



Department of Defense Human Resources - Enterprise Information Warehouse/Web (EIW)

Using standards to Federate and Integrate
Domains at DOD



Federation Defined

- Members of a federation agree to certain standards to interoperate and relate to each other for the common good.
- In a federation participants create their own policies, systems, facilities and delegate some authority to the federal authority

The United States of America is a federation



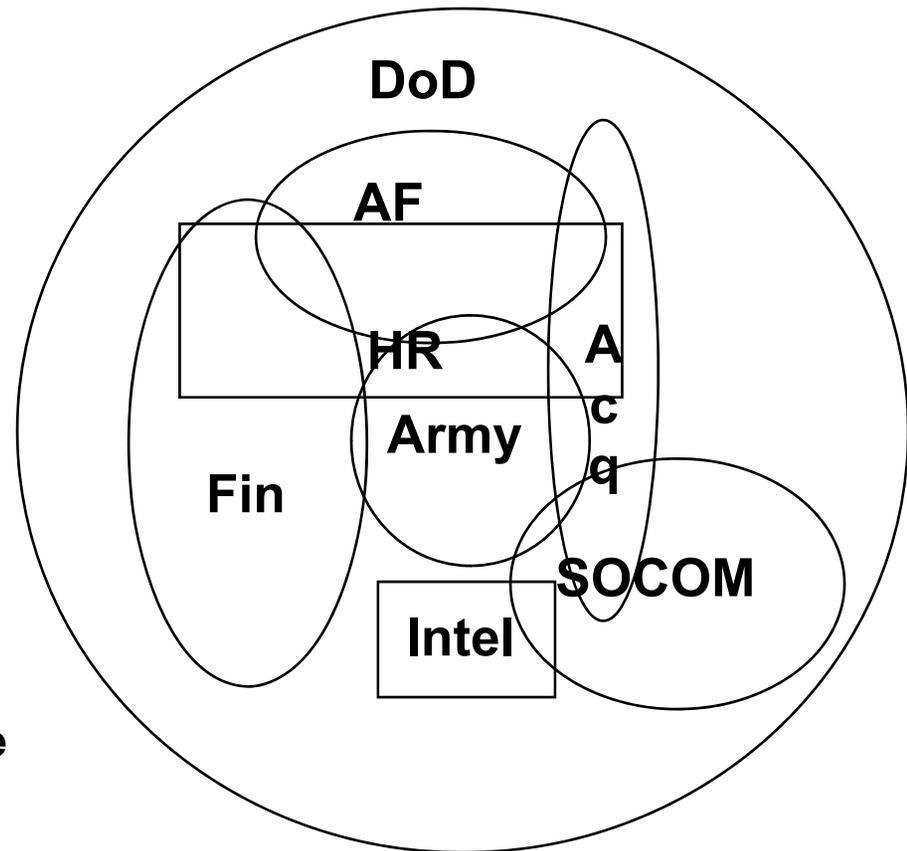
DOD is a federation





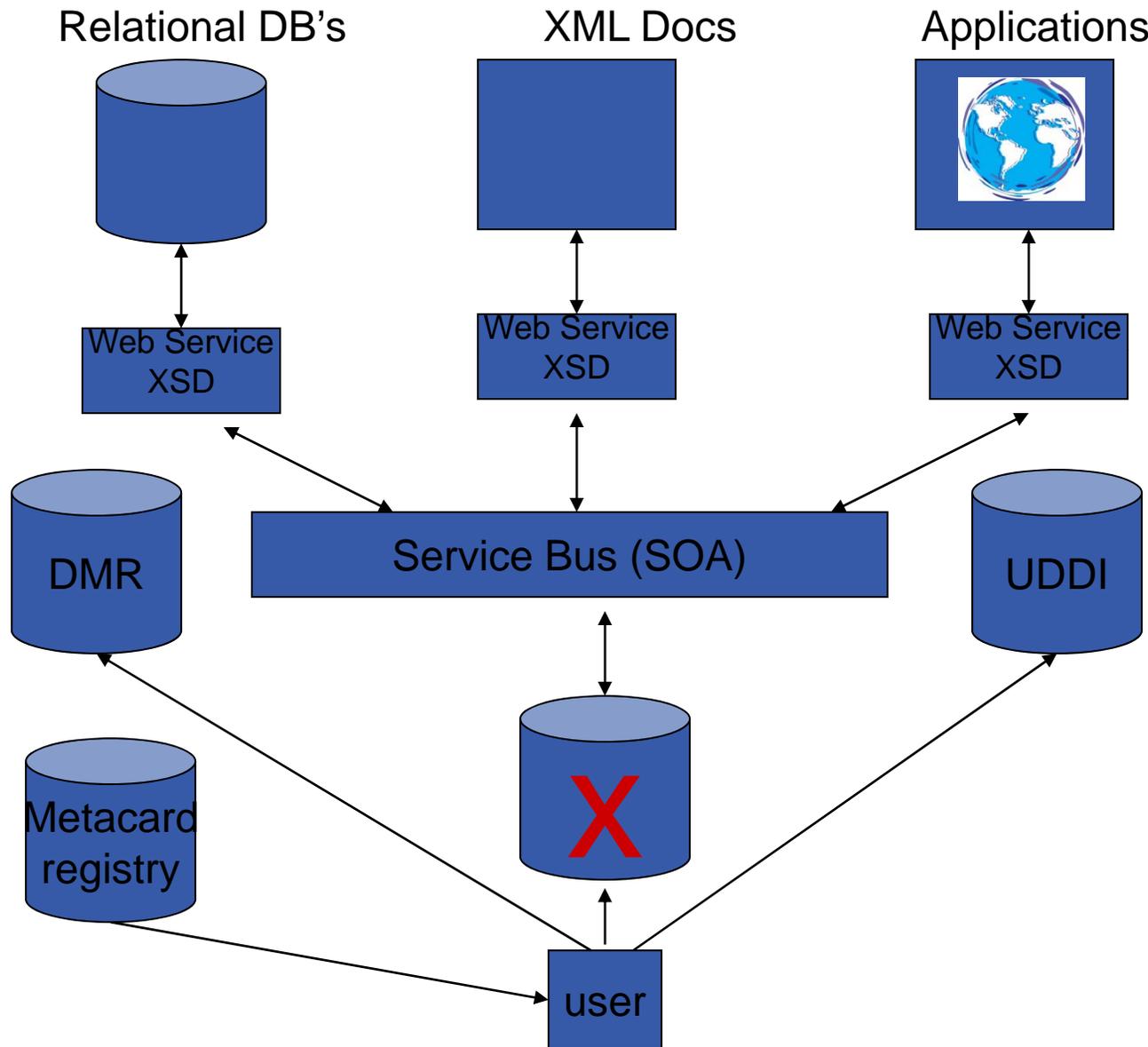
DOD Federation Problem

- **DOD is made up of many domains within domains**
 - ❖ Army, Navy AF, Marines, OSD, JFCOM, ...
 - Logistics, HR, Finance, Command and Control, intelligence ...
- **Each domain fields its own applications and creates its own information to execute its mission**
 - ❖ It is often not possible to federate and integrate applications within domains
- **In many cases it is *necessary* to share data within domains and across domains**
- **DOD will never meet its information sharing needs until it first solves the federation problem**





NCES – Current Policy

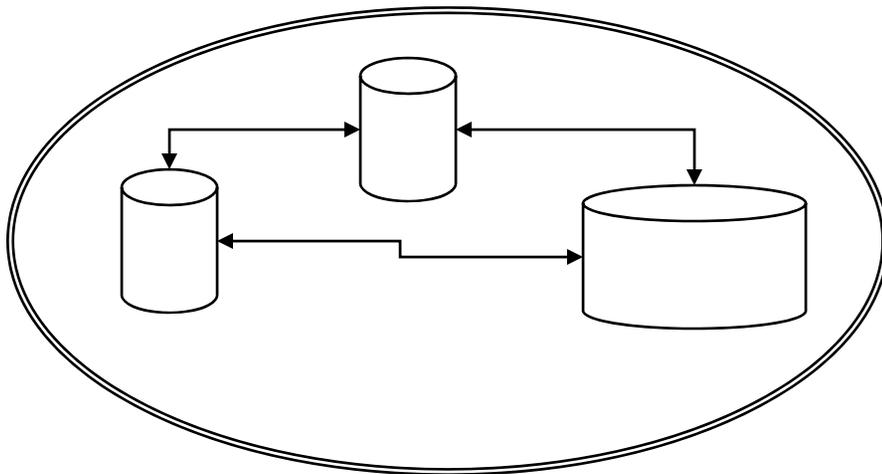


1. Information Systems
2. Application information wrapped in web services
3. Service bus connects web services to DOD networks
4. No way to know how the information is related
5. No integrated data for the user, No federation



Information Federation

- **Federating communications networks does not mean that “information” is federated.**
 - ❖ Many people think that NCEs is an information federation architecture, it isn't
 - ❖ It is a network federation architecture
- **TCP/IP, HTTP, XML let machines know how to read packets of bytes, but they do not tell machines what the information contained in the packets means**
 - ❖ It is easy to tell a machine what a TCP/IP packet means so we just build the knowledge into the hardware or firmware
- **Semantics must be added to the packets so that machines know how the information contained in the packet is related and what each term means.**
 - ❖ It is complex to tell a machine what a term means, so we build ontologies
 - ❖ A “tank” in one domain is related to liquid, and in another domain “tank” is related to vehicle
- **If a machine can understand the similarity or difference of meaning of terms in domains, then we have a federated information architecture**



≠

***Information
Federation***



Federation Solution

■ Enterprise Information Web

- ❖ Any information from any system can be shared with any other system on the Global Information Grid

