Application of the PORTAL-DOORS System for Use by Clinical Trial Registries

Carl Taswell, MD, PhD
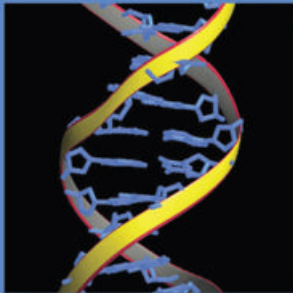
Presented 15 March 2009 at

AMIA Summit on Translational Bioinformatics

San Francisco, CA

Talk in 3 Slides

- Brain Imaging as Motivating Context
- The Blueprint for PORTAL-DOORS
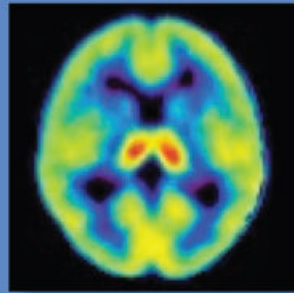- The Essence of PORTAL-DOORS

2007 SNM Image of the Year

Wagner p15N, Journal of Nuclear Medicine Vol 48 No 7 August 2007

The Blueprint for PORTAL-DOORS

DOORS to the Semantic Web and Grid With a PORTAL for Biomedical Computing

Carl Taswell, *Member, IEEE*

Abstract—The semantic web remains in the early stages of development. It has not yet achieved the goals envisioned by its founders as a pervasive web of distributed knowledge and intelligence. Success will be attained when a dynamic synergism can be created between people and a sufficient number of infrastructure systems and tools for the semantic web in analogy with those for the original web. The domain name system (DNS), web browsers, and the benefits of publishing web pages motivated many people to register domain names and publish web sites on the original web. An analogous resource label system, semantic search applications, and the benefits of collaborative semantic networks will motivate people to register resource labels and publish resource descriptions on the semantic web. The Domain Ontology Oriented Resource System (DOORS) and Problem Oriented Registry of Tags and Labels (PORTAL) are proposed as infrastructure systems for

registries are proposed with scientific problem-oriented designs that avoid the engineering-technology-oriented restrictions of existing registries.

Sections II–IV review the background and motivation for DOORS, PORTAL, and BioPORT. Section II explains key concepts of the current semantic web and grid, and summarizes how they are driving the transformation of software architecture from designs based on closed-world computing to those based on open-world computing. Section III reviews the literature and current state-of-the-art in the life sciences web and grid, and summarizes the opinions of leading commentators in the bioinformatics community on existing barriers that impede development. Section IV defines the meaning and scope of biomedical

Manuscript received 10/31/2006; online 8/3/2007; in print 3/5/2008

The Essence of PORTAL-DOORS

- PORTAL-DOORS for the semantic web modeled on the success of IRIS-DNS for the original web
- *PORTAL-DOORS designed to address three major problems: cybersilos, search engine consolidation, and transition barriers*
- Significant benefits for translational bioinformatics:
  - Distributed registry-directory system
  - Important applications for clinical trials
  - Important applications for complex information systems such as those necessary for the study of gene-brain-behavior relationships

Outline of Talk in 3 Sections

- Brain Imaging and Informatics

- PORTAL-DOORS for the Semantic Web

- A Vision for the Future

  - PORTAL-DOORS Applications in Brain Imaging
 - PORTAL-DOORS Applications in Clinical Trials

Brain Imaging and Informatics
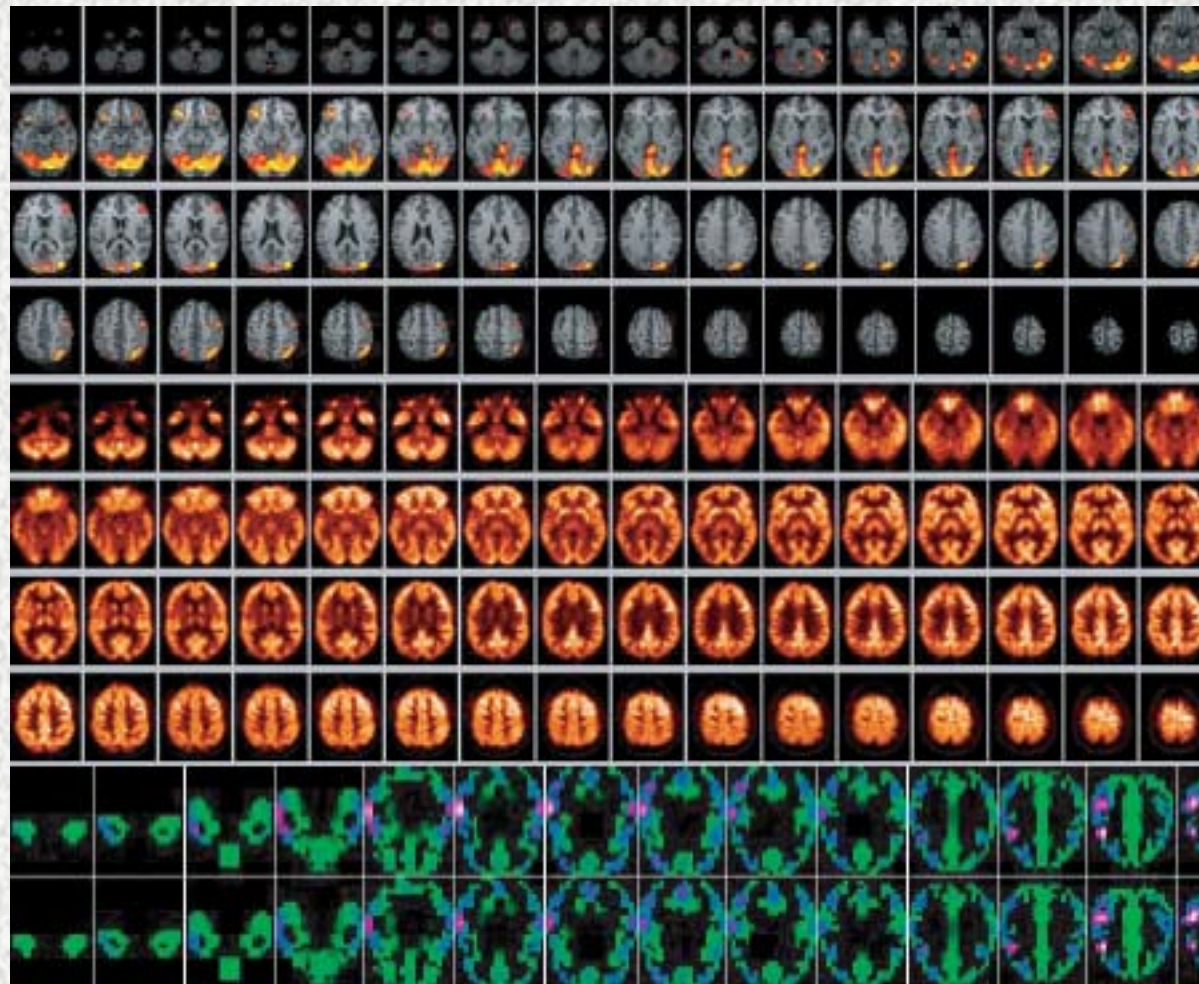
How to Visualize the Brain?

- Ex vivo anatomic and histopathologic methods require cadavers, specimens or biopsies

- In vivo physiologic methods (EEG, EMG, SSEP, VEP, AEP) remain limited

- Radiologic methods (CT, MR) provide excellent structural information

- *Nuclear medicine methods (PET, SPECT) provide excellent functional information*

- Future PET-MR? Benefits of both PET and MR?

