

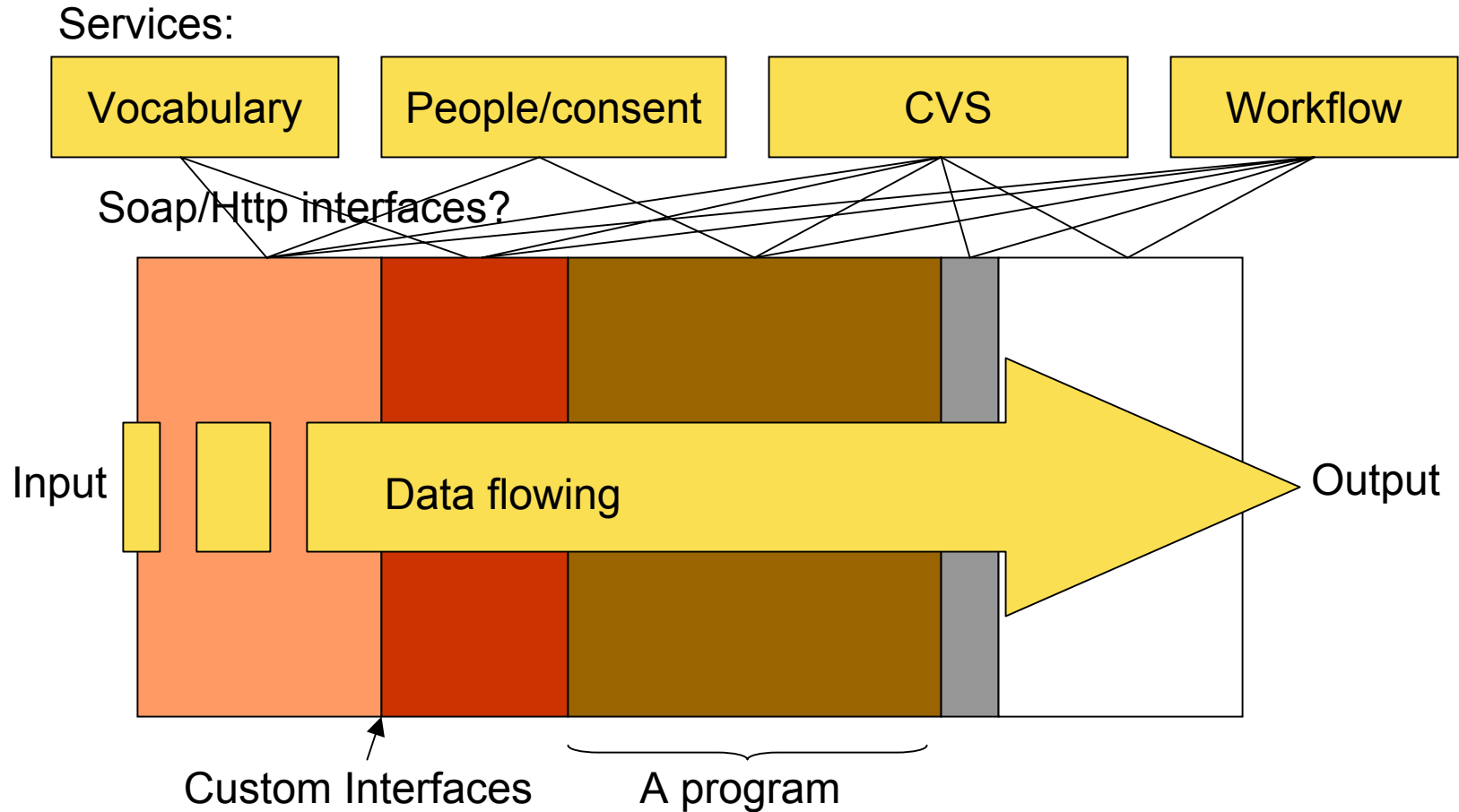
Working to improve workflow

- Workflow solution
 - Unify data processing pipeline solutions at various sites to become single mBIRN solution
- Benefits of single mBIRN solution
 - Set up collaborations faster
 - Greater exploration of recalculated experiments
 - Ability to routinely explore complex parameter spaces