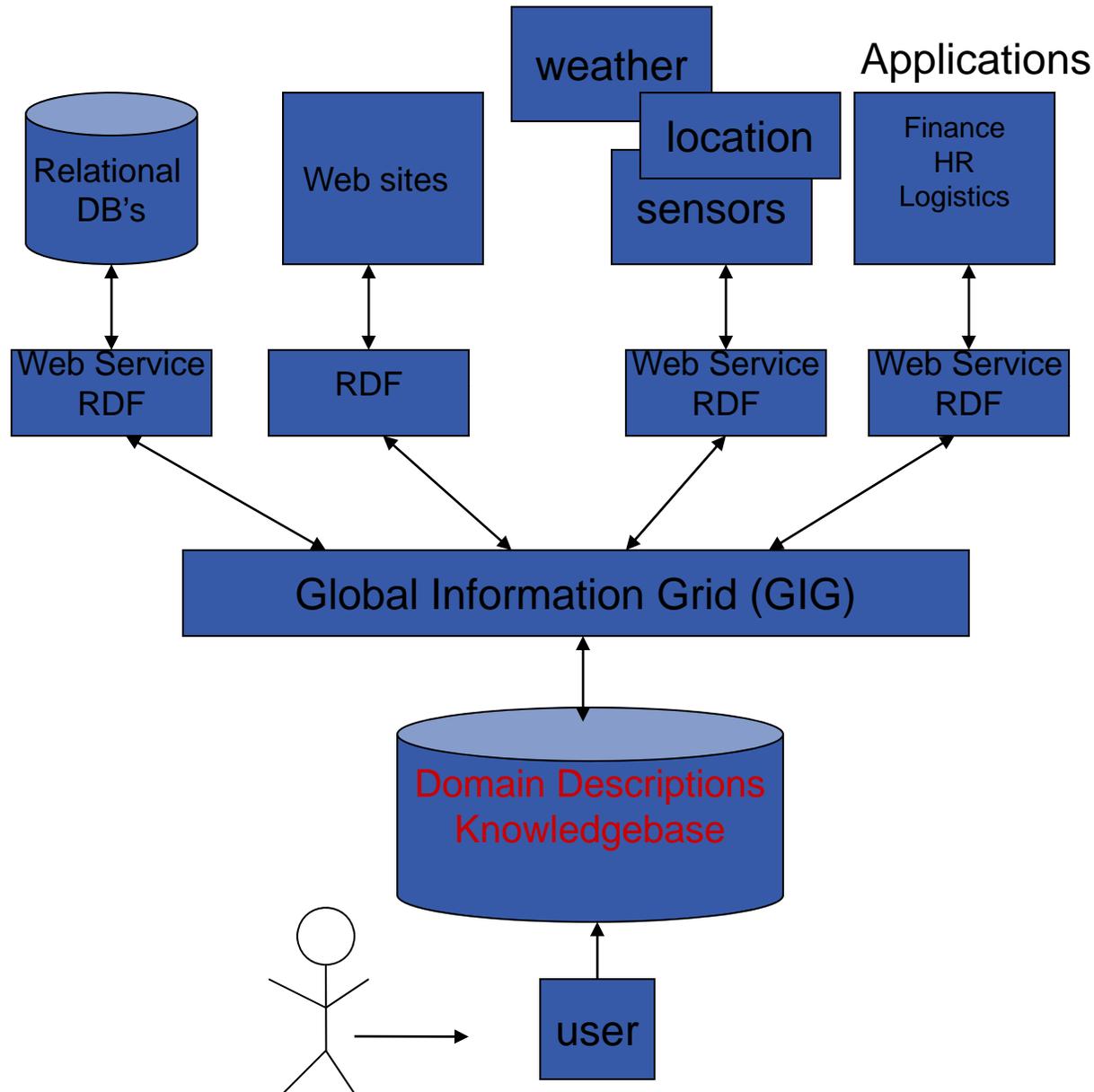
■ Steps

- ❖ Describe all of the artifacts in each domain using standards (RDF, OWL)
 - We currently do this description work, but we do not use standards – Excel, Word, Powerpoint, Visio
 - The formal description of a domain is called a Domain vocabulary
- ❖ Use these standards based descriptions to say how domains are related
 - this is the big missing piece of the current approach
- ❖ Use these standards to say how all of the data in each domain is related to the domain vocabulary
- ❖ Query the Domain vocabularies for any information

■ The result is an Enterprise Information Web that meets the goals of information sharing as laid out in numerous DOD policy statements



Enterprise Information Web



1. Information Systems
2. Expose as RDF web services or SPARQL endpoints
3. GIG contains self described data
4. GIG is a big federated knowledgebase of any information
5. Any authorized user or system can query the GIG for any information



Federation Progress

- **DOD has led the charge and made progress on solving the Information Systems federation problem**
 - ❖ Transmission Control Protocol/Internet Protocol – TCP/IP
 - This standard has allowed data networks to be federated
 - ❖ V.6, SMTP
 - Facilitates the federation of telecommunications networks and email

- **In the past ten years standards to federate information have arisen**
 - ❖ DARPA creates the Defense Agent Markup Language program in 1998 to facilitate information federation
 - ❖ W3C takes the work funded by DARPA and creates the Resource Description Framework (RDF) and Ontology Web Language (OWL) specifications
 - ❖ These specs are an integrated part of the W3C stack – HTTP, HTML, XML, XSD, namespaces, URI, and URL

- **Taken together they form the standards on which an Enterprise Information Web can be formed**



Leverage Existing Investment

- **Enabling the GIG as an EIW leverages all of the existing infrastructure**
 - ❖ Same networks, same security, same applications, same organizations
 - ❖ COIs already formed and new COIs will have to do their work in standards based tools instead of proprietary tools

- **DOD is doing a lot of this description work now, it simply requires some redirection**
 - ❖ Must use standards like any other federation

- **The result of this relatively minor change and expense will be an astounding advance in information management capability**



HR EIW Background

- On Jan 16th 2009, Deputy Secretary of Defense issued a memorandum with a directive regarding enterprise level personnel and pay information requirements:
 - ❖ To meet the Department's requirements for enterprise-level information visibility to support the needs of OSD and the Combatant Commands, the enterprise will continue to manage the delivery of these information capabilities under the leadership of the Office of the Deputy Chief Management Officer (DCMO). This will include establishing an enterprise level information warehouse and the necessary functional and technical requirements to enable the delivery of this capability in close coordination with the delivery of the personnel and pay transaction systems by the individual military departments.