Brain Imaging for Alzheimer's Disease
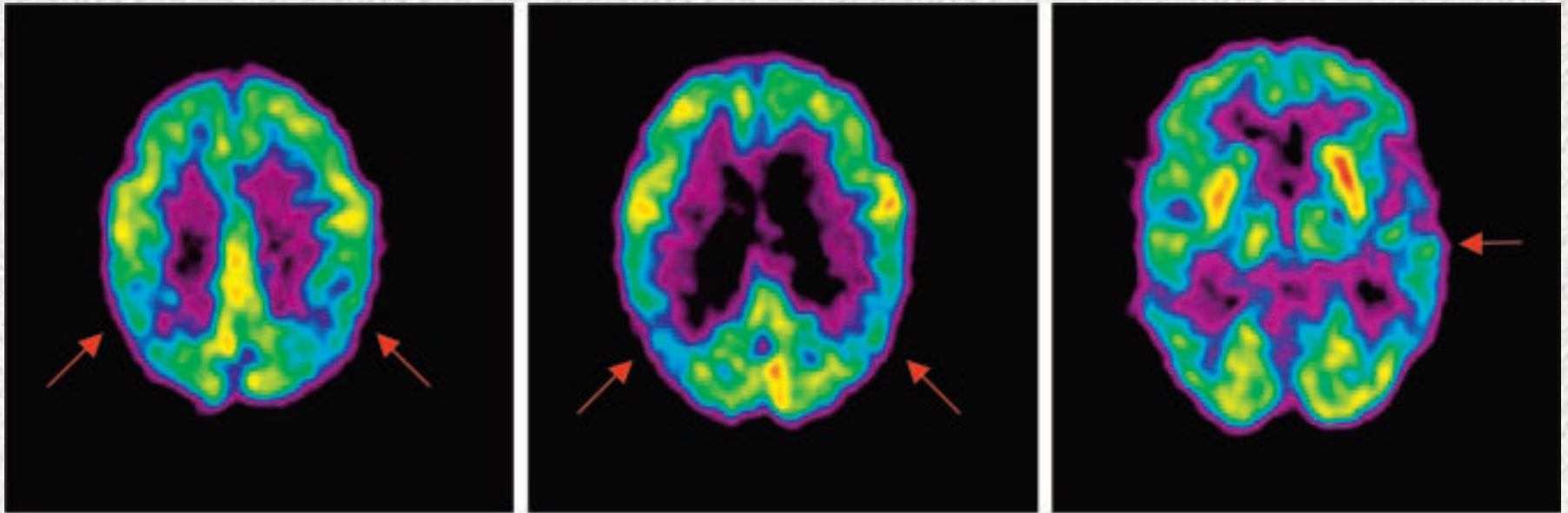
18F-FDG PET images of early Alzheimer's disease.