Improve workflow

- Problems to be solved
 - Identify bottlenecks in workflow
 - Simplify interfaces between processing systems
 - Adapt to culture focused on programming (vs. databases)
- Requirements for pipeline
 - Build completely open source solution that is based upon existing workflow expression standards
 - Ensure architecture can solve current mBIRN use-cases (SASHA, MIRIAD, AD)

Data Pipeline Overview



Data Pipeline Vocabulary Service



- The Vocabulary service is where programs in the pipeline will go to look up and register terms, performing QA and preemptive additions of new terms.
- An important function is to provide tables that map one similar term to another similar term, and contain lists of acceptable elements for various data elements.

Data Pipeline Person Service



- The People/consent service is where lookups are made during the data processing to make map aliases (usually numbers of some kind) to the person, and assure that permission exists in the form of consent to proceed with the movement of data.
- It's most important function is to provide tables that map one identifier to another identifier for that person.
- It can also be used to obtain minimal demographics and provide matching services to resolve people based upon fuzzy data.

Data Pipeline CVS Service



- The CVS service is where all processing programs/scripts are cataloged and versioned. Programs that were used in the transformation process are available in the version for which the transformation was performed.
- An important function is to provide tables of transformations that can be used to look up and recreate a transformation.
- This is especially essential with regard to realizing flow diagrams.

Data Pipeline Workflow Service

People/consent

CVS

Workflow

- The Workflow service is the control center of the pipeline. It records processes that have been run and are set up to run, what versions are running and what the error status is of each completed job.
- An important function is to provide schedules for running Jobs.
- Also monitors computer and file space, permissions

Pipeline Schematic

A row in the vocabulary database

Alzheimer's disease, late stage	SN8745	\\Diagnoses\neurology\dementia\Alzheimer's\late_stage
---------------------------------	--------	---