HR EIW Problem Statement

■ **DoD currently lacks the capability to quickly and accurately account for personnel, manage troop strength, and war plan based on enterprise level authoritative, real-time personnel information**

❖ Impact: challenges the objective of having the right personnel in the right place at the right time to fight battles and win wars.

❖ Sample COCOM personnel data needs:

➤ Requirement 28.0: Provide timely and accurate information on the location, status, and identity of units, personnel, equipment, and supplies with emphasis on Personnel including patients.

➤ Requirement 124.0: Identify the availability of personnel within the active forces, or within the reserves and other sources when analyzing mobilization alternatives, to satisfy filler, replacement and force expansion personnel requirements.



HR EIW Solution Statement

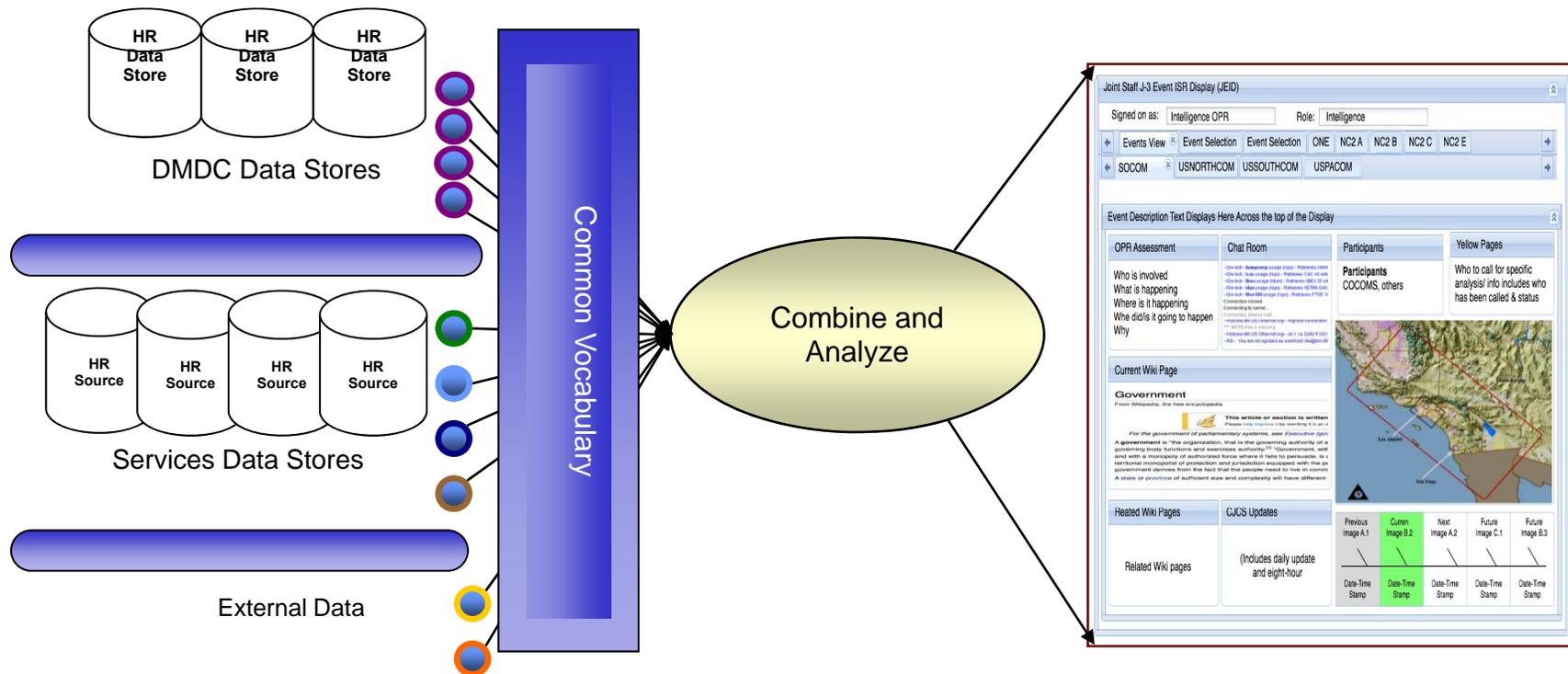
The HR EIW is a mechanism for reaching into service applications to satisfy enterprise HR information needs. It accomplishes three things:

- 1. It reports real-time, authoritative HR information on-demand.**
- 2. It creates HR enterprise information standards.**
- 3. It supports IT flexibility.**