DIAGNOSIS OF NEURODEGENERATIVE DEMENTIAS • Silverman
JOURNAL OF NUCLEAR MEDICINE Vol. 45 No. 4 April 2004
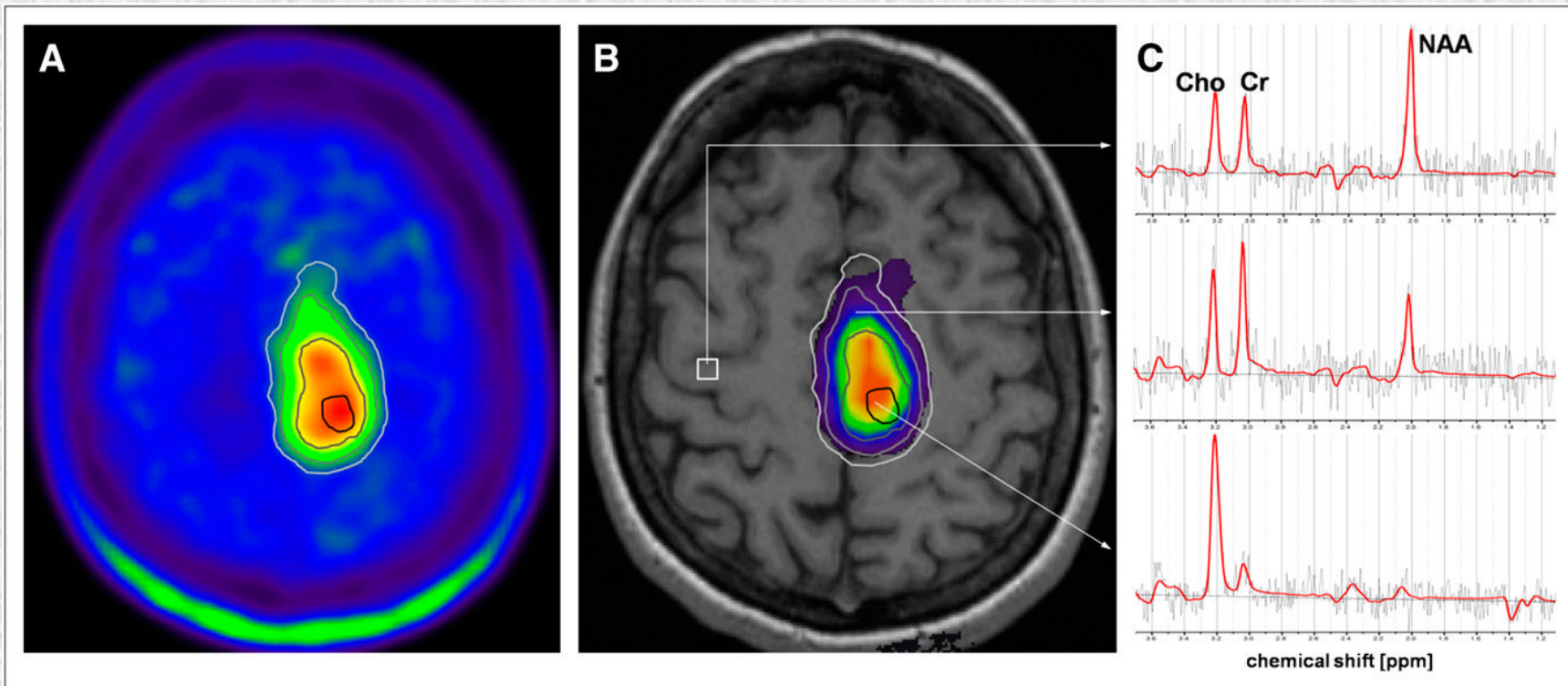
18F-FDG PET for Various Dementias
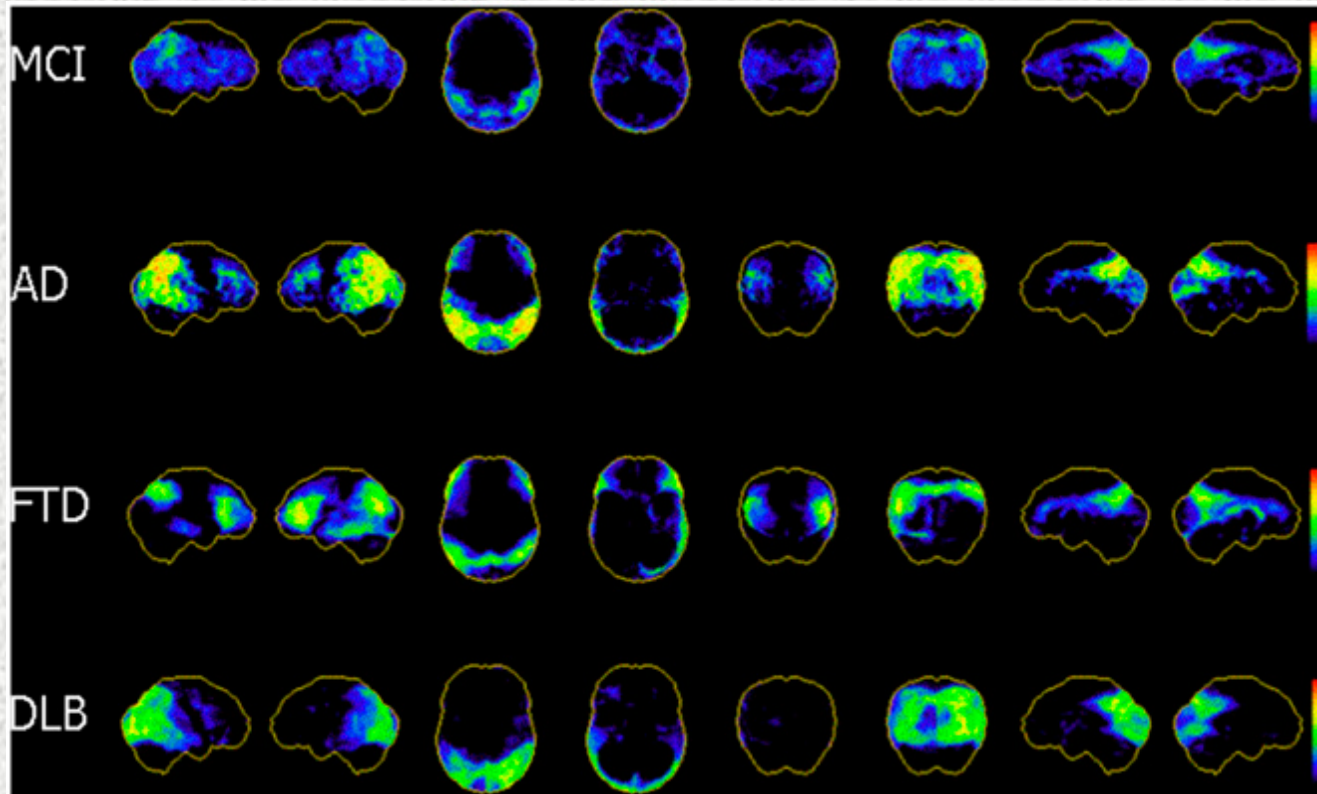
18F-FDG PET DIAGNOSIS OF EARLY DEMENTIA • Mosconi et al.
JOURNAL OF NUCLEAR MEDICINE Vol. 49 No. 3 March 2008
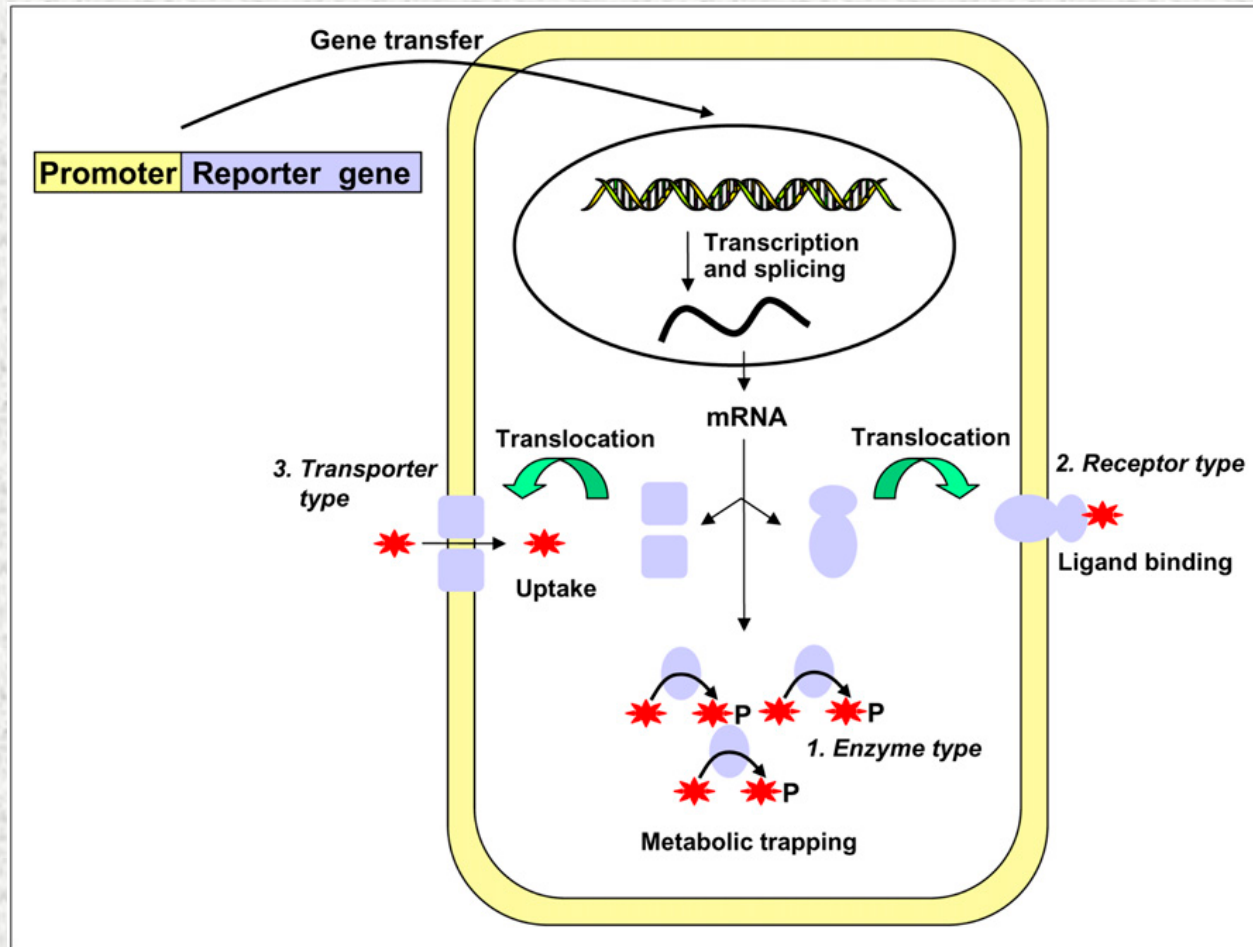
Bi-modal Imaging (PET+MRI) for Parkinson's Disease
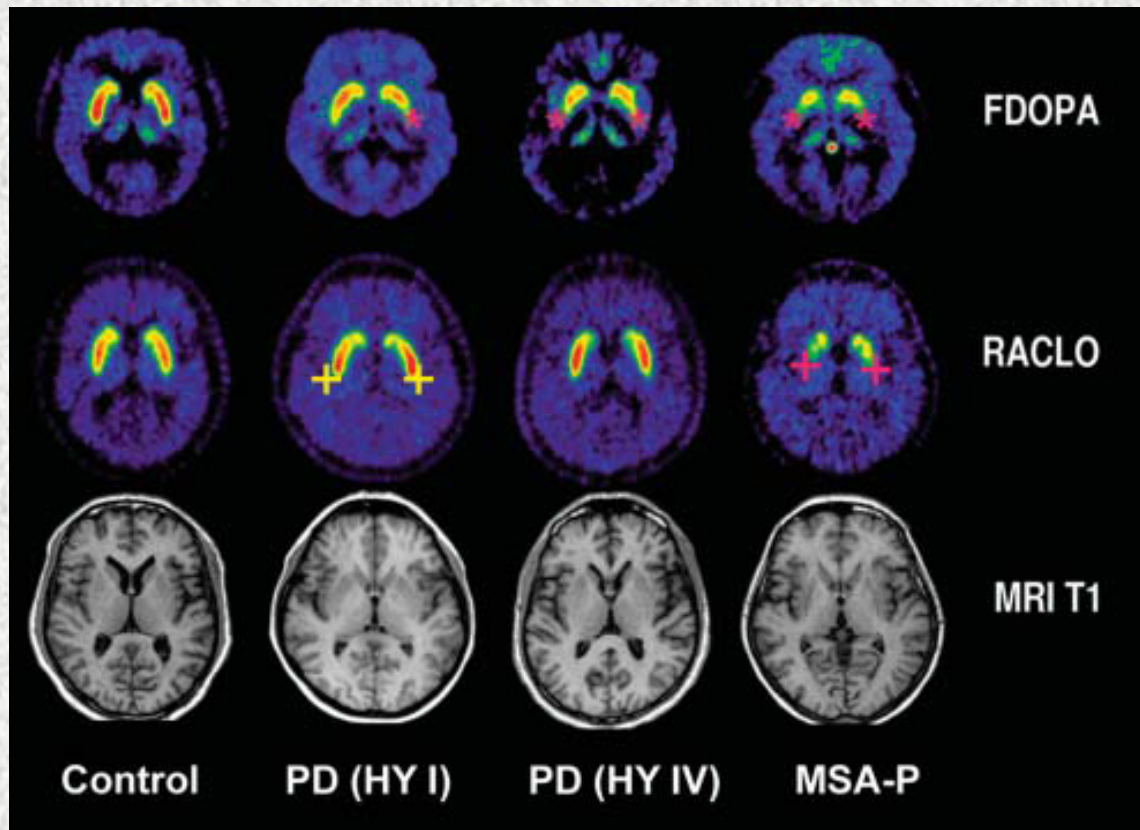
BRAIN RECEPTOR IMAGING • Heiss and Herholz
JOURNAL OF NUCLEAR MEDICINE Vol. 47 No. 2 February 2006

