Secure Knowledge Management

Dr. Bhavani Thuraisingham
The National Science Foundation

September 2004

Outline

0 Background on Knowledge Management

0 Aspects of Secure Knowledge Management

0 Secure Knowledge Management Technologies

- E.g., Secure Semantic Web

0 Secure Knowledge Management Strategies, Processes and Metrics

0 Note on Privacy

0 Directions

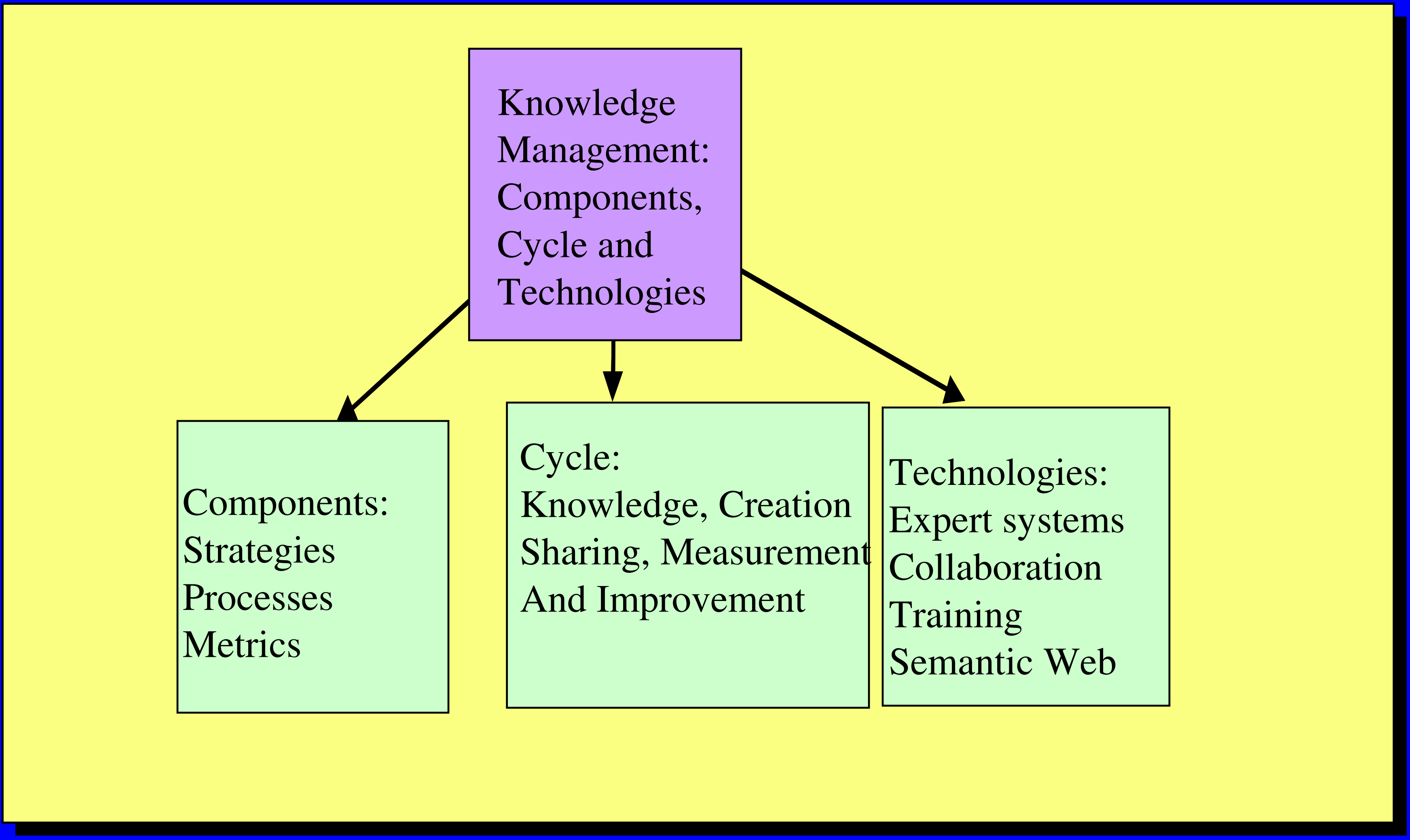
What is Knowledge Management?

0 Knowledge management, or KM, is the process through which organizations generate value from their intellectual property and knowledge-based assets