HR EIW At a Glance

Multiple Sources

Single view



Combine multiple data and service sources into single view



Activities and Benefits

WHAT WE'RE DOING

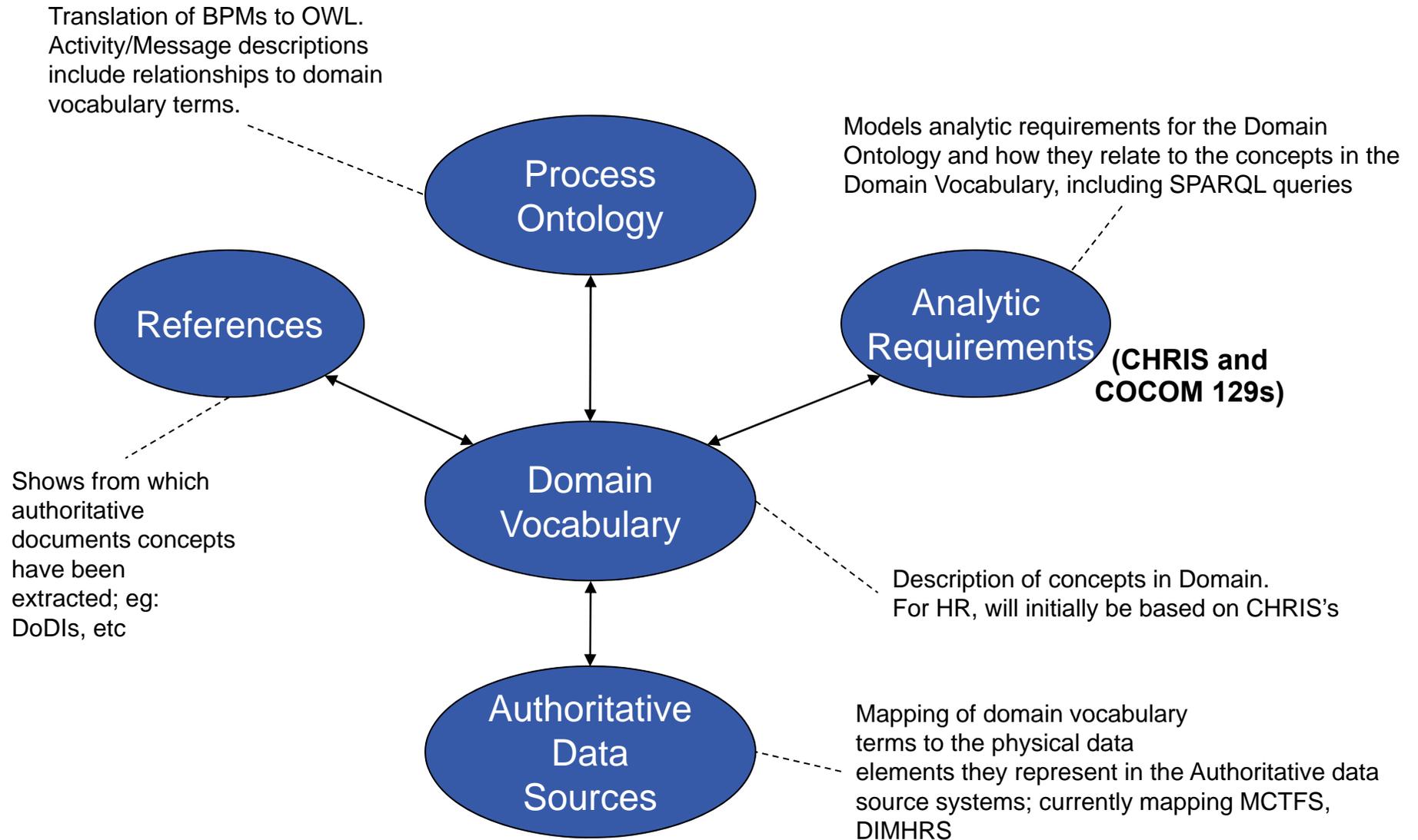
- **Building a DOD HR ontology (or Common Vocabulary) using W3C standards called RDF/OWL resulting in:**
 - A conceptual model that is queryable in a run-time environment;
 - Unambiguously described information in the DoD HR Domain
 - A common vocabulary for information integration (federation)

MAJOR BENEFITS

- **A virtual web of DoD authoritative source HR systems**
- **Answers for any HR information requirements with existing HR systems or ability to show the gaps**
- **A plug-and-play federated environment so new systems or analytical needs can come online and go offline without disrupting the overall environment**
- **Near real-time, on-demand, dynamic and authoritative data**



HR Ontology Architecture





Domain Ontology Defined

- **The Domain Ontology is a conceptual description of the domain covered by the relevant business processes**
 - ❖ The “domain” is defined by the business processes and rules, the information sources, and any reports which are required
- **Instances in this ontology are the same instances which are currently stored in information sources (i.e. databases)**



CHRIS Modeling

- **Each CHRIS is mapped to a property in the ontology**
- **A Property has a Domain and Range**
 - Domain – Defines types of things that can *have* a value for the property
 - Ranges – Defines types of things that can *be* the value for the property
- **If a CHRIS has Permitted Values defined, they are represented as an enumerated list of range classes**
 - Represents that there is a specific view of the range of the property



HR EIW Progress: CHRIS Metrics

P&R Supplied CHRIS	CHRIS Modeled	% Total CHRIS Modeled	CHRIS Aligned to DIMHRS	# CHRIS Signed Off
328	328	100%	247	121

Legend

- P&R Supplied CHRIS:** # of CHRIS to complete (this number will fluctuate throughout project)
- CHRIS Modeled:** progress metric; # CHRIS that have been modeled
- CHRIS Aligned to DIMHRS:** progress metric; # of CHRIS processed through DIMHRS model alignment activity
- CHRIS signed Off:** completion metric: CHRIS statements that have been “signed off” by P&R as accurately aligned and modeled