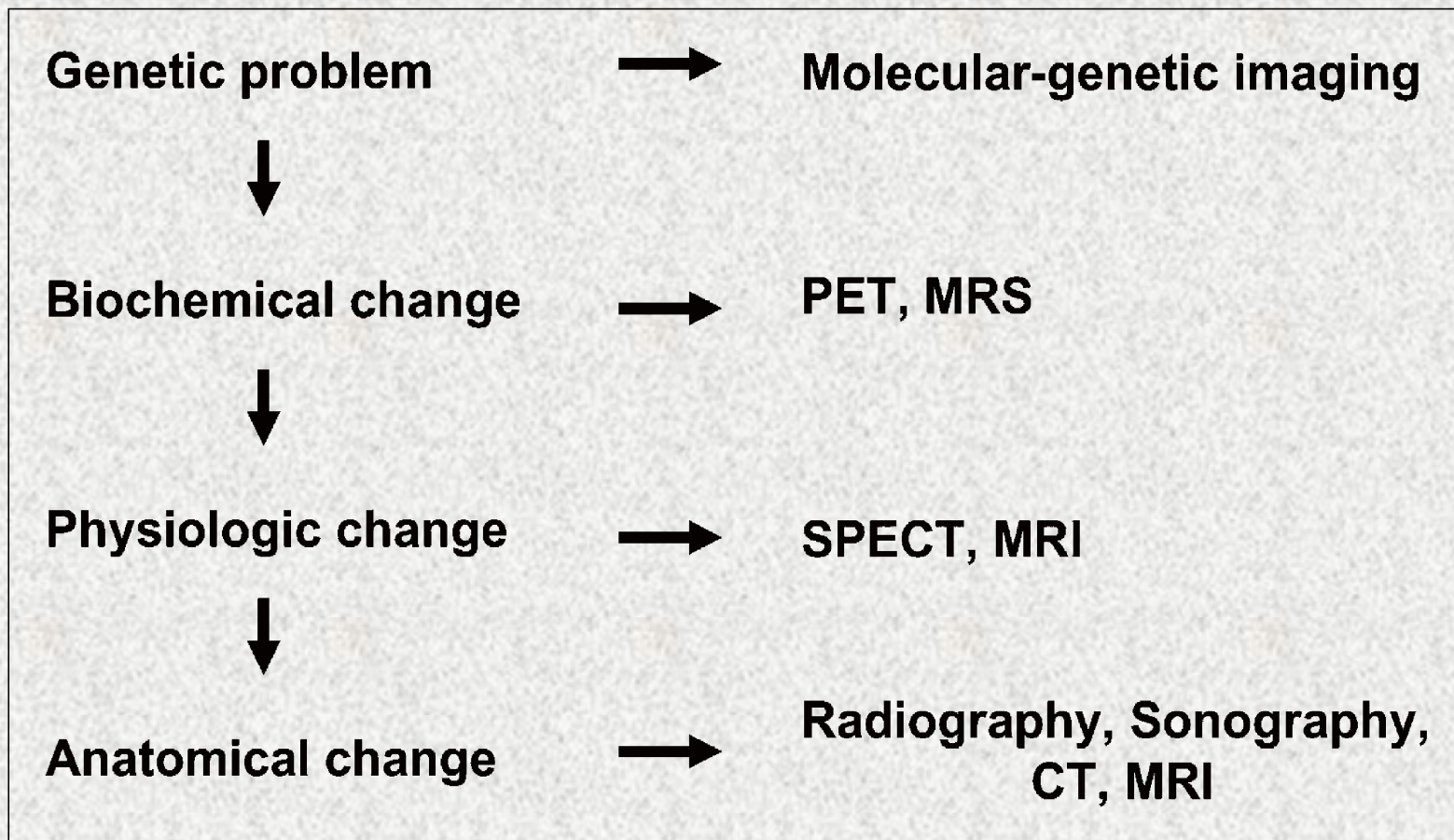# Tri-modal Imaging (PET+MRI+MRS) for Glioblastoma multiforme

METABOLIC IMAGING OF HUMAN GLIOMAS • Stadlbauer et al.
JOURNAL OF NUCLEAR MEDICINE Vol. 49 No. 5 May 2008

PET/SPECT Molecular Genetic Imaging

IMAGING OF REPORTER GENES • Kang and Chung
JOURNAL OF NUCLEAR MEDICINE Vol. 49 No. 6 (Suppl) June 2008

Molecular Genetic Imaging Enables Early Detection

Genetic problem → Molecular-genetic imaging

↓

Biochemical change → PET, MRS

↓

Physiologic change → SPECT, MRI

↓

Anatomical change → Radiography, Sonography, CT, MRI

IMAGING OF REPORTER GENES • Kang and Chung
JOURNAL OF NUCLEAR MEDICINE Vol. 49 No. 6 (Suppl) June 2008

Applications of PET and SPECT in Nuclear Neurology/Psychiatry

- Dementias (Alzheimer's, Parkinson's, Pick's, Huntington's, vascular, etc)

- Cerebrovascular diseases (hemorrhage, occlusion, TIA, stroke)

- Epilepsy (temporal lobe, extratemporal)

- Brain Trauma/Death (perfusion changes)

- Brain Tumors (recurrence vs post-radioTx)

- Psychiatric Disorders (psychotic, affective, substance abuse and dependence)

WHO Global Burden of Disease

Cause	DALYs
Infectious and parasitic diseases	350,332,571
Neuropsychiatric conditions	193,278,495
Cardiovascular diseases	148,190,083
Malignant neoplasms	75,544,632
Sense organ diseases	69,380,870

DALY = YLL + YLD is Disability Adjusted Life Year equal to Years Life Lost plus Years Lost to Disability. Note that this measure "DALY" does NOT assess burden in terms of financial costs nor psychological costs to family and community caring for the "disabled" individual. Statistics for year 2002 as revised in 2004 Annual Report of the WHO.

Brain imaging: What do we need?

