net/OCRe/OCRe-Start-Here#"
    xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
    xmlns:research="http://purl.org/net/OCRe/research.owl#"
    xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
    xmlns:owlapi="http://www.semanticweb.org/owlapi#"
    xmlns:study_protocol="http://purl.org/net/OCRe/study_protocol#"
    xmlns:clinical2="http://purl.org/net/OCRe/clinical.owl#"
    xmlns:valset="http://sig.uw.edu/valueset#"
    xmlns:statistics="http://purl.org/net/OCRe/statistics.owl#"
    xmlns:owl2xml="http://www.w3.org/2006/12/owl2-xml#"
    xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
    xmlns:fma3="http://sig.biostr.washington.edu/fma3.0#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:study_design="http://purl.org/net/OCRe/study_design.owl#"
    xmlns:study_exposure="http://www.hsdb.org/StudyExposure.owl#"
    xmlns:study_protocol2="http://purl.org/net/OCRe/study_protocol.owl#"
    xmlns:clinical="http://purl.org/net/OCRe/clinical#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns:swrl="http://www.w3.org/2003/11/swrl#"
    xmlns:ontologies="http://purl.org/net/OCRe/"
    xmlns:owl2="http://www.w3.org/2006/12/owl2#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
  <rdf:Description rdf:about="http://sig.uw.edu/valueset#set">
