

# Information Flow in the Federal Enterprise Redux

Governance, Information Sharing and Privacy

richard.murphy@gsa.gov

## Quote

*Systems, scientific and philosophic, come and go. Each method of limited understanding is at length exhausted. In its prime each system is a triumphant success: in its decay it is an obstructive nuisance.*

Alfred North Whitehead, Adventures of Ideas

## Purpose

1. Examine new principles of governance: The Natural Law of Federation
2. Explain information sharing based on information flow
3. Introduce policy interaction model as a next generation, executable social contract for citizen privacy

## Power and Federation

*Each State, in ratifying the Constitution, is considered as a sovereign body, independent of all others, and only to be bound by its own voluntary act. In this relation, then, the new Constitution will, if established, be a federal, and not a national constitution.*

James Madison, The Federalist Papers, #39

## The Natural Law of Federation

1. Law of Approximation: Fractal Society; Fractal Web (Tim BL)  
Our architectural patterns should reveal events from the world in which we live. Life on this planet is a Complex Adaptive System! Christopher Alexander calls this Living Structure<sup>1</sup>.

2. Law of Emergence: Discovery and adaptation precede determinism. We don't all need to agree first, agree on everything, or agree on the same thing. And we need to share information beyond our scope of control.
3. Law of Generativity: Cooperation precedes compliance. Advancement implies tolerance, independent invention, free extension, language mixing, and partial understanding.

### Application of Federation to Information Sharing

1. Supply & Demand Model for Information
2. Discovery and Description v. CUO Approach
3. Shared Concept v. Intersection or Union
4. Languages, Logics, Models and Theories

### Information Flow

*There is the Intentional Interpretant, which is a determination of the mind of the utterer; the Effectual Interpretant, which is a determination of the mind of the interpreter, and the Communicational Interpretant, or say the Cominterpretant, which is a determination of that mind into which the minds of the utterer and the interpreter have to be fused in order that any communication should take place. This mind may be called the commens. It consists of all that is and must be, well understood between utterer and interpreter, at the outset, in order that the sign in question should fulfill its function. This I proceed to explain.*

Charles Sanders Peirce, Spring 1906

### Highlights in Information Theory

1. A Mathematical Theory of Communication, Claude Shannon, 1948

---

1. The Nature of Order: An Essay on the Art of Building and the Nature of the Universe, Christopher Alexander, Center for Environmental Studies, 2002

2. Knowledge and the Flow of Information, Fred Dretske, 1979
3. Information Flow: The Logic of Distributed Systems, Barwise and Seligman, 1997
4. Information Flow Framework, Robert E. Kent, 2006

### Information Flow as Metaphor

Claim: a's being of type  $\alpha$  carries the information that b is of type  $\beta$

Example: The espionage ring tone carried the information that Ginny was the person calling Rick

### Information Flow Principles

**First Principle (P1):** IF results from regularities in a distributed system

**Second Principle (P2):** IF crucially involves both types and particulars (tokens)

**Third Principle (P3):** It is by virtue of regularities among connections that information about some components of a distributed system carries information about other components

**Fourth Principle (P4):** The regularities of a given distributed system are relative to its analysis in terms of information channels

### Developing the FEA Example ...

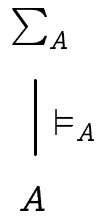
1. Consider information sharing among DIA, DHS and DOJ. Does the behavior of the members reflect federation or central authority? How does The Natural Law of Federation better inform information sharing?
2. FEA reference models provide the domain restriction on types and tokens in the distributed system. Federation members choose how they implement the reference models. (BTW - Does everyone know about FEA-RMO?)

3. Explain information sharing among members of the federation with information flow: classifications, infomorphisms, constraints, local logics and channels

## Classifications

**Definition 1.** *A classification  $A = \langle A, \Sigma_A, \models_A \rangle$  consists of a set  $A$  of objects to be classified called tokens of  $A$ , a set  $\Sigma_A$  of objects used to classify the tokens, called the types of  $A$ , and a binary relation  $\models_A$  between  $A$  and  $\Sigma_A$  that tells one which tokens are classified as being of which types.*