- Bioware: Radiopharmaceuticals with radioisotopes tagged to carrier molecules that can cross the blood-brain barrier and bind to receptors in the brain

- Hardware: Efficient multi-modal scanners that can perform integrated PET-MR, PET-CT, SPECT-MR, SPECT-CT, tri-modal combos

- *Software: Effective imaging informatics systems useful for all stakeholders and contexts*

PORTAL-DOORS for the Semantic Web
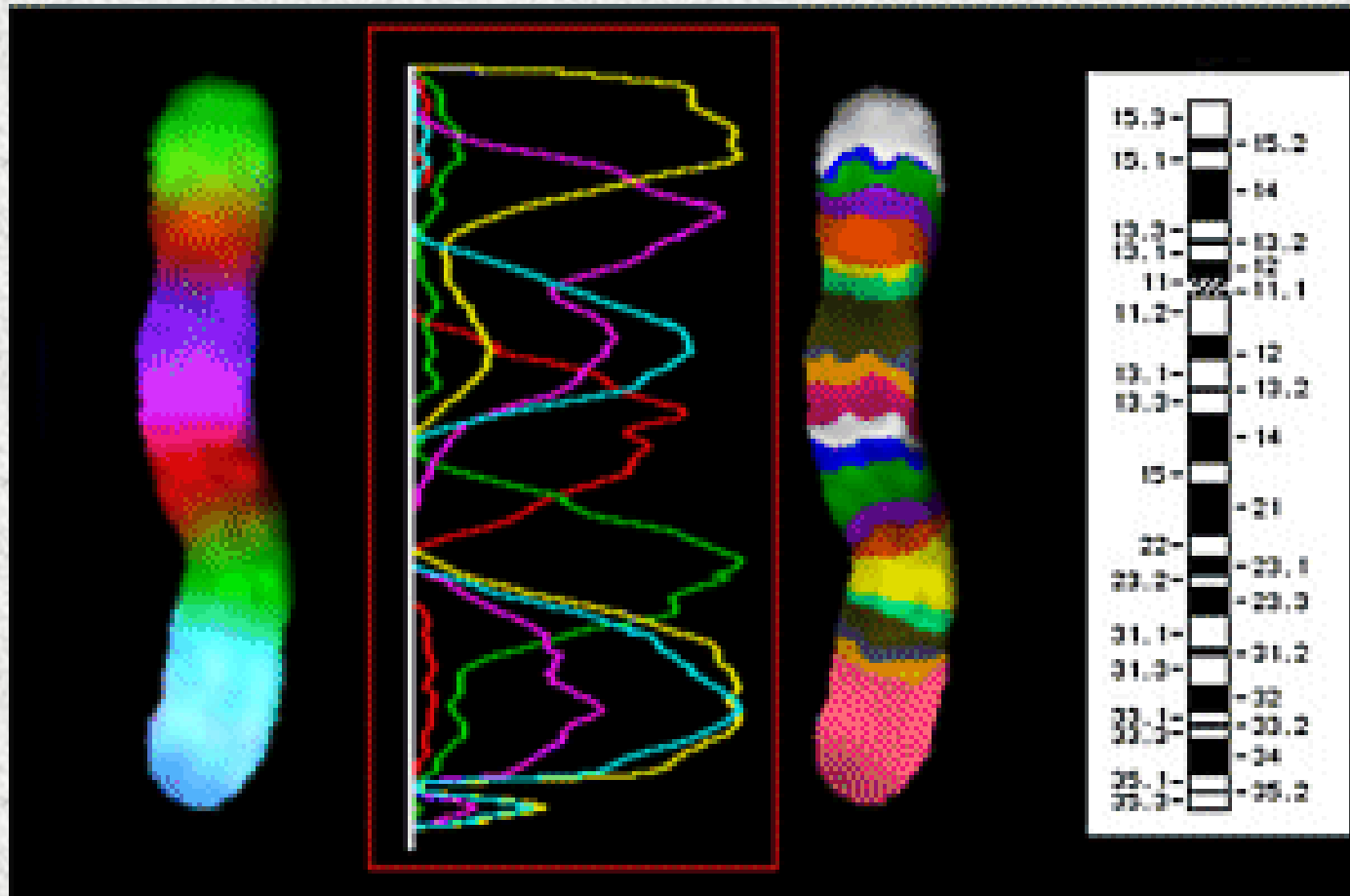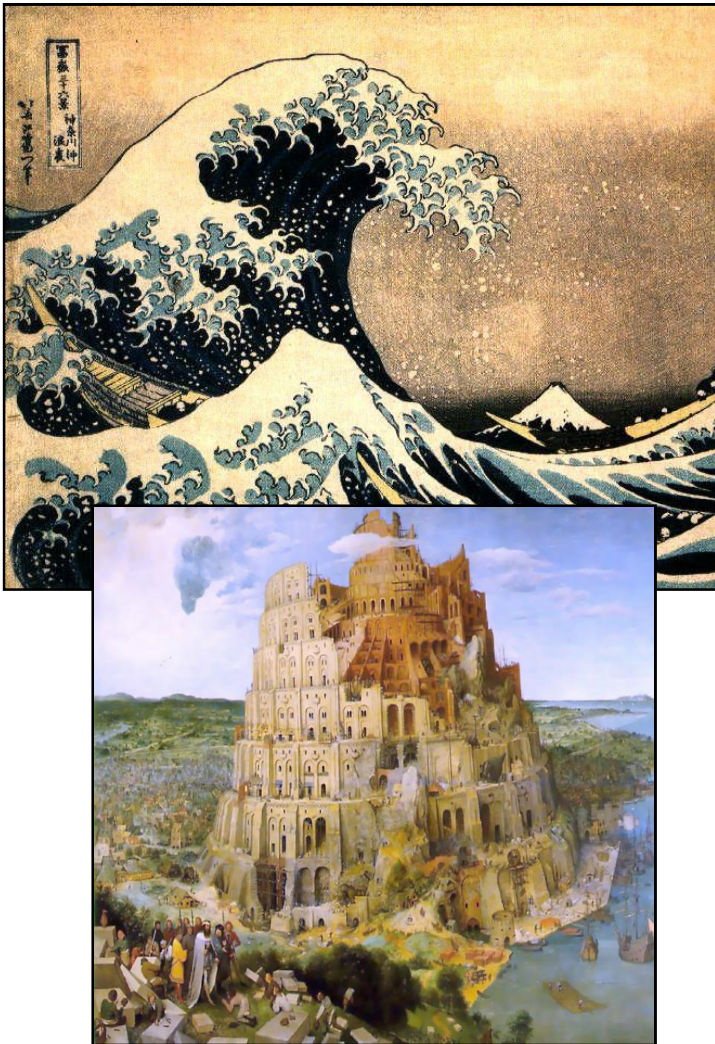
Together, we face:

- **A Biomedical Information Tsunami**
  - Overwhelming volume of data
  - Multitude of sources

AND

- **An Informatics Tower of Babel**
  - Each cancer research community speaks its own scientific "dialect"
  - Integration critical to achieve promise of molecular medicine

The Problem of Cybersilos

- caBIG and caGRID built for cancer and oncology
- What about heart disease and cardiology? Or brain disease and neurology/psychiatry? Or genetics and all fields of medicine?
- What about chemistry and organic molecules? Or physics and radioisotopes?
- What about plants, animals, and biodiversity? Or all sciences and "astronomy to zoology"?
- *Have traditional silos been replaced by cybersilos? Or truly interoperable informatics systems networked without barriers to the freely flowing exchange of information?*

Semantic System

- A lexical ("dumb") system is an informatics system in which words are processed as character strings that have no meaning to the processing agent;

- A semantic ("smart") system is one in which words have defined meaning to the agent processing them with logic-based reasoners;

- *Semantic search can be very efficient while lexical search can be very inefficient*

Semantic Search & Applications

- Current web search often yields irrelevant or innumerable results that lose practical usefulness (too costly in time to review)

- Fast accurate delivery of relevant information remains important motivating driver for development of semantic web search applications (find needle in haystack)

- *Translational medical research with drug development, clinical trials, and personalized medicine (including smart search through records linked by genetic pedigree)*

How many hits? Question answered?