    <valset:hasMember rdf:resource="http://purl.org/net/OCRe/OCRe-Start-Here#tel"/>
    <valset:hasMember rdf:resource="http://purl.org/net/OCRe/OCRe-Start-Here#mailto"/>
    <valset:hasMember rdf:resource="http://purl.org/net/OCRe/OCRe-Start-Here#x-text-fax"/>
    <valset:hasMember rdf:resource="http://purl.org/net/OCRe/OCRe-Start-Here#http"/>
  </rdf:Description>
</rdf:RDF>
```

View Query Manager Further Work

- Additional query languages
- Graphical user interfaces
- Results visualization
- Optimize execution time
- Views over views
- Deliver to NCBO

Projects

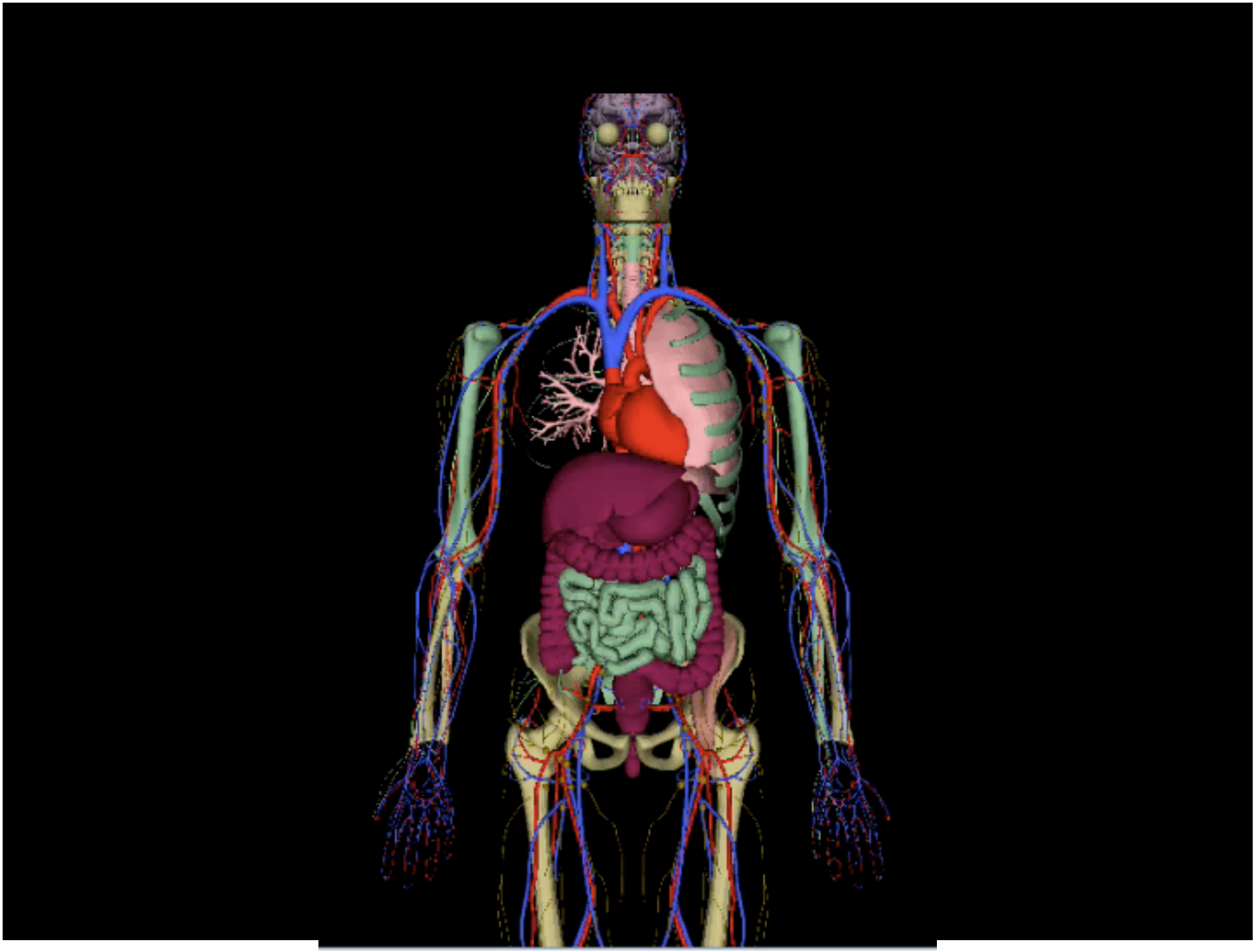
- Foundational Model of Anatomy