HR EIW Technology Architecture Roadmap

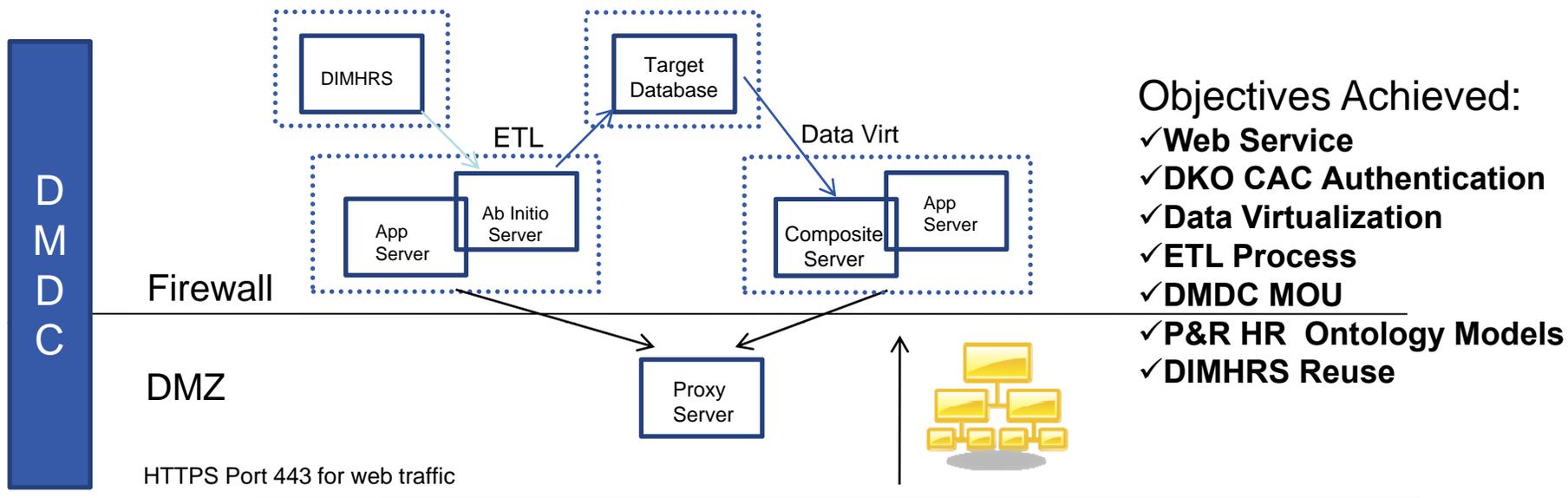
▪ **Current 2-year Schedule**

- Implement or Build Standards-Based Semantic Information Management Platform
 - ✓ Modeling PoD
 - o **RDF Warehouse**
 - o RDF Services / SPARQLizer
 - o Federated SPARQL Engine
- Implement or Build SPARQL-driven Business Intelligence (BI) Platform
- Risk Mitigation Strategy includes Implementation of Relational Warehouse (ETL)