Semantic & Open-World Computing

- Berners-Lee et al (2006 Science 313:769) leading efforts to develop the new interdisciplinary field of "web science"

- Transition from closed-world architectures (static, centralized, monolithic) to open-world environments (dynamic, modular, distributed)

- *W3C Semantic Web built on XML, RDF, OWL, SPARQL with ontologies and reasoners that enable computing with "smart" (rather than "dumb") systems capable of understanding words with meaning in context*

W3C Semantic Web Stack
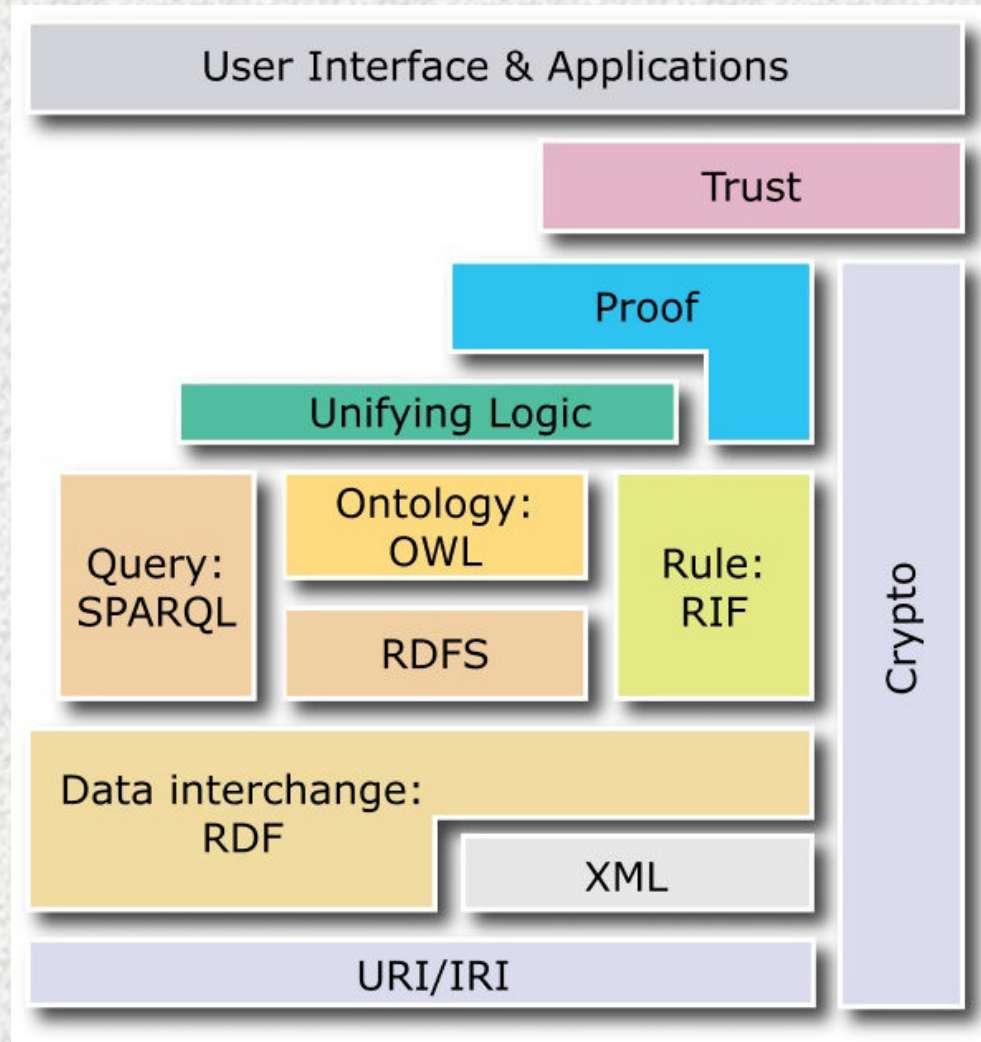
Life Sciences Web & Grid

- Significant development of biomedical ontologies and semantic systems in clinical informatics and bioinformatics (eg, see NAR annual issues on databases and web servers)

- Numerous commentators and critics (Philipi, Neumann, Cannata, Buetow, DeRoure, etc) discuss many problems of current systems

- *Almost all concluded that a new infrastructure is required for growth of the semantic web and grid*

- Good (2006) only exception who blamed people rather than technology as barriers to progress

PORTAL-DOORS vs IRIS-DNS

- Problem Oriented Registry of Tags And Labels (PORTAL) designed as an IRIS analogue

- Domain Ontology Oriented Resource System (DOORS) designed as a DNS analogue

- *PORTAL-DOORS for the semantic web modeled on the success of IRIS-DNS for the original web*

- PORTAL-DOORS uses an analogous paradigm with resource labels instead of domain names

- Taswell 2008 IEEE TITB 12(2):191-204

Resource Labels vs Domain Names

- Resources with labels are more general and flexible than domains with names
- *A resource may be any entity, abstract or concrete, offline or online*
- Its label may be any URI or IRI
- Its non-semantic tags may be any tokenized names or phrases
- Its semantic description may be any set of RDF triples referencing any ontologies

Examples of Resources

- Persons, patients, investigators, organizations

- Online virtual entities, offline physical entities

- Data services, databases, database records

- Analysis services, data processing tools

- Authored information, books, journals, papers, web sites, web pages

- Many more examples and categories

PORTAL-DOORS Data Records

PORTAL-DOORS System

Network of **PORTAL registries**

Network of **DOORS directories**

PORTAL record

Required:
1. **Entity label**
2. Entity owner
3. DOORS directory
4. Record status

Permitted:
5. **Entity tags**
6. Entity crossrefs
7. Other metadata
8. Owner signature

DOORS record

Required:
1. Entity label
2. **Entity locations**
3. PORTAL registry
4. Record status

Permitted:
5. **Entity description**
6. Record distribution
7. Record provenance
8. Record signature

Register labels/tags

Publish locations/descriptions

Agent

PORTAL-DOORS network server interaction

Search PORTAL

Search DOORS

User

Lexical server network

Semantic server network

PORTAL-DOORS Server Network

Bootstrapping Combined Design:
-- NEXUS registrars
-- PORTAL registries
-- DOORS directories
-- NEXUS servers operate as
   registrar only or as integrated
   registrar/registry/directory