- Ontology Views
- Data management
- Data integration
- <u>Visualization</u>

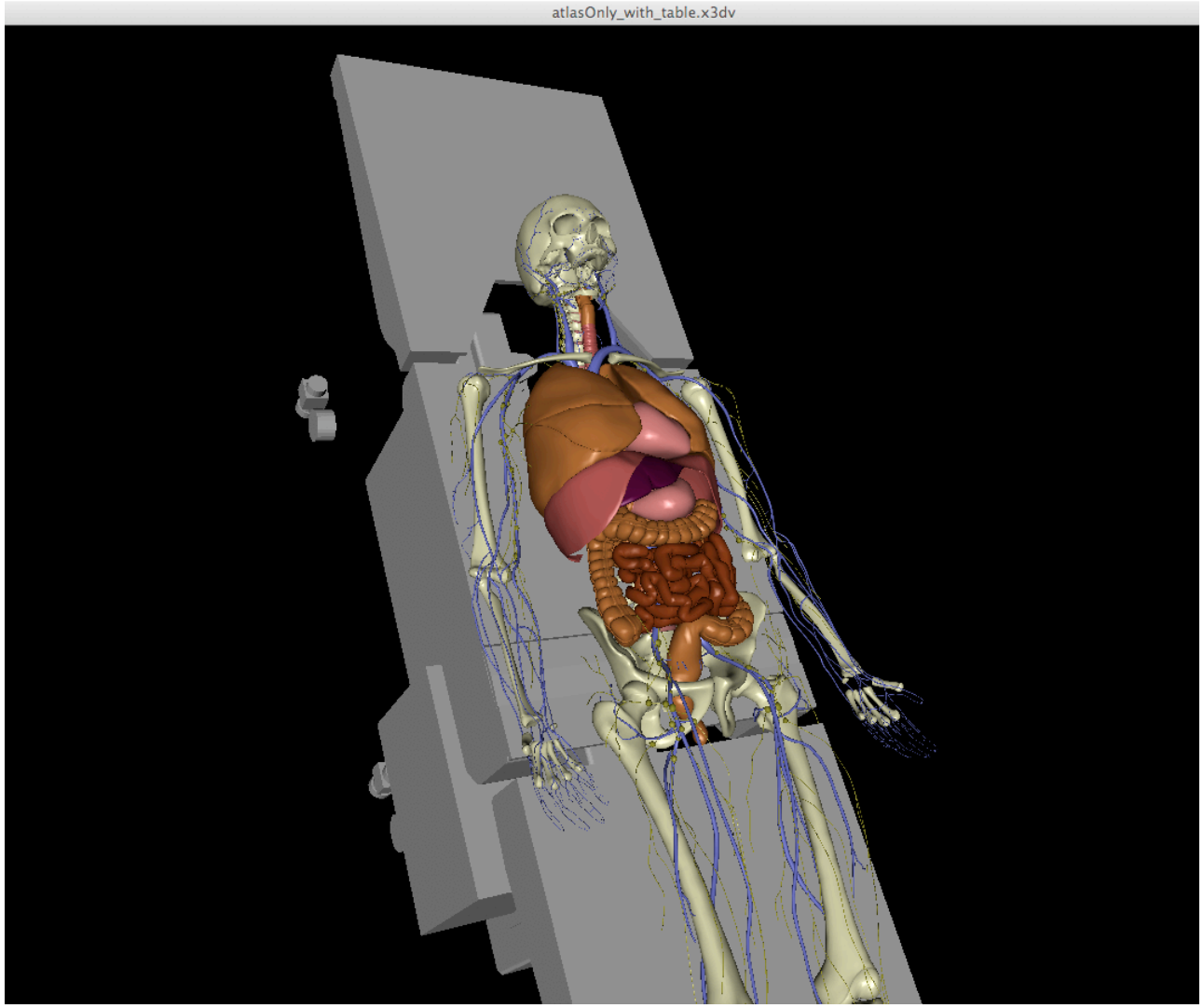
Visualization

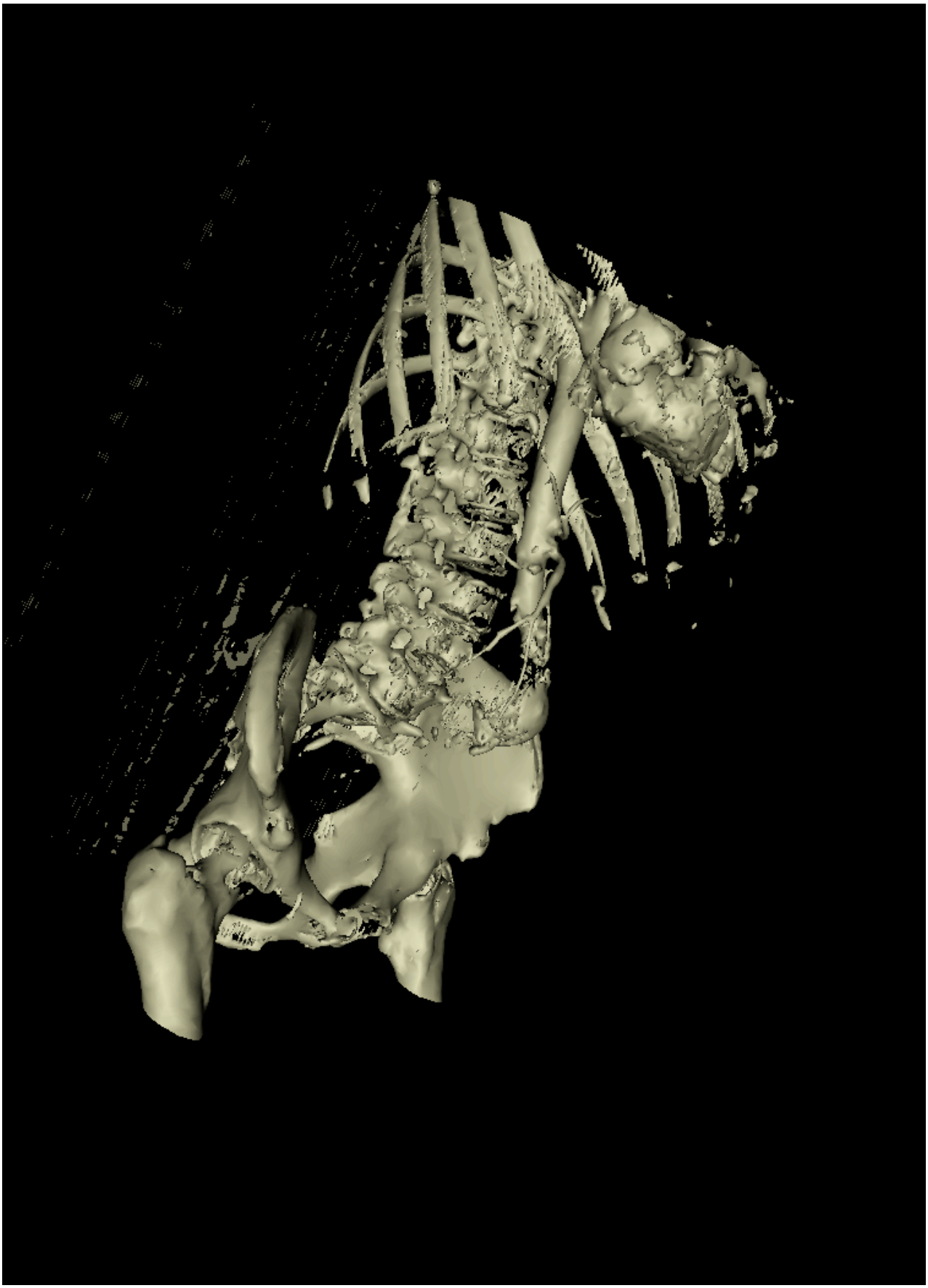
Wayne Warren and Trond Nilsen

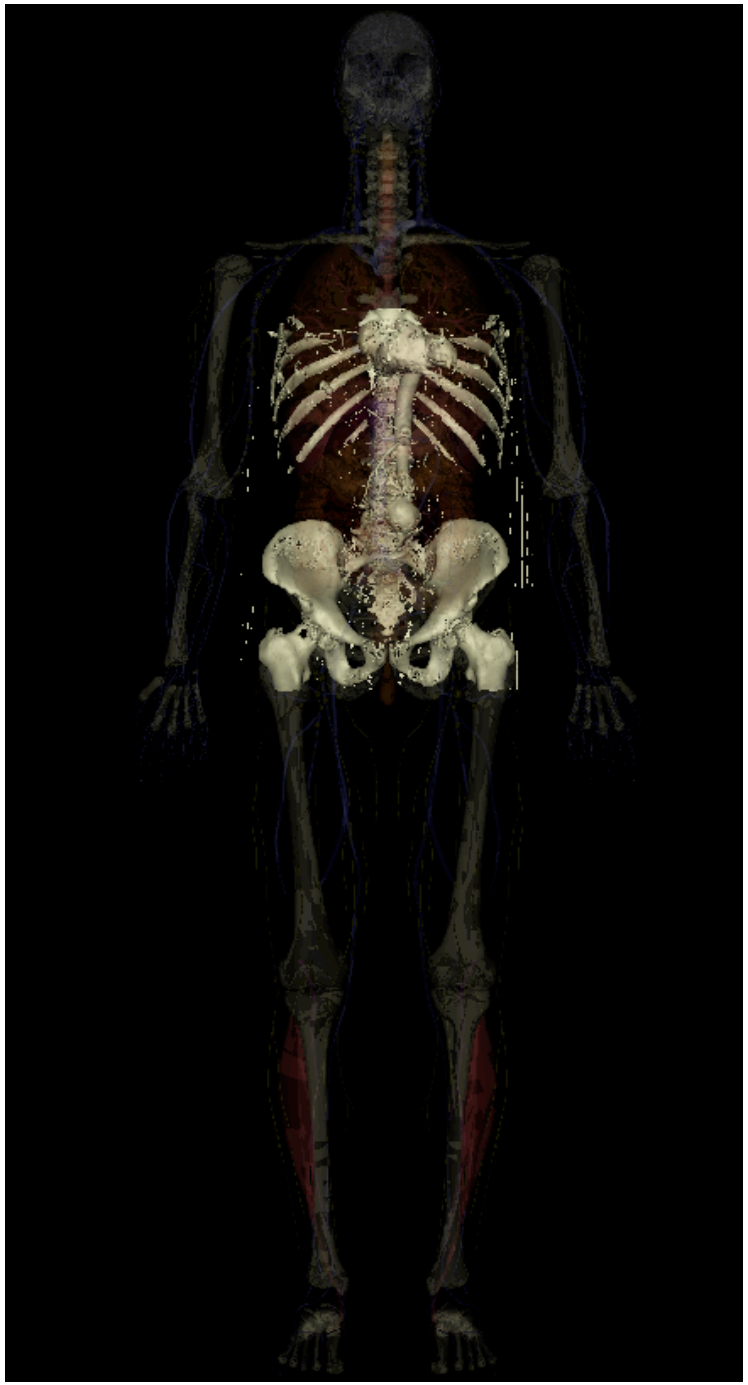
The Intelligent Virtual Cadaver

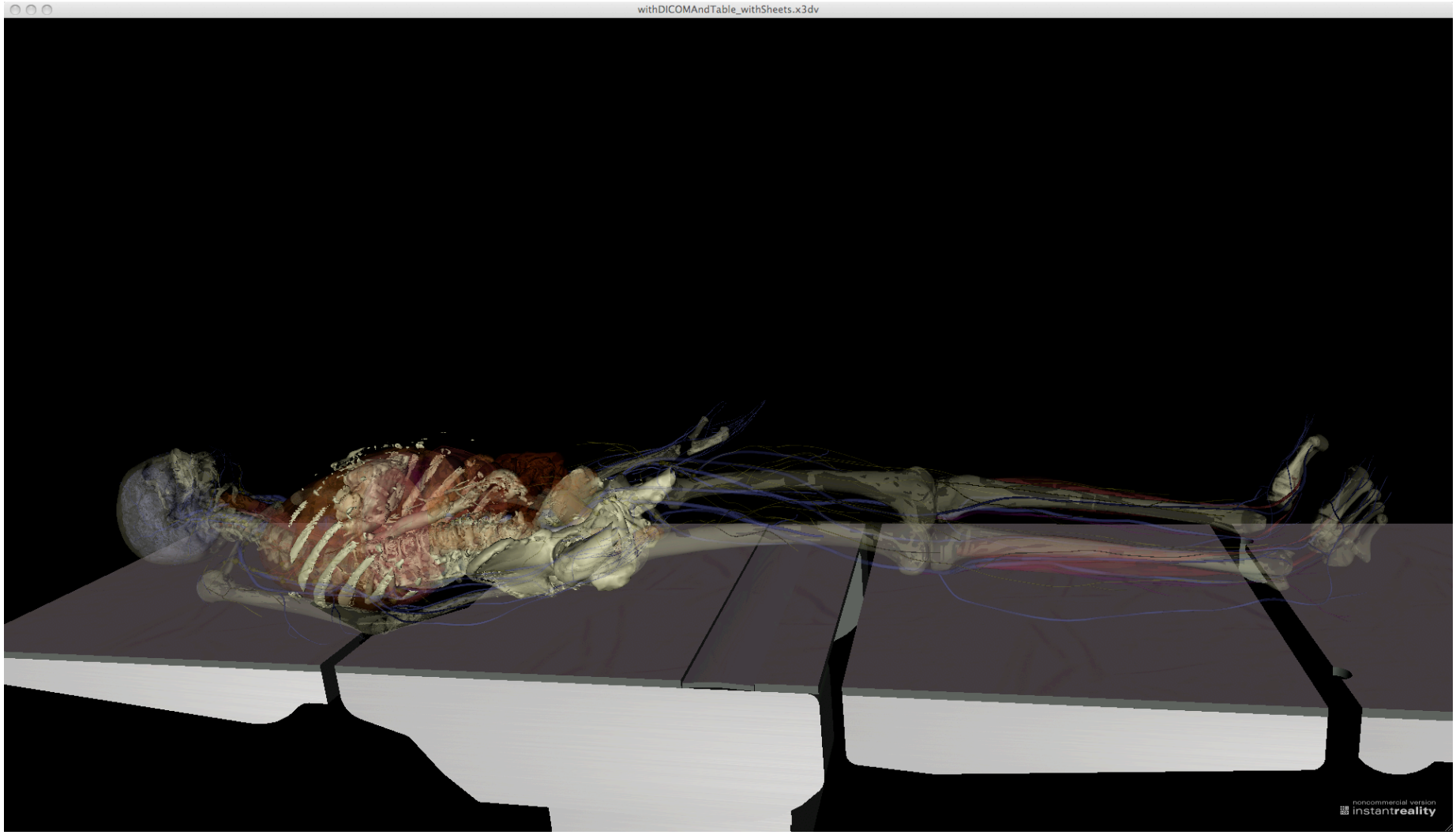
- Collaboration with ISIS
- Uses Biolucida

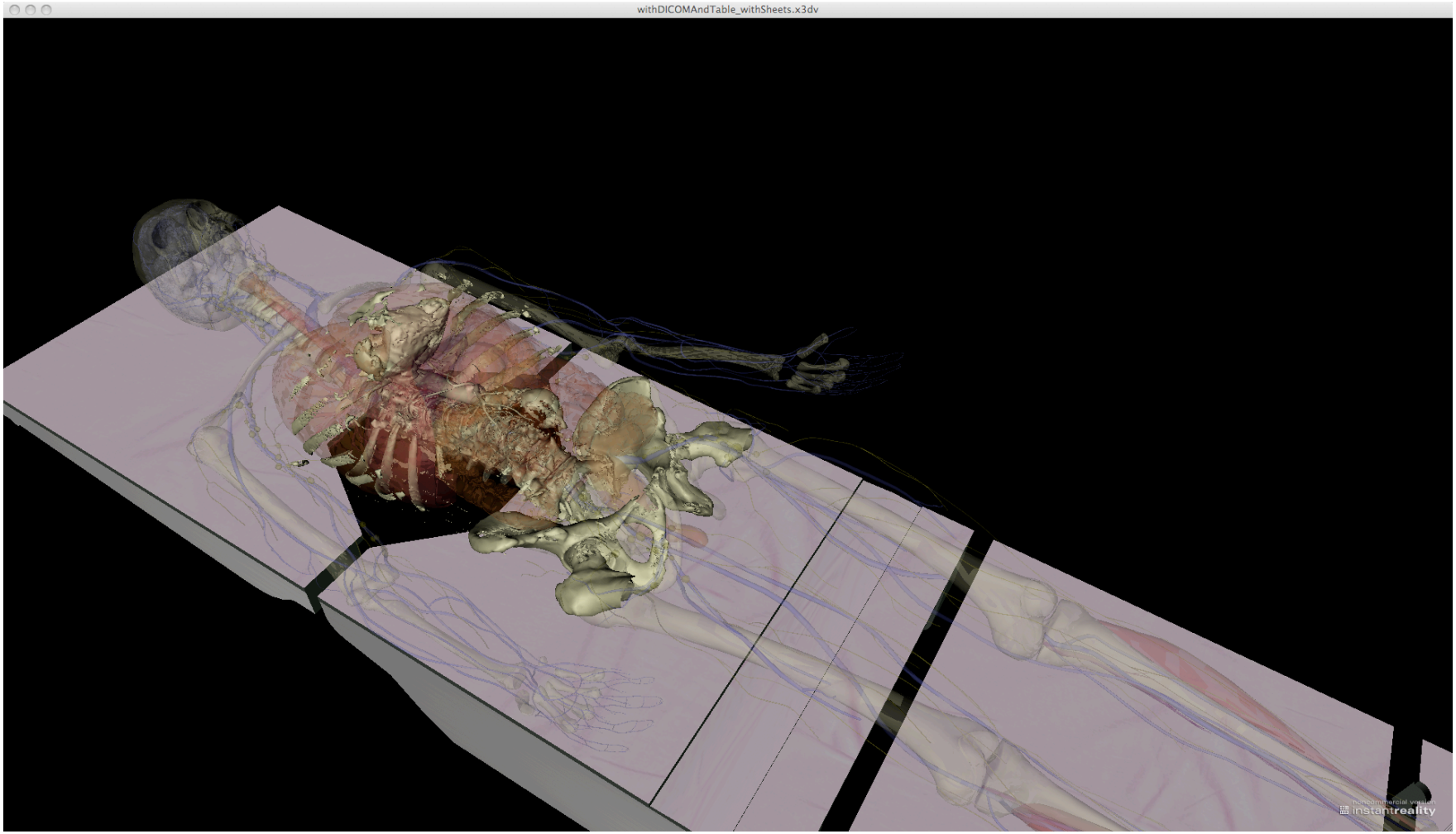



atlasOnly_with_table.x3dv






withDICOMAndTable_withSheets.x3dv
noncommercial version
instantreality


withDICOMAndTable_withSheets.x3dv
instantreality

Projects

- Foundational Model of Anatomy
- Ontology Views