A row in the master subject-consent database

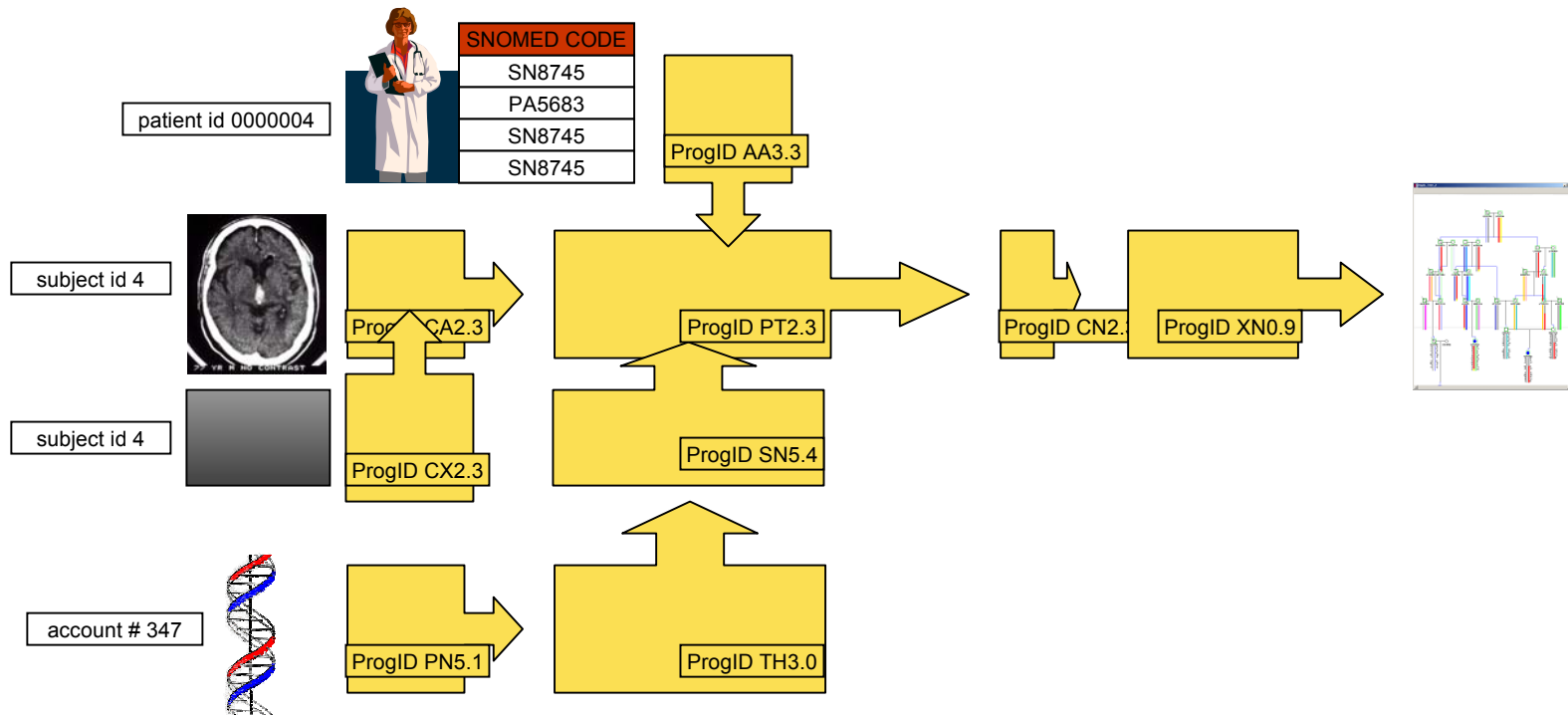
Shawn Kennedy (research subject) Master ID 34	patient id 0000004	subject id 4	account # 347	Consented for study X and Y
---	--------------------	--------------	---------------	-----------------------------

A row in the CVS database

Gradient un-warping	Current version CA2.3	5/02/2004	Available at http://www.site.edu/un-warping
---------------------	-----------------------	-----------	--

A row in the Flow reporting & control database

12 images processed	Prog ID CA2.3	11/02/2004	Owner: Jorge	No errors
---------------------	---------------	------------	--------------	-----------