Original Separate Design:
-- PORTAL registrars/registries
-- DOORS directories

NEXUS Root

NEXUS A Pri/Sec   NEXUS B Pri/Sec   NEXUS C Pri/Sec

DOORS Root

DOORS X Pri/Sec   DOORS Y Pri/Sec   DOORS Z Pri/Sec

PORTAL Root

BrainWatch Pri/Sec   ManRay Pri/Sec   GeneScene Pri/Sec
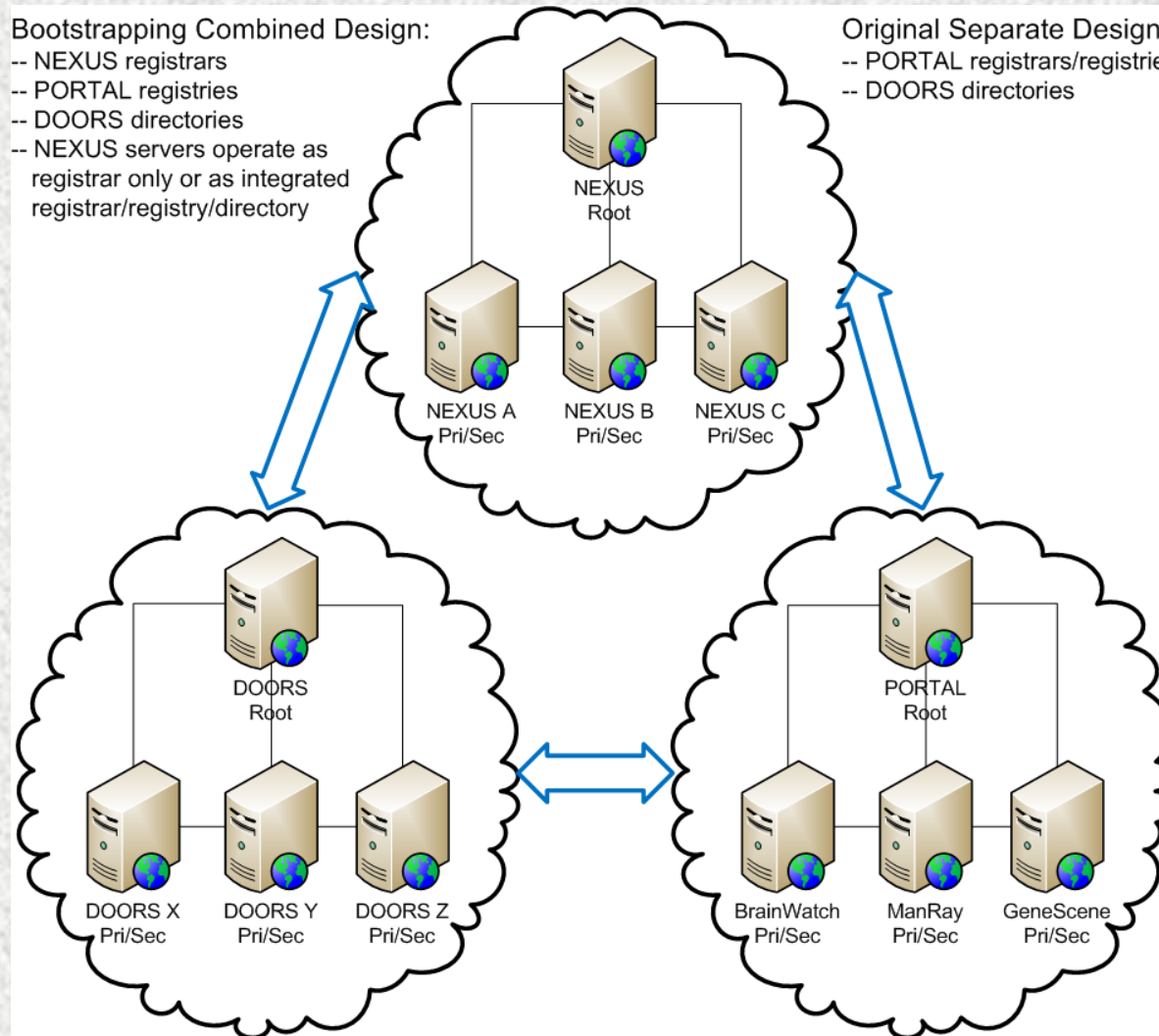
PORTAL-DOORS: Some Key Principles

- Assures globally unique identification of resources while promoting interoperability
- Enables cross-registry and cross-directory searches between different problem domains
- Problem oriented, not technology restricted
- *A distributed infrastructure that permits localized control of policies and content*
- A hybrid bootstrap and bridge to transition from old lexical web to new semantic web

PORTAL-DOORS: Some Other Points

- Physical infrastructure, content control, and content itself are all distributed and shared
- *Analogous to DNS where data records are distributed and mobile with request forwarding and response caching*
- Compare Wikipedia where content is centralized but control of content is shared
- Compare Google where infrastructure is distributed but not the control of content

Democratization of Search?

- PORTAL-DOORS interlinks registries, directories, databases across fields, disciplines, specialties
- PORTAL-DOORS supports mass collaboration via its hierarchical and distributed but decentralized and localizable infrastructure
- *PORTAL-DOORS provides a democratized solution to the problems of search engine consolidation*
- Mowshowitz and Kumar in Feb 2009 IEEE Computer (p108) discuss the *realities and risks* of search engines that restrict access to information

Initial PORTAL-DOORS Sites

- ## PORTAL-DOORS Project
  - www.portaldoors.org
  - with read only access for Resource Users
- ## GTG Registrar at
  - www.telegenetics.net
  - with read/write access for Resource Agents

Resource Entity Types

Initial PORTAL Registries

Agent in Edit Mode at GTG Registrar

3/15/2009 (c) GTG   Presented at AMIA STB, San Francisco   Carl Taswell, MD, PhD

Generation of Content

- Human agents initially

- Webbots and converters when available

- Trade-off in quality of content produced versus rate of content production

- Trade-off not applicable to situations where *existing databases only need an appropriate interface for inbound queries and wrappers for outbound responses*

A Vision for the Future
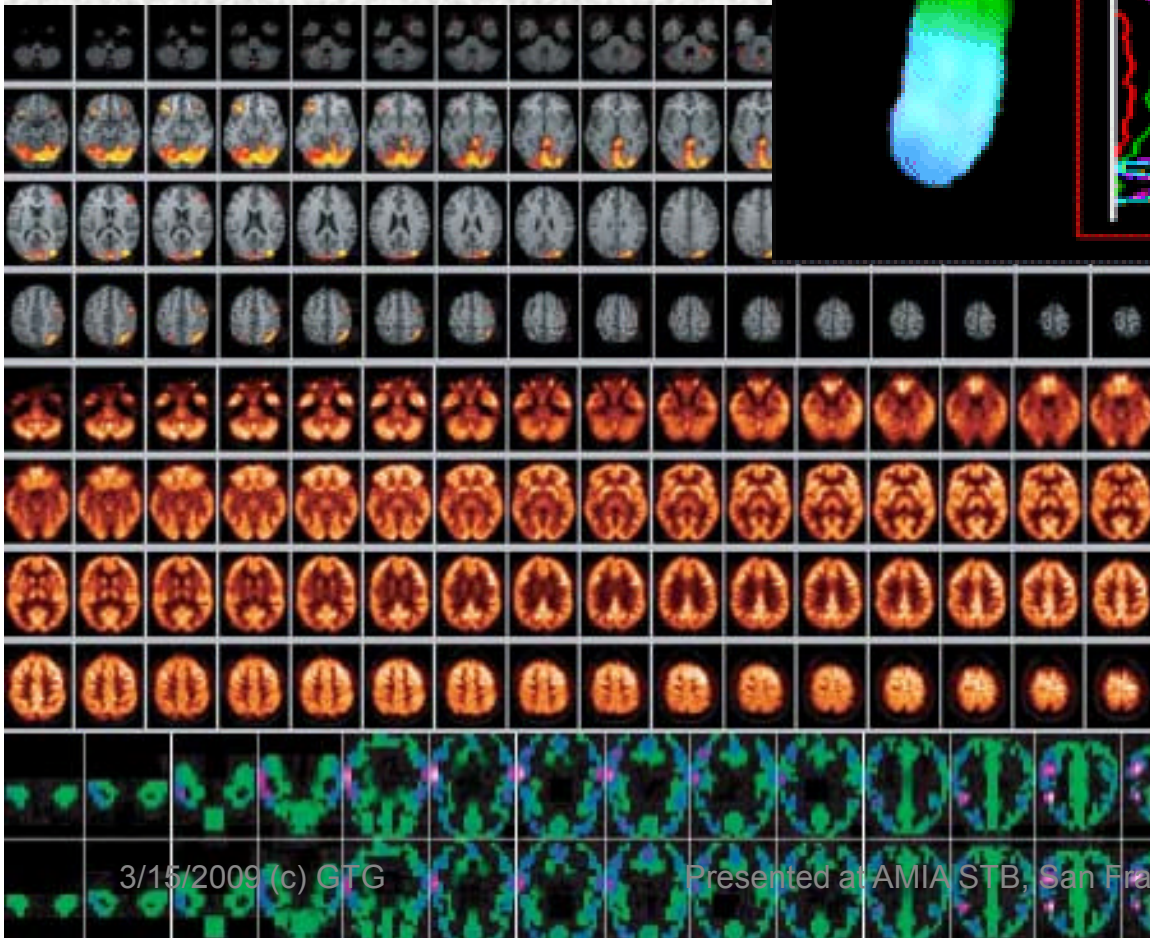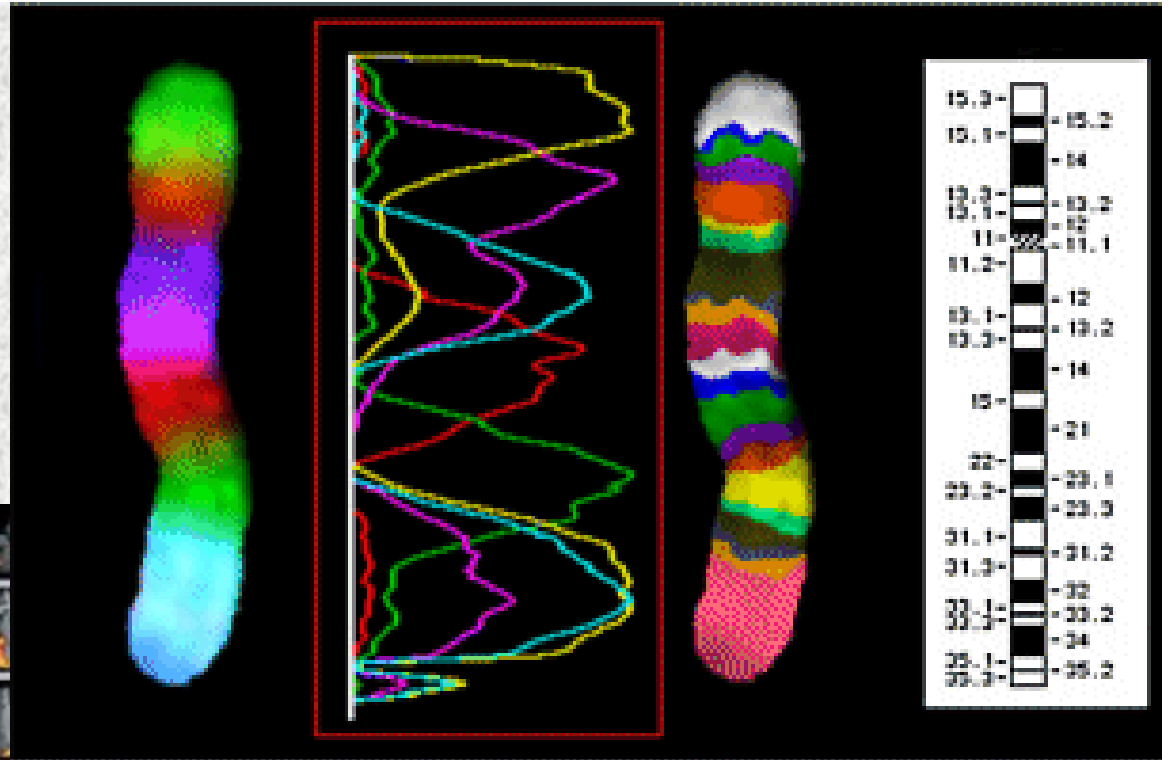
Biomedical Informatics for
Gene – Brain – Behavior Relationships