- Data management
- Data integration
- Visualization

Foundational Model of Anatomy: Challenges

- Representing development
- Conversion to OWL 2
- Modularity
- How to allow multiple contributors while also maintaining principles

Ontology Views: Challenges

- Incorporating semantics
- User interface and results visualization
- Response time

Data Management: Challenges

- How to create a custom data management tool for little cost
- How to utilize standards and ontologies from the outset

Data integration: Challenges

- How to make it lightweight
- How to ensure interoperability
- Privacy and security
- Scalability and response time
- Integration with larger efforts
- Usability

Visualization: Challenges

- Integration with other tools
- Content creation
- Usability

Funding

- Ontology views
- RadLex
- OcRE
- UW Neuroproteomics Center
- ITHS
- ISIS
- caBIG
- NLM fellowships
- Facebase
- IBIC

Credits

- Cornelius Rosse
- Onard Mejino
- Todd Detwiler
- Joshua Franklin
- Ron Shaker
- Xenia Hertzenberg
- Nicola Dell
- Nolan Nicols

Marianne Shaw

Dan Suciu

Linda Shapiro

Credits

- Dan Cook
- John Gennari
- Mark Musen
- Daniel Rubin
- Natasha Noy
- Jessica Turner
- Max Neal
- Melissa Clarkson
- Peter Tarczy-Hornoch

Towards a Structural Information Framework