Figure 1 - Classification diagram



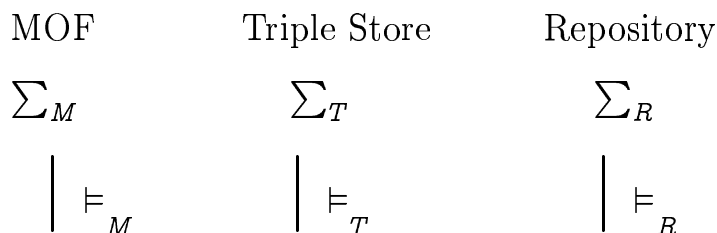
Comment: Classifications formalize what we might think of as components, modules or context. (P2)

### Developing the FEA Example ...

The classifications are a Meta Object Facility (MOF) that holds a Unified Modeling Language (UML) artifact, a triple store that holds a Web Ontology Language (OWL-DL) artifact, and a meta data repository that holds an artifact based on XML Topic Maps (XTM).

Let's say DIA has the MOF, DHS has the Triple Store, and DOJ has the XML Repository.

Figure 1 Classifications (Technical Components)

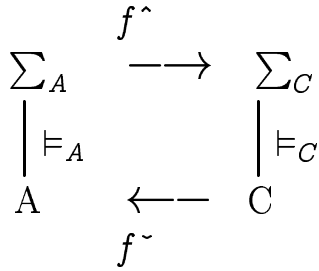


$M$  $T$  $R$ 

## Infomorphisms

**Definition 2.** If  $A = \langle A, \Sigma_A, \models_A \rangle$  and  $C = \langle C, \Sigma_C, \models_C \rangle$  are classifications then an infomorphism is a pair  $f = (f^\wedge, f^\sim)$  of functions satisfying the analogous biconditional:  $f^\sim(c) \models_A \alpha$  iff  $c \models_C f^\wedge(\alpha)$  for all tokens  $c$  of  $C$  and all types  $\alpha$  of  $A$ . An infomorphism is represented concisely as  $f: A \rightleftarrows C$

Figure 2 - Infomorphism Diagram



Comment: DIA, DHS, and DOJ need to share, or morph, their information while preserving meaning. We can think of infomorphisms as interpretation. (P3) Information provenance preserves trust.

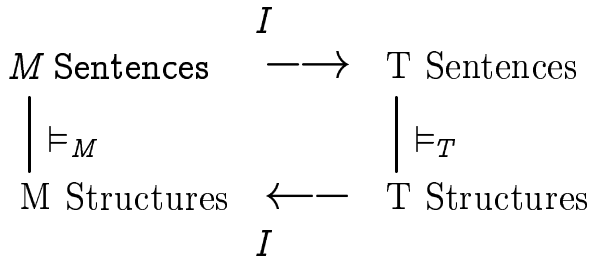
## Constraints

**Definition 3.** Let  $A$  be a classification and let  $\langle \Gamma, \Delta \rangle$  be a sequent of  $A$ . A token  $a$  of  $A$  satisfies  $\langle \Gamma, \Delta \rangle$  provided that if  $a$  is of type  $\alpha$  for every  $\alpha \in \Gamma$  then  $a$  is of type  $\alpha$  for some  $\alpha \in \Delta$ . We say that  $\Gamma$  entails  $\Delta$  in  $A$ , written  $\Gamma \vdash_A \Delta$ , if every token  $a$  of  $A$  satisfies  $\langle \Gamma, \Delta \rangle$ . If  $\Gamma \vdash_A \Delta$  then the pair  $\langle \Gamma, \Delta \rangle$  is called a constraint supported by the classification  $A$ .

Comment: Constraints express the regularities that allow information to flow. (P1) Unfortunately, our federation is irregular in its use of languages, logics, models and theories.

## Infomorphism as Structure Preserving Transformation

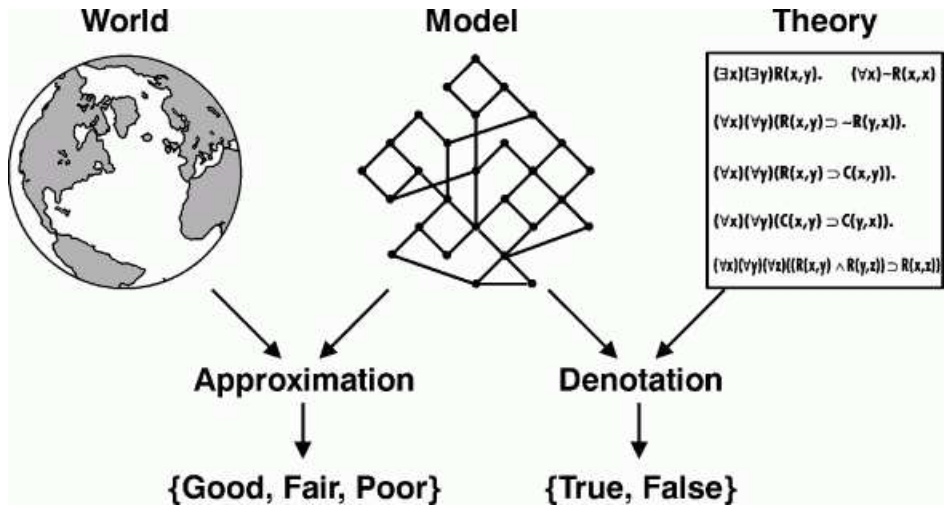
The M Sentences represent UML Class, Association, and Class; the T Sentences represent OWL Class, Property, Class