0 KM involves the creation, dissemination, and utilization of knowledge

0 Reference: http://www.commerce-database.com/knowledge-management.htm?source=google

Aspects of Knowledge Management



Knowledge Management: Components, Cycle and Technologies

Components:
Strategies
Processes
Metrics

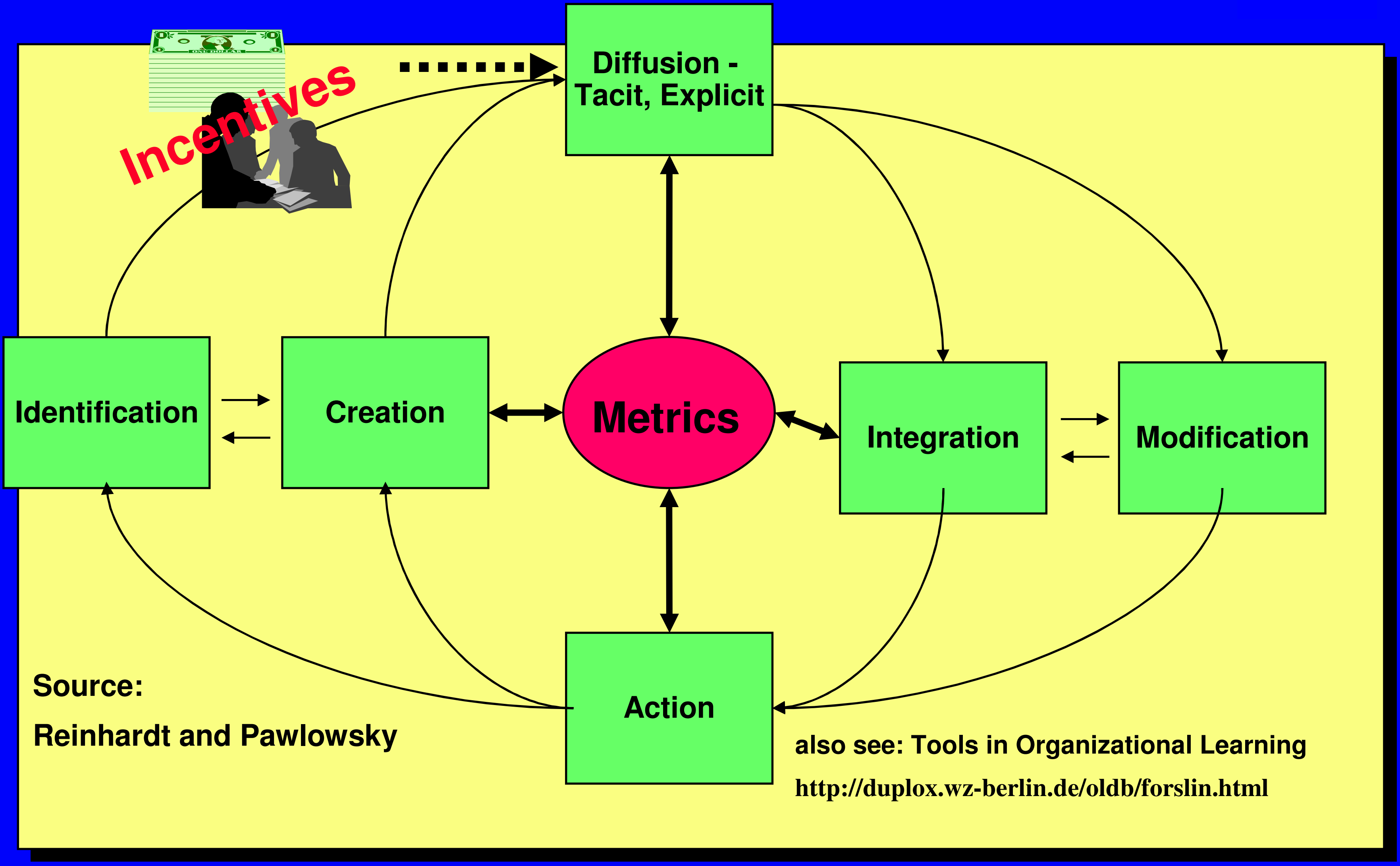
Cycle:
Knowledge, Creation
Sharing, Measurement
And Improvement

Technologies:
Expert systems
Collaboration
Training
Semantic Web

Knowledge Models

- **Level 1: Highest Level**
  - **Mental models utilized by psychologists**
  - **Social models (e.g. social network models) used by sociologists**
- **Level 2: Mid-level**
  - **Models utilized by expert systems**
  - **Process modeling**
- **Level: Bottom level**
  - **Models understood by machines**
  - **E.g., rule-based, frame-based, etc.**

Organizational Learning Process



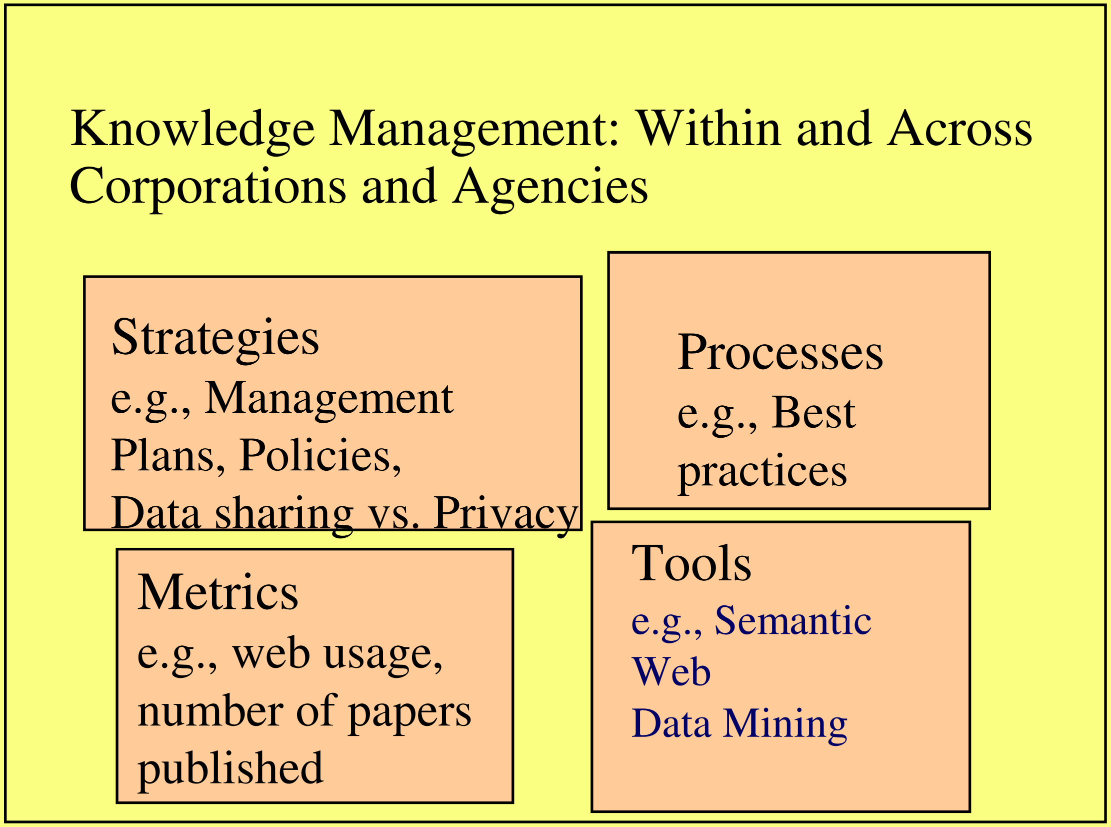
Source:

Reinhardt and Pawlowsky

also see: Tools in Organizational Learning

http://duplox.wz-berlin.de/oldb/forslin.html

Knowledge Management: Strategies, Processes, Metrics and Tools



Knowledge Management: Within and Across Corporations and Agencies

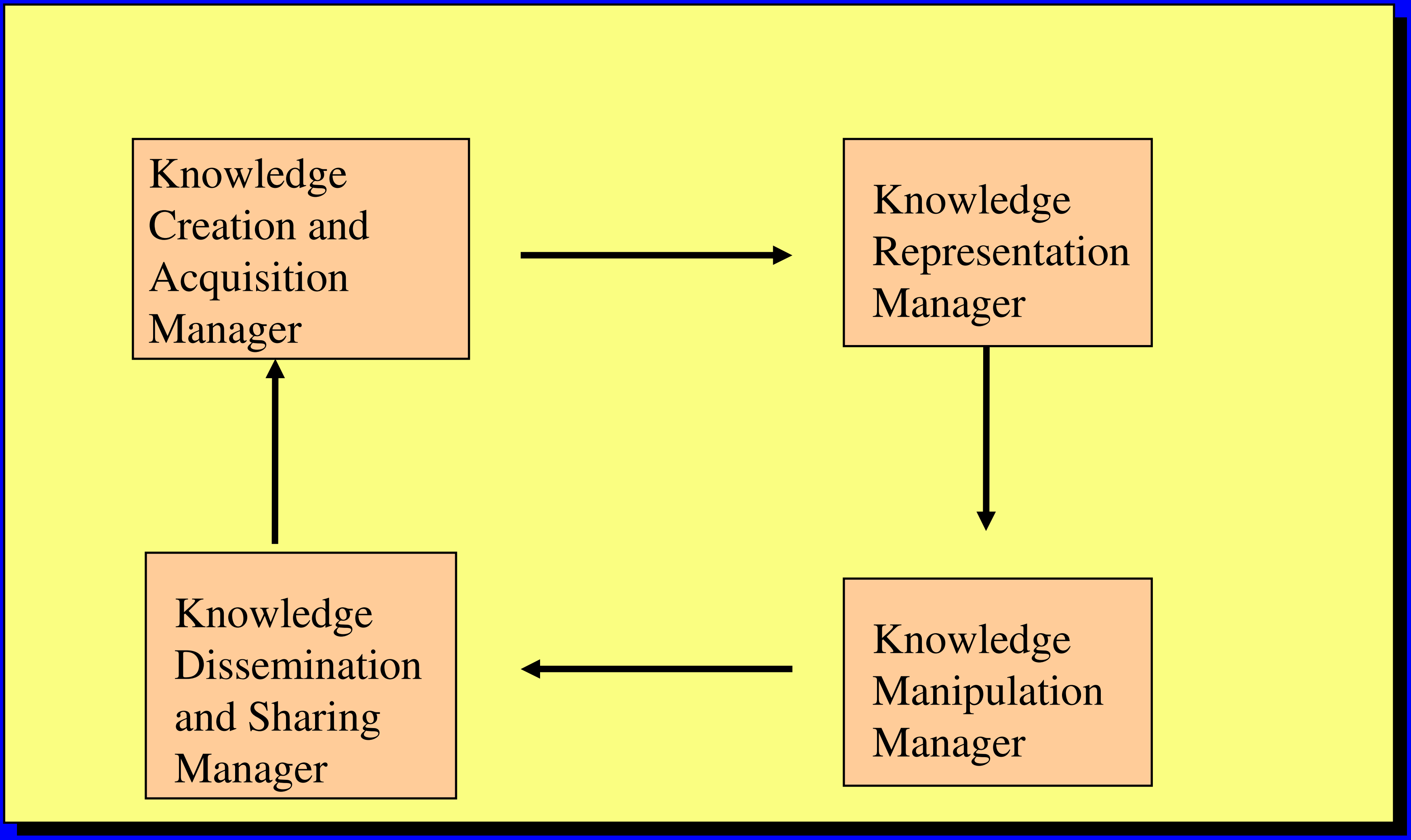
Strategies
e.g., Management Plans, Policies, Data sharing vs. Privacy

Processes
e.g., Best practices

Metrics
e.g., web usage, number of papers published

Tools
e.g., Semantic Web
Data Mining

Knowledge Management Architecture



Knowledge Creation and Acquisition Manager

Knowledge Representation Manager

Knowledge Manipulation Manager

Knowledge Dissemination and Sharing Manager

Secure Knowledge Management

0 **Security for Organizational Learning**

0 **Secure Knowledge Management Architecture**

0 **Protecting the intellectual property of an organization**

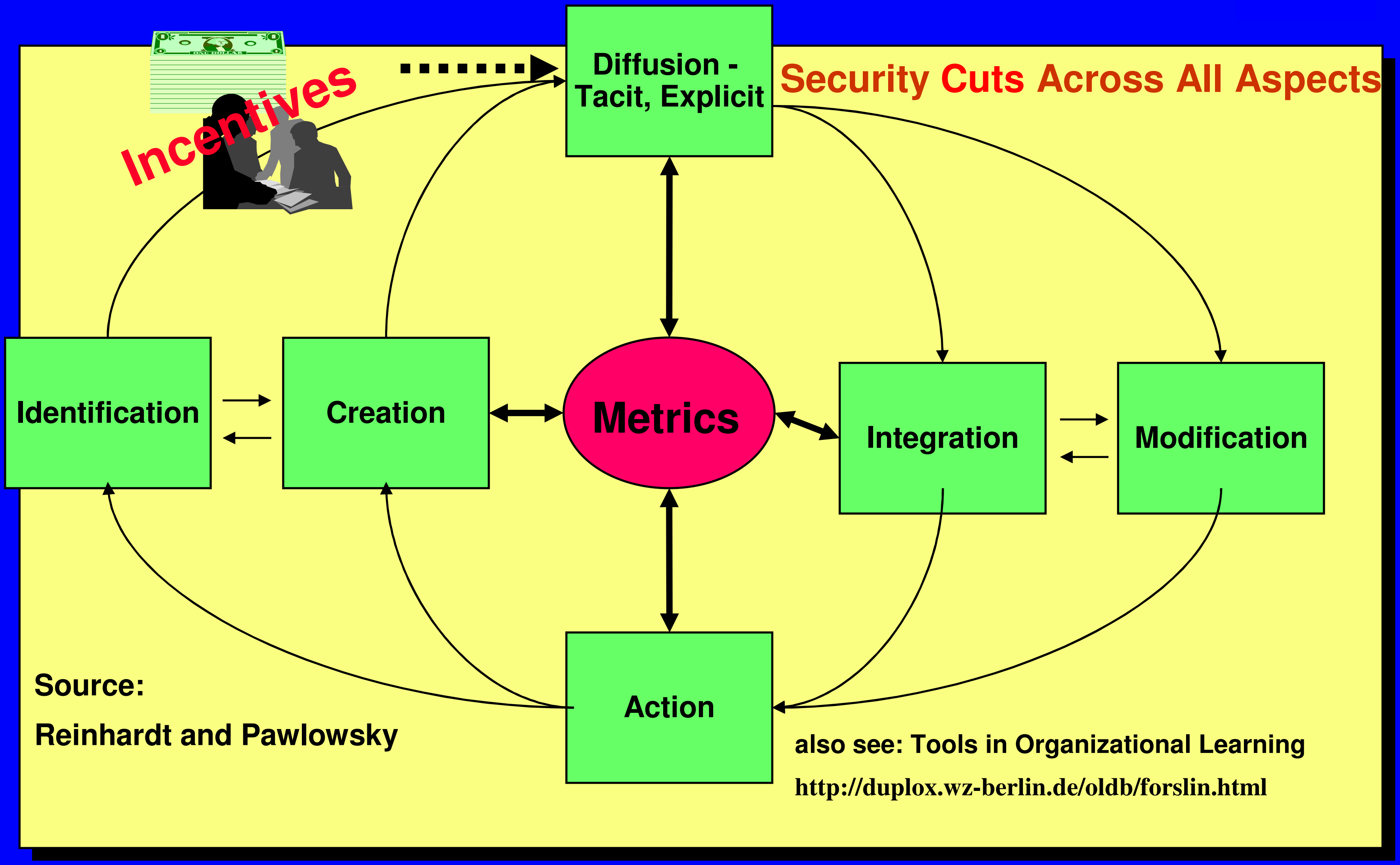
0 **Integrating security policies across organizations**