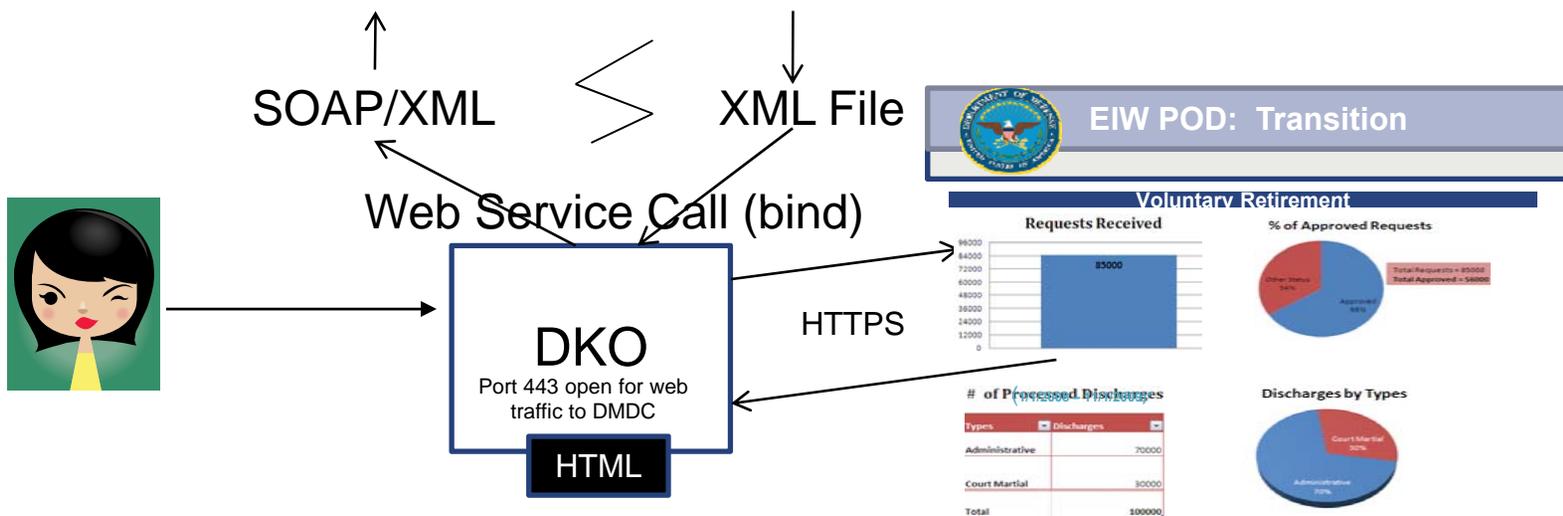
▪ **Current Plan for Years 3 & 4**

- Select, Scale, & Deploy Operational Technology

Backend PoD1 Architecture

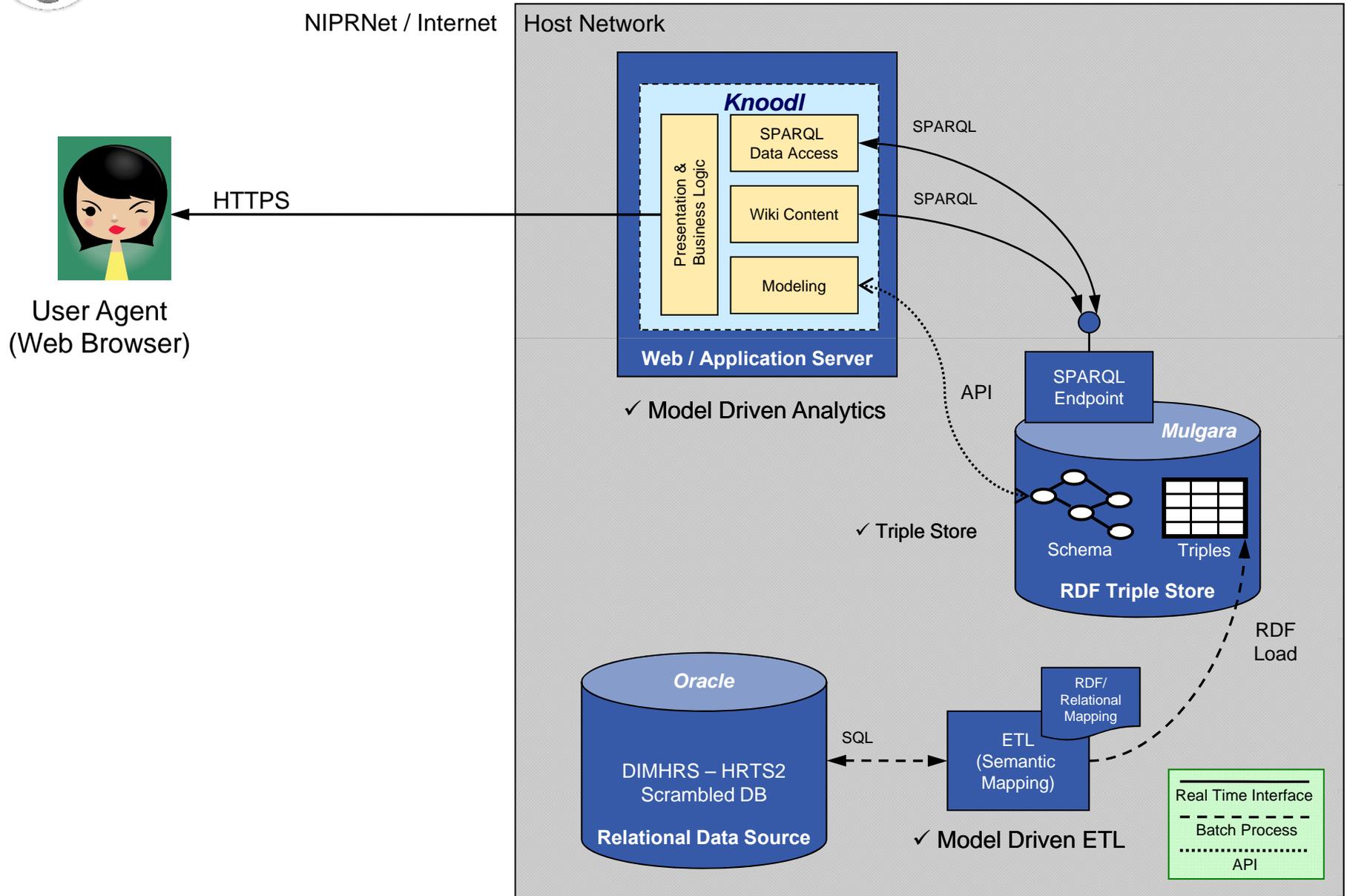


- Objectives Achieved:**
- ✓ Web Service
 - ✓ DKO CAC Authentication
 - ✓ Data Virtualization
 - ✓ ETL Process
 - ✓ DMDC MOU
 - ✓ P&R HR Ontology Models
 - ✓ DIMHRS Reuse





RDF Warehouse Architecture (POD2)