Comment: Assume a structural constraint on the languages. We lose decidability by going from ALHOIN(D) to OWL-Full. The constraint remains, but the models and theories differ.

### Models and Theories

Models are approximations of the world and theories are precise executable representations of models. We use tools like Specware and SNARK to provide assurance our theories are correct.



John Sowa, Signs Processes and Language Games

Local Logics

**Definition 4.** A local logic  $\mathcal{L} = \langle A, \vdash_{\mathcal{L}}, N_{\mathcal{L}} \rangle$  consists of a classification  $A$ , a set  $\vdash_{\mathcal{L}}$  of sequents (satisfying certain structural rules) involving the types of  $A$ , called the constraints of  $\mathcal{L}$ , and a subset  $N_{\mathcal{L}} \subseteq A$ , called the normal tokens of  $\mathcal{L}$ , which satisfy all the constraints of  $\vdash_{\mathcal{L}}$ . A local logic is sound if every token is normal; it is complete if every sequent that holds of all normal tokens is in the consequence relation  $\vdash_{\mathcal{L}}$ .

Comment: We can have sins of commission (soundness) and omission (completeness). Let's use Description Logics to illustrate the need for semantic preserving transformations.

### Description Logic Expressiveness<sup>2</sup>

Nomenclature	Meaning	DL-Core SHIN(D)	UML ALHOIN(D)	OWL-DL SHOIN(D)	ER ALN(D)	Topic Maps AL(D)
AL	Atomic Concept Universal Concept Bottom Concept Atomic Negation Intersection Value Restriction Limited Existential Quantification	x	Atomic Concept Value Restriction Limited Existential Quantification (AL-)	x	Atomic Concept Value Restriction Limited Existential Quantification (AL-)	Atomic Concept Value Restriction (AL - -)
C	Full Negation or Complement	X		X		
E	Full existential Quantification	x		x		
H	Role Hierarchies	x	x	x		
I	Inverse Roles	x	x	x		
N	Unqualified Number Restrictions	x	x	x	x	
O	Enumerated Classes			x		
R*	Transitive Roles	x		x		
V	Union Constructor	x		x		
(D)	Datatypes	x	x	x	x	x

2. "A Description Logic for Use as the ODM Core," Lewis Hart and Patrick Emery, EDOC 2004

## Infomorphism as Semantic (semi) Preserving Transformation

Assume the M Sentences represent UML Class, Association, and Class; the T Sentences represent OWL Class and property and the T Structures represent the OWL individuals

$$\begin{array}{ccc}
 & I & \\
 M \text{ Sentences} & \xrightarrow{\quad} & T \text{ Sentences} \\
 \left| \vDash_M \right. & & \left| \vDash_T \right. \\
 M \text{ Structures} & \xleftarrow{\quad} & T \text{ Structures} \\
 & I &
 \end{array}$$

Comment: We need a structure that classifies according to OWL-DL, not OWL Full, but targeting the local logic requires additional information we didn't have at design time. We can think of this as changing the channel or creating living structure.

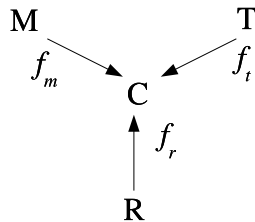
### Channel Theory

**Definition 5.** *An information channel consists of an indexed family  $C = \{ f_i : A_i \rightleftarrows C \}_{i \in I}$  of infomorphisms with a common codomain  $C$ , called the core of the channel.*

Comment: Consider *the commens* in the Peirce quote above. What we typically think of as COSMO is the ontology that defines the classification at the core of the channel. (P4) Debates centered around defining the union are an obstructive nuisance.