0 **Security for knowledge management: strategies, processes and metrics**

0 **Trust Management and Negotiation**

0 **Secure knowledge management technologies**

Secure Organizational Learning Process



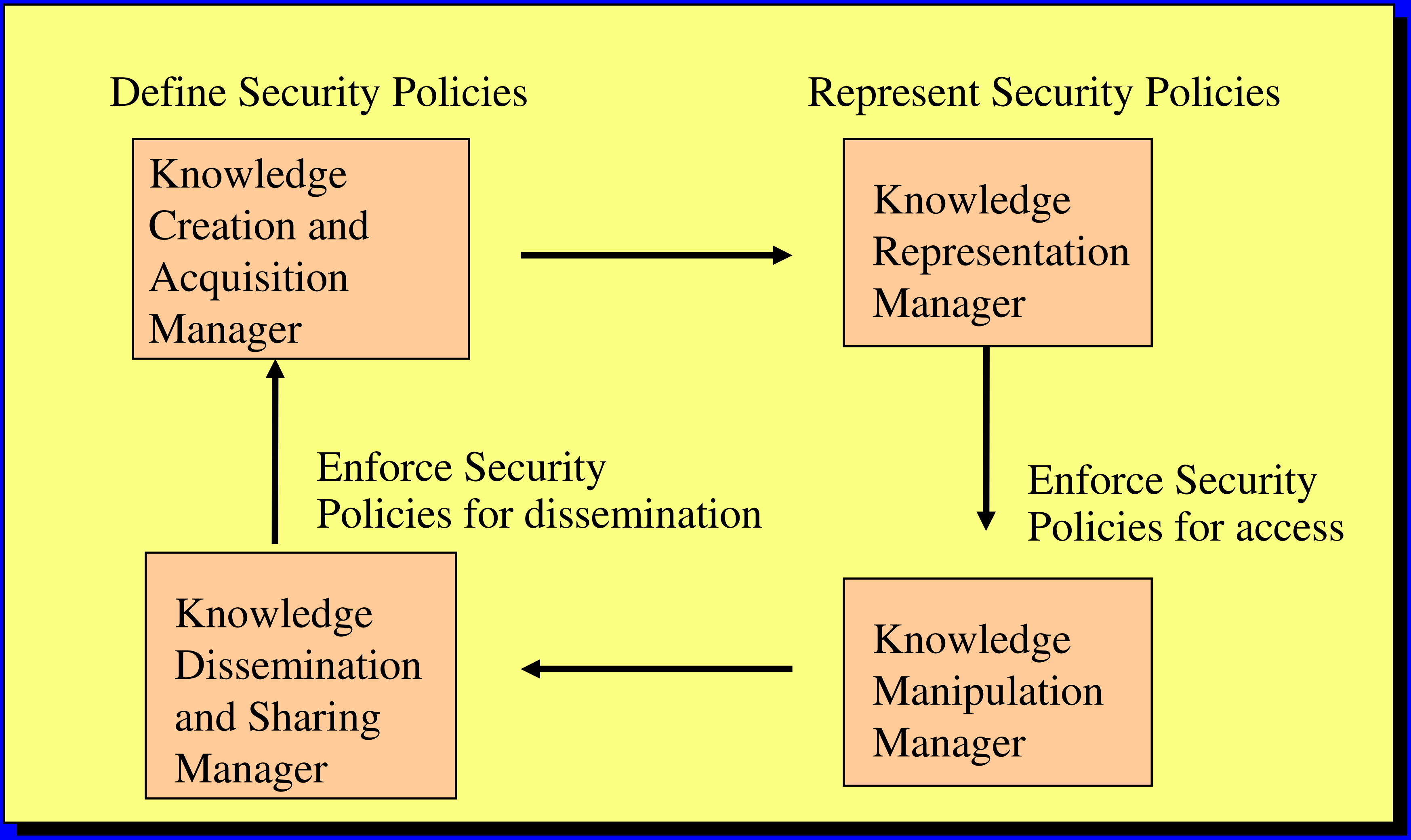
Source:

Reinhardt and Pawlowsky

also see: Tools in Organizational Learning

http://duplox.wz-berlin.de/oldb/forslin.html

Secure Knowledge Management Architecture



Define Security Policies

Represent Security Policies

Knowledge Creation and Acquisition Manager

Knowledge Representation Manager

Enforce Security Policies for dissemination

Enforce Security Policies for access

Knowledge Dissemination and Sharing Manager

Knowledge Manipulation Manager

Protecting Intellectual Property

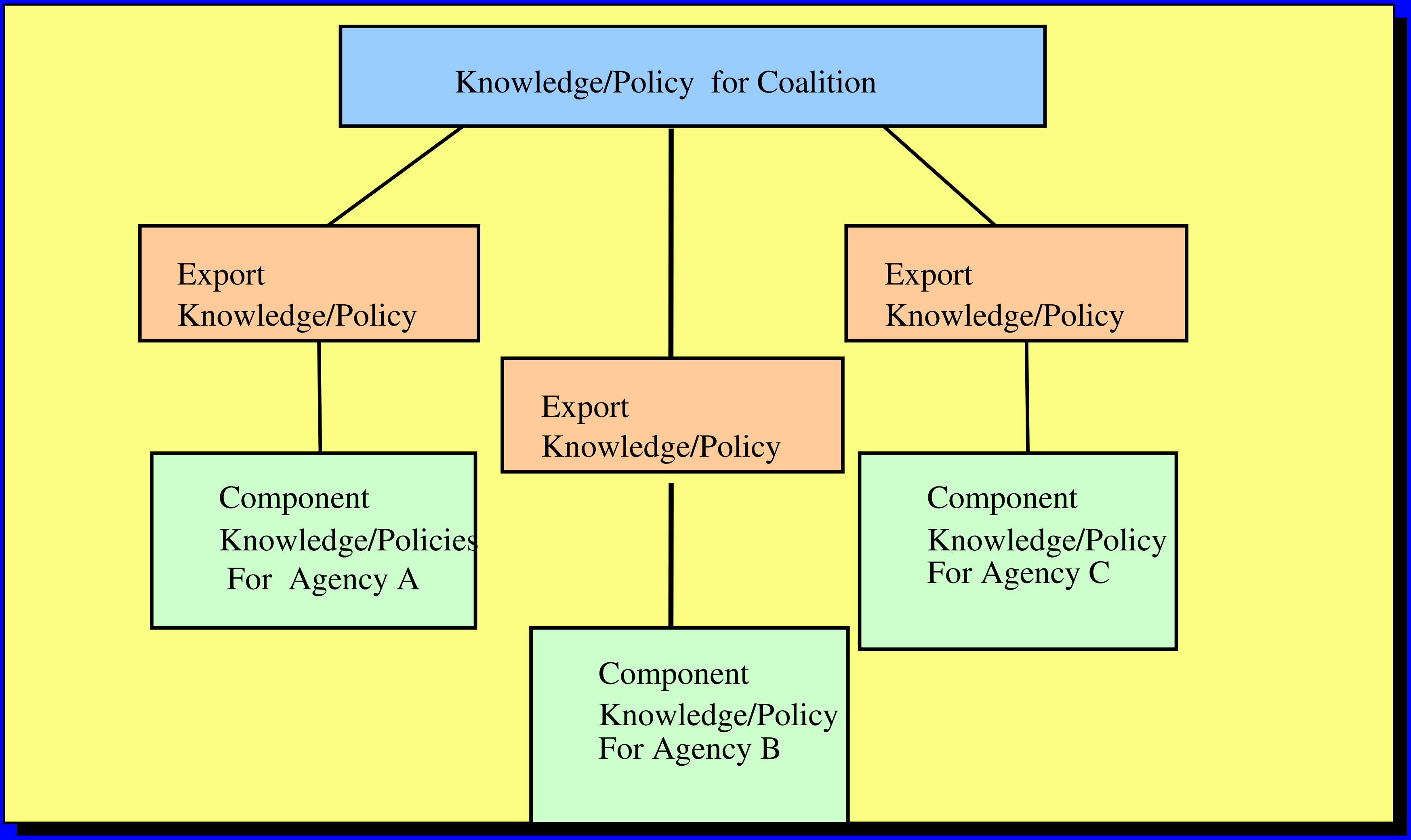
0 How can the assets of an organization be protected from competitors?

- Role-based access control policies
- Usage control policies
- Trust models
- Secure workflow and process management

0 What are the disadvantages of not openly sharing the intellectual property?

- Need an economic analysis study and experimentation

Integrating Policies: Knowledge Management for Coalitions



Knowledge/Policy for Coalition

Export Knowledge/Policy

Export Knowledge/Policy

Export Knowledge/Policy

Component Knowledge/Policies For Agency A

Component Knowledge/Policy For Agency B

Component Knowledge/Policy For Agency C

Secure Knowledge Management: Strategy, Process and Metrics

0 Strategy

- Motivation for knowledge management and how to structure a knowledge management program
- Incorporate security into the planning process

0 Process

- Use of Knowledge management to make existing practice more effective
- How can security be incorporated into the existing practices?

0 Metrics

- Measure the impact of knowledge management on an organization
- What is the impact of security on the metrics? Example, how does role-based access control affect the management of intellectual property?

Trust Management and Negotiation

0 Design a Trust Model

- Investigate the current trust models. Identify the inadequacies of current trust models and design a model for the organization

    = Components include trust management, trust negotiation as well as economic tradeoffs

0 Design a Language for specifying Trust policies

- Start with languages such as XML, RDF and Web Rules language and incorporate features for trust management and negotiation

0 Design and develop techniques for enforcing the trust policies

- Automated Trust Negotiation: A attempts to access database D based on access control policies; However before A can access D, triggers go off and owner of D exchanges credential information with A

0 Related Issues:

- Digital Rights Management, Content Management, - - - -

Secure Knowledge Management Technologies

0 Integration of Multiple Secure Data and Applications Technologies to provide security for Real and Virtual Organizations