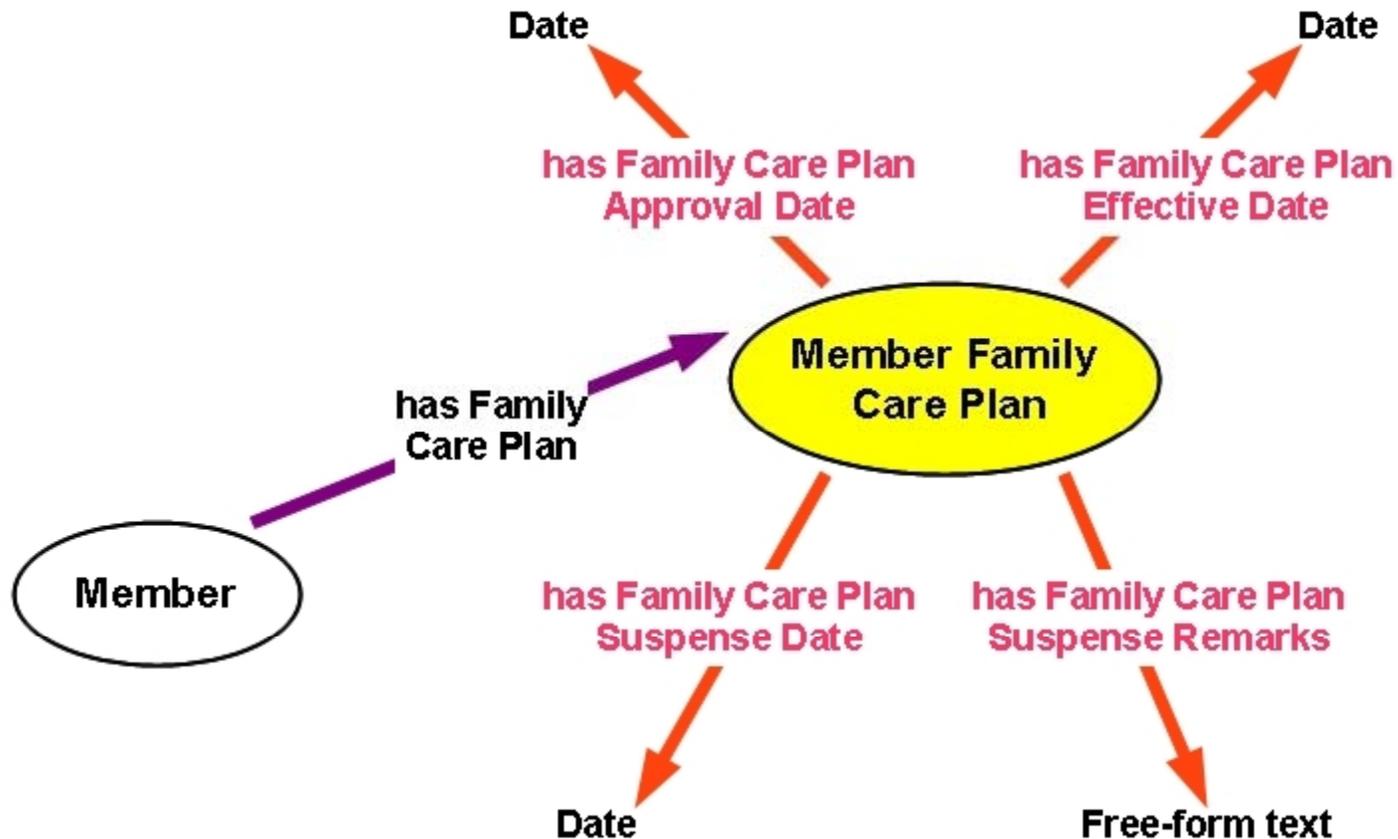




Backups

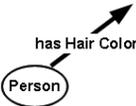
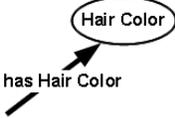
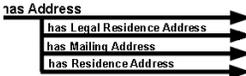


An Ontology Graph





Ontology Graph Key

	OWL Construct	Relationship to CHRIS Definition	Description	
	Class		Text inside a bubble represents a class. In this example, "Person" is the class being represented.	
	Object Property		An arrow connecting two classes represents an object property. These arrows are always purple. In this example, the property "has Hair Color" is being represented by the purple arrow.	
	Datatype Property		An arrow connecting a class to some datatype element represents a datatype property. These arrows are always orange. In this example, the property "has Social Security Number" is being represented by the orange arrow.	
	Domain		A class with an arrow pointing away from it represents the domain of a property. The domain is the class that can have a specific property. In this example, the class "Person" has the property "has Hair Color."	
	Range		A class with an arrow pointing toward it represents the range of a property. The range is the class that can be referred to by a specific property; it constrains the values of this property. In this example, the property "has Hair Color" refers to the class "Hair Color." So, the values for the property "has Hair Color" must be part of the class "Hair Color."	
	Subclass		A bubble inside another bubble represents a subclass. In this example, the class "Member" is a subclass of the class "Person."	
	Subproperty		An arrow originating from a longer arrow represents a subproperty. In this example, the property "has Address" is a superproperty, and it has subproperties "has Legal Residence Address," "has Mailing Address," and "has Residence Address."	
		A Class that has Allowed Values of a CHRIS	Green text inside a bubble represents a class that contains Allowed Values of a CHRIS. In this example, the class "Blood Type" contains allowed values of the CHRIS "Blood Type" (ie, A, AB, B, or O).	
has Blood Type		A property that represents a CHRIS	A property that appear in pink font represents a CHRIS. In this example, the property "has Blood Type" represents the CHRIS "Blood Type."	