Applications of PORTAL-DOORS in Functional Brain Imaging

- BioPORT, GeneScene, ManRay, and BrainWatch as registries for resources relevant to functional brain imaging

- Pharmacogenomic molecular imaging informatics for Alzheimer's disease and other neurodegenerative disorders

- *Knowledge management system for tracking radiotracers and imaging biomarkers that enable mapping neural pathways in the brain*

Applications of PORTAL-DOORS in Clinical Trial Registries

- Promote efficiency, standardization and harmonization by managing lists of qualified investigational sites, INDs, and CMCs for radiopharmaceuticals

- Manage information about patients as the "resource" in multi-center clinical drug trials with imaging biomarkers

- Track the images themselves as the "resource" in data grids for multi-modal imaging trials

- *Applications determined by the investigator's definition of "resource" and focus of interest*

Pharmacogenomics Molecular Imaging Study
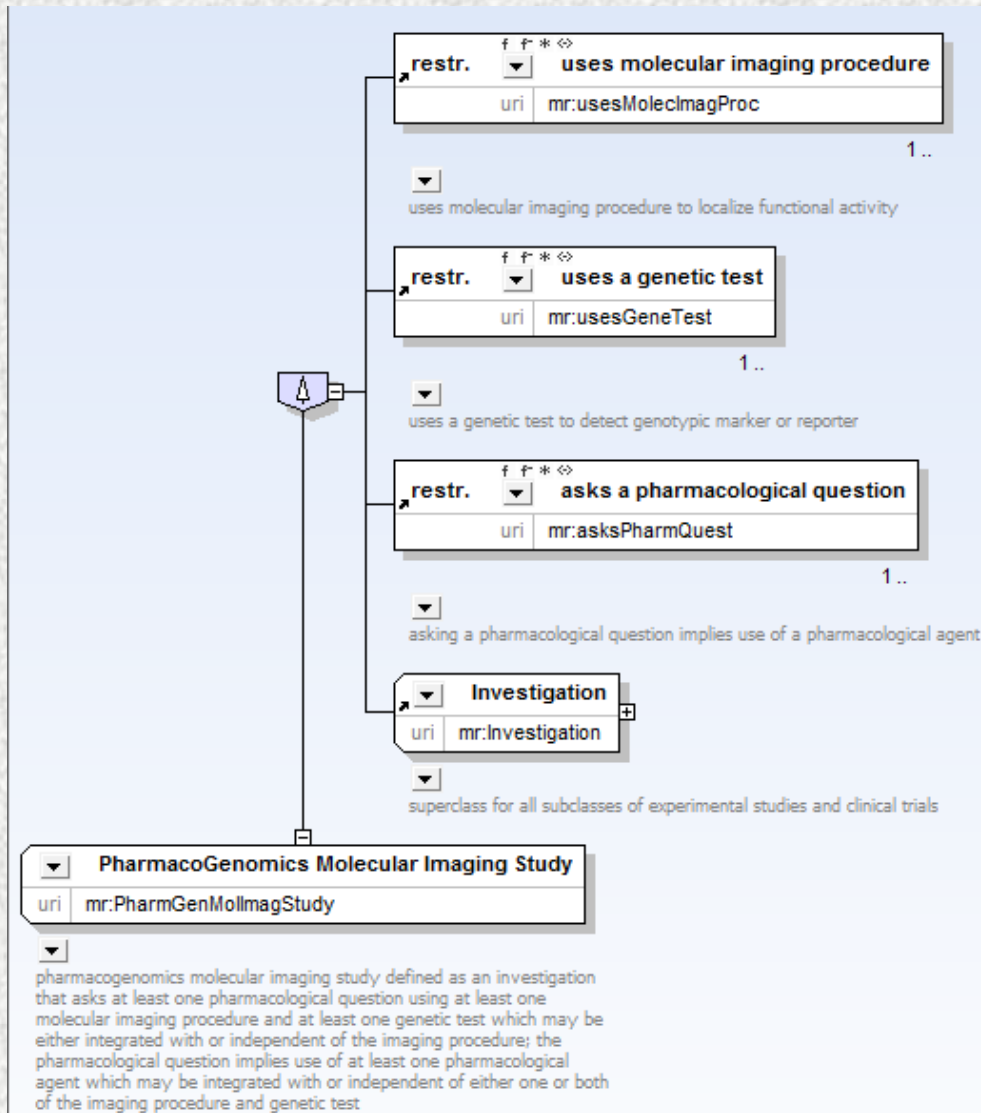
- Ontology class definitions (such as this example from ManRay ontology) enable new studies
- Drive from question to answer through semantic web of data
- Guided by map of PORTAL-DOORS interconnections

Roadmap for PORTAL-DOORS

- Version 0.5: Current live implementation with back-end database and front-end web browser client for lexical PORTAL functionality

- Version 0.6: Implementation as web service with both desktop and browser clients;

- Version 0.7: Semantic DOORS functionality

- Version 0.8: Revision of PORTAL primary server functionality

- Version 0.9: Revision of DOORS primary server functionality

- *Version 1.0: First "official" release of PORTAL-DOORS System*

- Version 2.0: Interacting primary and secondary server functionality with request forwarding and response caching amongst distributed servers

Thanks!   For More Details…

- Visit with me during Poster Session I at "Implementation of Prototype Biomedical Registries for PORTAL-DOORS" (Board 26)
- Contact me via websites:
  - www.portaldoors.org
  - www.telegenetics.net