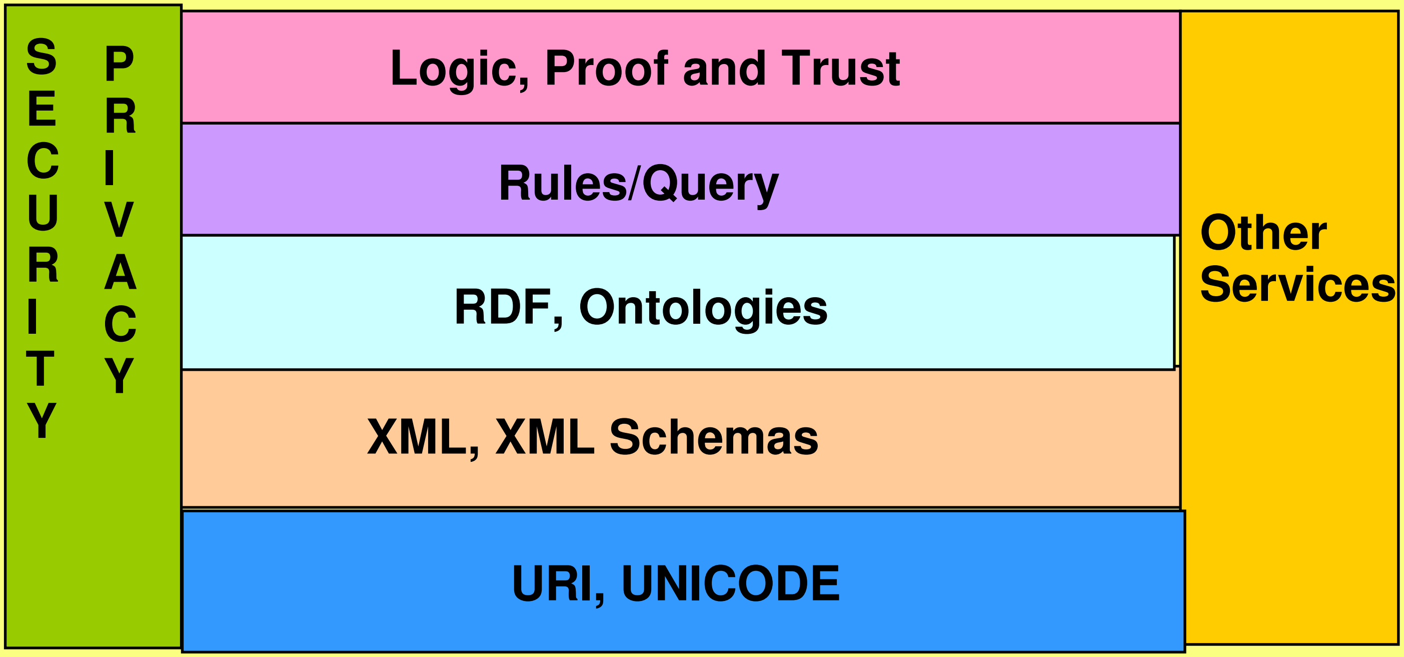
0 Examples technologies Include:

- Semantic Web
- Data Mining
- Supply Chain Management
- Multimedia information management
- - - - - - - - -

0 Integrating the Multiple Technologies is a major challenge; also we need to determine how to adapt these technologies for Real and Virtual Organizations?

Layered Architecture for Dependable Semantic Web

0Adapted from Tim Berners Lee’s description of the Semantic Web



SECURITY

PRIVACY

Logic, Proof and Trust

Rules/Query

RDF, Ontologies

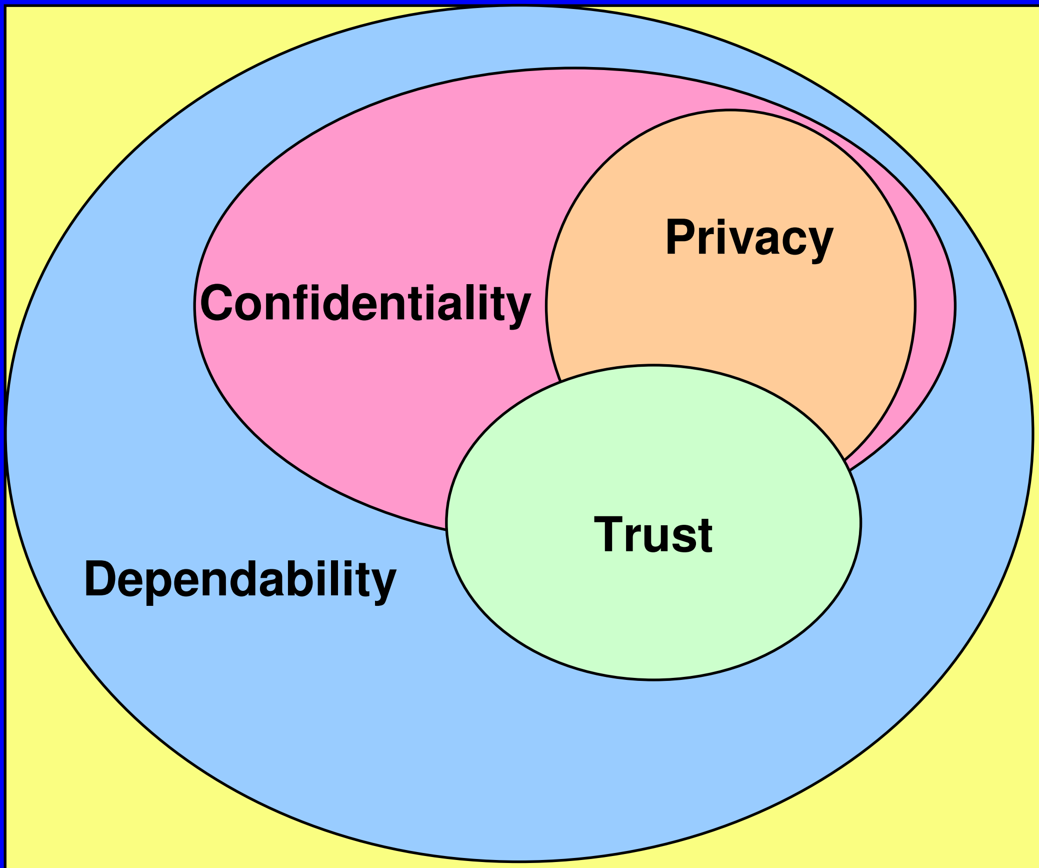
XML, XML Schemas

URI, UNICODE

Other Services

0 Some Challenges: Interoperability between Layers; Security and Privacy cut across all layers; Integration of Services; Composability

Relationships between Dependability, Confidentiality, Privacy, Trust



Dependability: Security, Privacy, Trust, Real-time Processing, Fault Tolerance; also sometimes referred to as “Trustworthiness”

Confidentiality: Preventing the release of unauthorized information considered sensitive