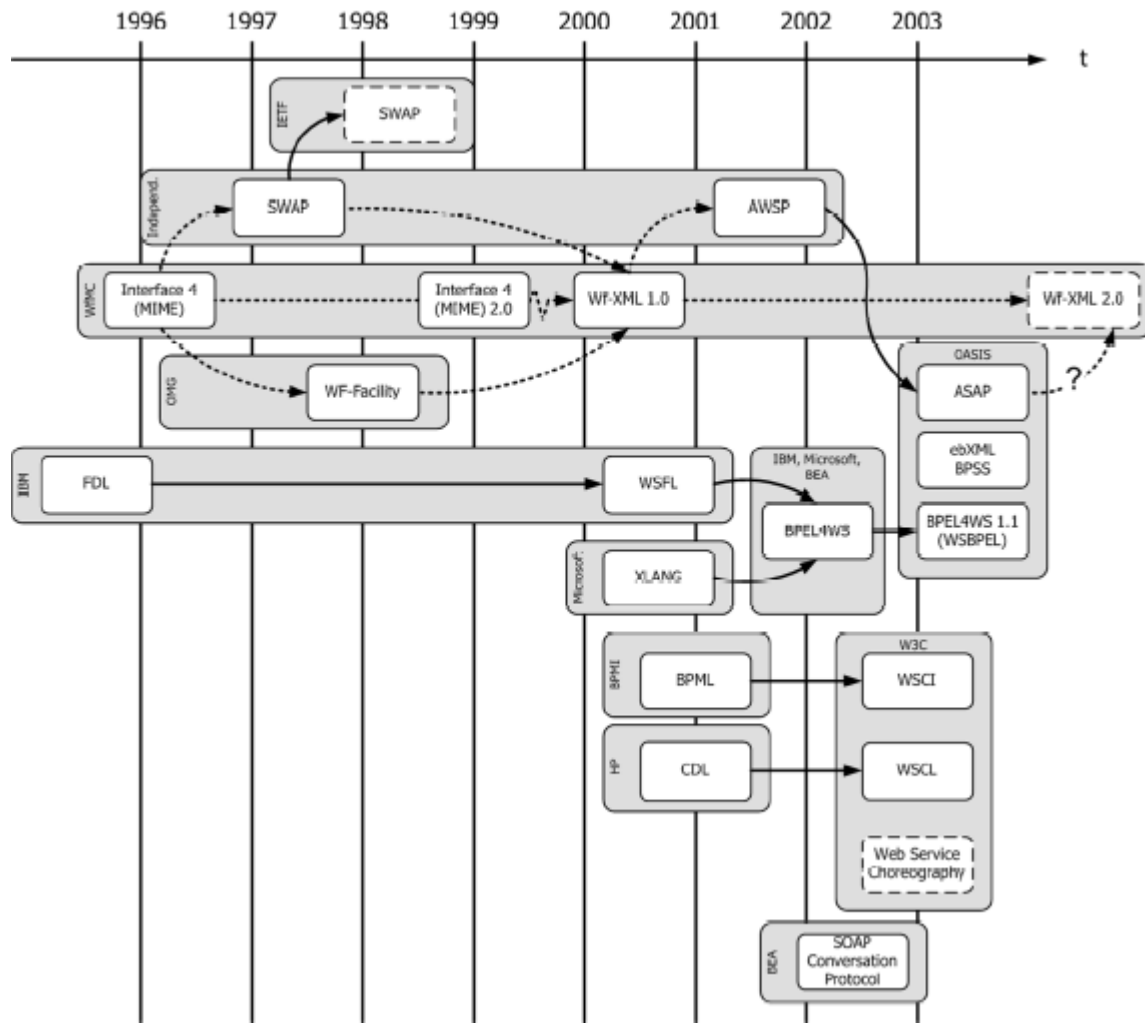
Standards

Established Standardization Players

- **Workflow Management Coalition**
 - Workflow Process Definition Language (WPDL/XPDL)
 - Workflow Process Interchange (IF 4/Wf-XML)
- **Object Management Group**
 - OMG Workflow Facility
 - Process Definition RFP (bom/99-10-03)
 - Extension of UML for Workflow Modeling
- **National Institute for Standards and Technologies/MIT**
 - Process Specification Language (PSL)
 - Process Interchange Format (PIF)

More Recent Standardization Players

- **Business Process Management Initiative (BPMI)**
 - Business Process Modeling Notation (BPMN)
 - Business Process Modeling Language (BPML)
 - Business Process Query Language (BPQL)
- **OASIS**
 - Business Process Execution Language for Web Services (BPEL)
 - ebXML Business Process Schedule Specification (BPSS)
 - Asynchronous Service Access Protocol (ASAP)
- **W3C**
 - Web Services Choreography Definition Language (WS-CDL)
 - based on Web Services Conversation Language (WSCL) and Web Services Choreography Interface (WSCI)
- **DARPA**
 - DARPA Agent Markup Language – Services (DAML-S)



Synopsis on Standards

- The most well developed standard so far appears to be WPDL/XPDL from the standpoint of number of supporters and open source packages using it. It also has a lot of work going on with it, evolving it into Wf-XML 2.0. The main site is <http://www.wfmc.org/> which is the home of the organization that manages this standard. Possibly use this one?