### Developing the FEA Example ...

Figure 2 - Classifications and the Channel in the Federation



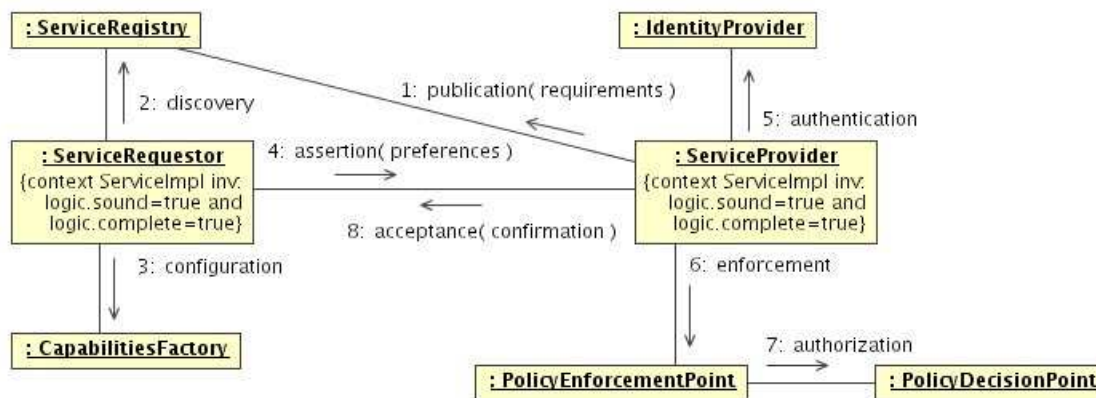
Summary: In the static representation, the information that can flow through the channel are the normal tokens that satisfy all the constraints of the system. (P4) With living structure enabled by semantic preserving transformation, DIA, DHS, and DOJ can more effectively share information.

### Policy, Privacy and the Social Contract

*The reason why men enter into society, is the preservation of their property; and the end why they chuse and authorize a legislative, is, that there may be laws made, and rules set, as guards and fences to the properties of all the members of the society, to limit the power, and moderate the dominion, of every part and member of the society.*

John Locke, Second Treatise on Government

### Policy Interaction Model



Comment: Consider Information Flow in the context of SOA. PIM is a next generation, executable social contract for the collection, use, maintenance, and disposition of private information and general policy conformance. First presented at TAMI/Portia Workshop, Cambridge, MA 2006

### Citizen Privacy Service ( A PDP )

Provides for inferencing allowed disclosures based on disclosure target, authorization, and intent as well as inferencing on denied requests and disclosure accounting.

US Privacy Act of 1974 in OWL-DL (sound and complete)

Public interface (GSA's WSDL); private implementation (TBD)

Open Source artifacts available here: <http://www.osera.gov>

### Information Provenance: Proof and Trust

Trust is based on information provenance including FOL proofs

JTP (Stanford) is a modular reasoning system that generated proofs in Proof Markup Language (PML) and Proof Protocol for Deductive Reasoning (PPDR).

Inference Web ([iw.stanford.edu](http://iw.stanford.edu)) includes a discoverable registry of automated reasoners of varying capabilities to meet our trust requirements.

### US Privacy Act IW-Proof Generation

How to we know a Congressional Request is a Granted Request?

```
jtp>load-kb;http://www.osera.gov/privacy.owl
```

```
jtp>ask;(holds rdfs:subClassOf privacy.owl:CongressionalRequest privacy.owl:GrantedRequest)
```

```
<Modus Ponens Inference Step: (holds rdfs:subClassOf privacy.owl:CongressionalRequest privacy.owl:GrantedRequest)
```

```
:- (instance-of privacy.owl:CongressionalRequest frame)
```

```
(:has-slot privacy.owl:CongressionalRequest rdfs:subClassOf) (holds rdfs:subClassOf privacy.owl:CongressionalRequest privacy.owl:GrantedRequest)
```

```
Inference Bindings: ?slot : ...
```

```
{{ [/1]: (instance-of privacy.owl:CongressionalRequest frame)
```

```
« Direct Assertion Inference Bindings: ?subj: ...
```

### Recap

1. The Natural Law of Federation better informs our approach to information sharing

2. Information Flow formalizes our approach to information sharing
3. The policy interaction model, enhanced by information provenance, ensures privacy for sharing information