Privacy: Preventing the release of unauthorized information about individuals considered sensitive

Trust: Confidence one has that an individual will give him/her correct information or an individual will protect sensitive information

Directions and Challenges for Securing the Semantic Web

- Secure Web Database Management and Secure Web Services are critical technologies for securing the semantic web
- Steps for Securing the Semantic Web
- Security and Ontologies
- XML Security for Securing the Semantic Web

Secure Web Database Management

0 **Secure web data management issues include:**

  - **Extending traditional security mechanisms for web databases**

    = **Integrating security policies**

    = **Secure query, indexing and transaction management strategies**

    = **Security impact for integrating heterogeneous databases**

    = **Access control mechanisms**

  - **Security specific for the web**

    = **Security for unstructured databases such as multimedia, XML and RDF documents**

    = **Security impact on Ontology management**

    = **Privacy violations due to data mining**

0 **Many developments reported in recent IFIP 11.3 Database Security Conference Proceedings**

Secure Web Services

0 How authenticity, confidentiality and integrity can be ensured in the presence of an untrusted UDDI?

0 Traditional techniques are not enough!

0 Possible solutions:

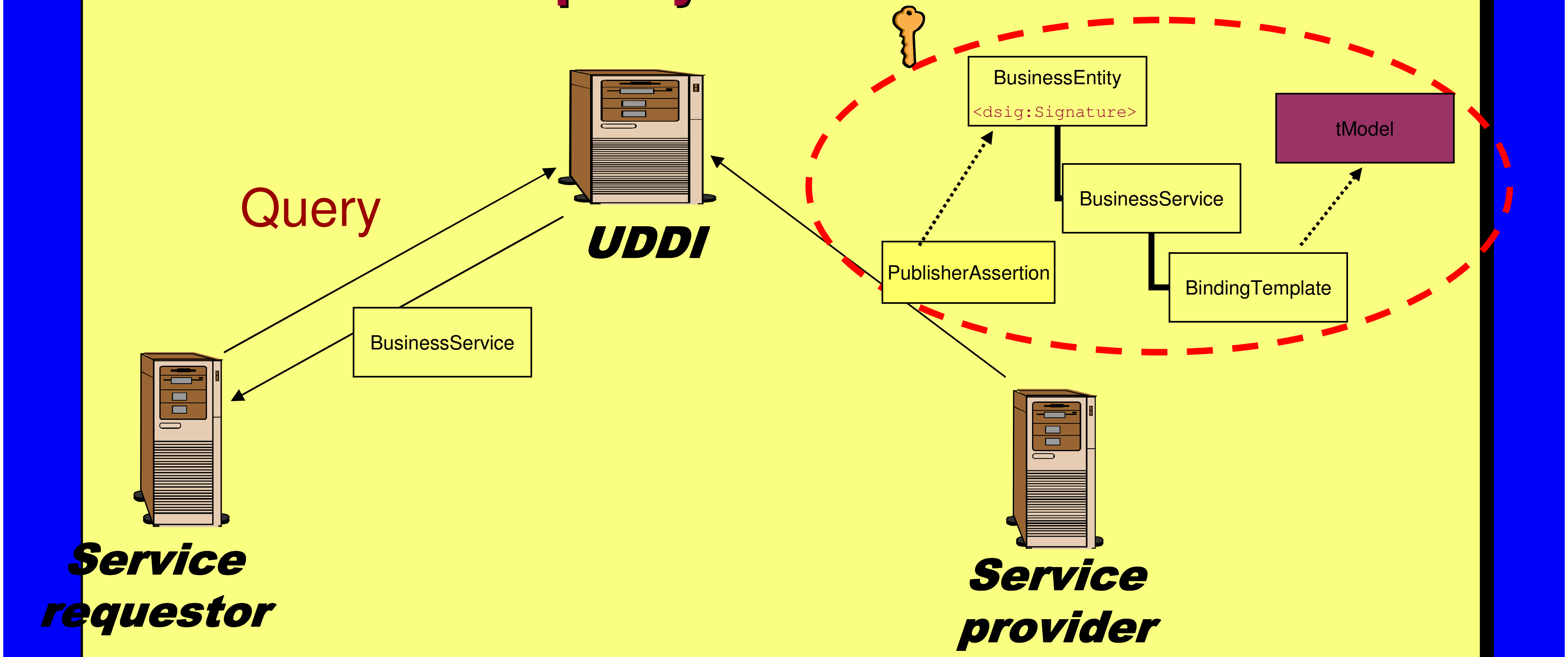
- Integrity, confidentiality: selective encryption of the data managed by the UDDI according to the specified access control policies
- Authenticity: Merkle hash trees

0 Additional security properties:

- Completeness
- Consistency

Authenticity

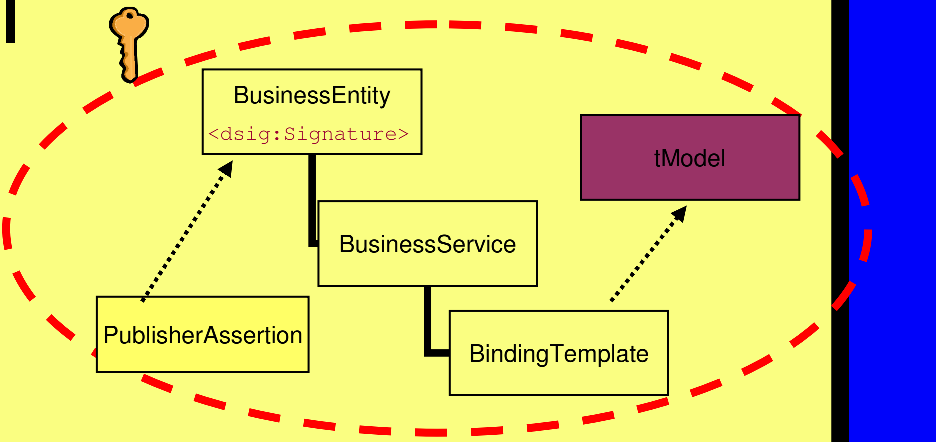
....traditional digital signatures do not fit well in third-party architectures!!