Requirements

(some kinds of requirements to consider are listed on the next few pages, specific requirements to be developed by mBIRN pipeline group)

Criteria: System Specifications

- System
 - Origin → Determines system functionality
 - Stand-alone or embedded system
 - Hardware: Host/Mainframe, Workstation, PC, Mac.
 - Server-OS; Client-OS / Web-based clients
 - Databases: proprietary solution, solutions from other vendors (included or to be purchased separately)
 - Scalability (multi-server operation etc.)
 - System stability and security
- Interfaces
 - Databases
 - Document Management Systems
 - Invoked Applications
 - Organizational Repositories
 - Modeling and Simulation Tools
 - Client Interfaces
 - Standard APIs (WAPI)

Criteria: Workflow-Modeling

- Process Modeling
 - Graphical editor
 - Model building blocks, inheritance, libraries
 - Dynamic modification at runtime
 - Form design (RAD option)
 - Control flow elements
 - Full-featured business process management suite
incl. Business Process Modeling + Simulation + Evaluations
 - Testing tools (Simulation)
- Organizational Modeling
 - Entity types (fixed / extensible)
 - Assignment mechanisms / policies
 - Authorization for invoked applications
 - Access of X.500 directories

Criteria: Workflow-Run-Time

- Monitoring of deadlines, reminder, escalation (to which level?)
- Deferred execution, re-submission of work-items
- Delegation of activities
- Substitution mechanisms / proxies
- Skipping/Insertion of activities
- Workload-dependent staff resolution
- Cancellation, roll-back of workflows with compensation mechanisms
- Run-time monitoring, status-tracking of workflow participants
- Notification of participants via e-mail
- Integrated groupware functions (group calendaring etc.)
- Customizable worklist
- Customizable logging of run-time events (audit trail)
- Analysis tools for audit trail data

Web Services

- Access remote service
- Protocol Based: Technology Independent
 - .net
 - Java
- XML Based: X-platform, X-OS
- WSDL definition of the interface
- UDDI to look for service

Resources

Websites to get started

- www.e-workflow.org
 - Case Studies
- www.workflow-research.de
 - Discussion Forum & Papers (& this Tutorial)
- www.processmining.org
 - Tools & Papers
- www.workflowdownload.com
 - Links to demo versions of commercial systems
- <http://dmoz.org/Computers/Software/Workflow/>
 - Directory of workflow web sites

Standards and Open Source Projects

- www.wfmc.org
 - Glossary, XPD, Wf-XML
- www.bpml.org
 - BPML
- www.bpmn.org
 - BPMN
- www.oasis-open.org
 - ASAP, BPEL, BPSS
- www.w3c.org
 - WS-CDL working group
- www.omg.org
 - Workflow Facility
- jawe.objectweb.org
 - XPD editor
- enhydra.objectweb.org
- www.jbpm.org
- www.openbusinessengine.org
- bonita.objectweb.org
- xflow.sourceforge.net
- wfmopen.sourceforge.net

Good Introductory Articles

- van der Aalst, W. M. P.; ter Hofstede, A. H. M.; Kiepuszewski, B.; Barros, A. P.: Workflow Patterns. *Distributed and Parallel Databases*, 14 (2003) 3, pp. 5-51.
- van der Aalst, W. M. P.: The Application of Petri Nets to Workflow Management. *The Journal of Circuits, Systems and Computers*, 8 (1998) 1, pp. 21-66.
- Basu, A.; Kumar, A.: Research Commentary: Workflow Management Issues in e-Business. *Information Systems Research*, 13 (2002) 1, pp. 1-14.
- Kumar, A.; Zhao, J. L.: Workflow support for electronic commerce applications. *Decision Support Systems*, 32 (2002) 3, pp. 265-278.
- Stohr, E. A.; Zhao, J. L.: Workflow Automation: Overview and Research Issues. *Information Systems Frontiers*, 3 (2001) 3, pp. 281-296.
- zur Muehlen, M.: Evaluation of Workflow Management Systems Using Meta Models. In: 32nd Hawaii International Conference on Systems Sciences (HICSS 1999). Ed.: Ralph Sprague, Jr., Wailea, HI, 1999.