Merkle Signature

- An alternative way to sign an XML doc

- By applying a unique digital signature on an XML doc it is possible to ensure the authenticity of:
  - the whole document
  - any portion of it



- It uses a different way to compute the digest of XML docs, based on the Merkle tree authentication mechanisms

Steps to Securing the Semantic Web

0 **Flexible Security Policy**

- **One that can adapt to changing situations and requirements**

0 **Security Model**

- **Access Control, Role-based security, Nonrepudiation, Authentication**

0 **Security Architecture and Design**

- **Examine architectures for semantic web and identify security critical components**

Steps to Securing the Semantic Web (Concluded)

0 **Securing the Layers of the Semantic Web**

- **Secure agents, XML security, RDF security, secure semantic interoperabiolity, security properties for ontologies, Security issues for digital rights**
- **Much of the research is focusing on XML security; Next step is securing RDF documents**

0 **Challenge: How do you integrate across the layers of the Semantic Web and preserve security?**

Security and Ontologies

0 **Access control for Ontologies**

- **Who can access which parts of the Ontologies**
- **E.g, Professor can access all patents of the department while the Secretary can access only the descriptions of the patents in the patent ontology**
- **Can we apply the research on secure metadata management for secure ontology management?**

0 **Ontologies for Security Applications**

- **Use ontologies for specifying security policies**
- **Integrating heterogeneous policies may involve integrating ontologies and resolving inconsistencies**

XML Security

0 Some ideas have evolved from research in secure multimedia/object data management

0 Access control and authorization models

- Protecting entire documents, parts of documents, propagations of access control privileges; Protecting DTDs vs Document instances; Secure XML Schemas

0 Update Policies and Dissemination Policies

0 Secure publishing of XML documents

- How do you minimize trust for third party publication

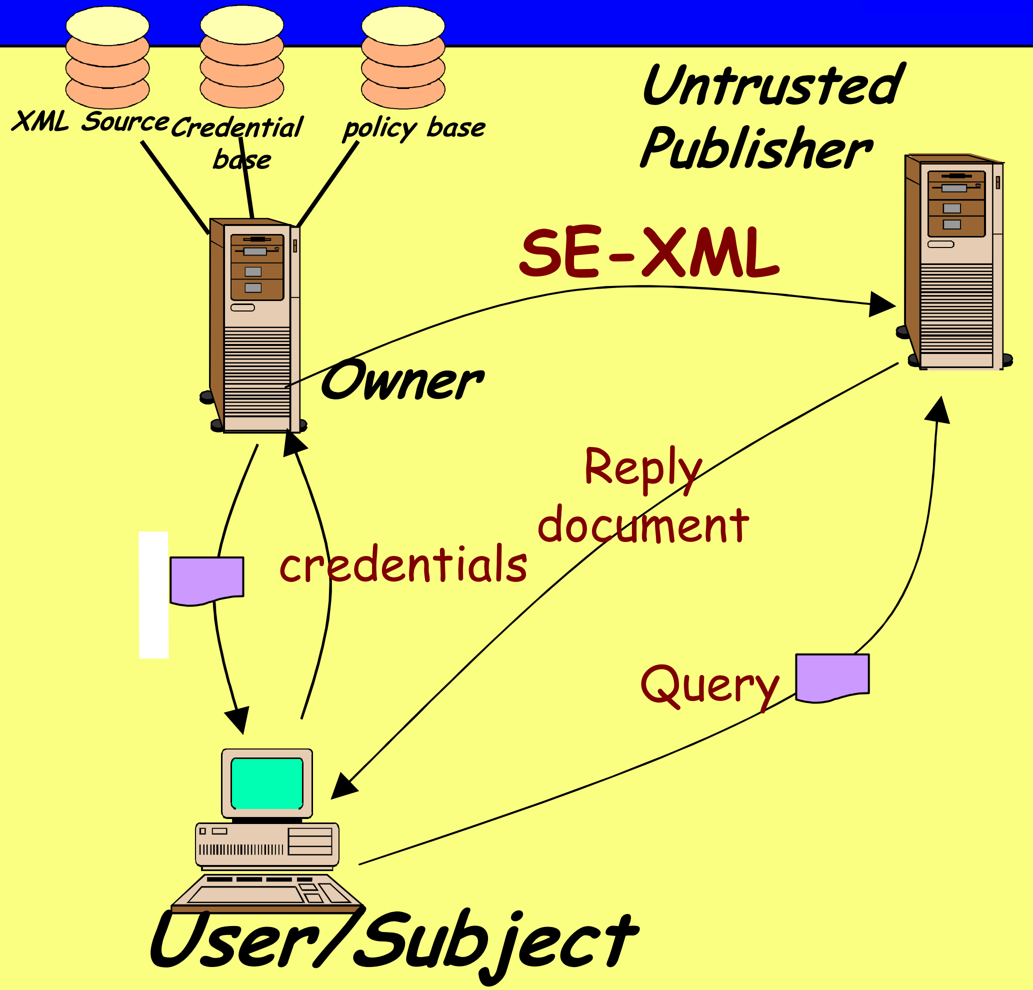
0 Use of Encryption

0 Inference problem for XML documents

- Portions of documents taken together could be sensitive, individually not sensitive

Third-Party Architecture

- The Owner is the producer of information. It specifies access control policies
- The Publisher is responsible for managing (a portion of) the Owner information and answering subject queries
- Goal: Untrusted Publisher



What are the Next Steps and Challenges for Secure Semantic Web? - I

0 We need to continue with XML security research as well as work with standards

- W3C standards are advancing rapidly; security research, prototypes and products must keep up with the developments
- Researchers, vendors and standards organizations must work together

0 Secure XML DBMSs (query, transactions, storage, - - -)

0 RDF Security

- When you bring in semantics, many challenges for security
- Need to develop security models for RDF documents

0 Secure Ontologies

- Two aspects; one is to develop protection models for Ontology databases; other is to use ontologies for ensuring security and privacy

What are the Next Steps and Challenges for Secure Semantic Web? - II

0 **Secure semantic interoperability**

- **What can we learn from secure database interoperability and federated databases?**

0 **Trust and digital rights management**

- **How do you trust the contents of a document? How do you pass digital rights when documents are disseminated?**

0 **Security for domain specific semantic webs**

- **Do we need multiple security policies and models?**

0 **Secure interoperability across the layers of the semantic web**

- **This will be a major challenge even when security is not being considered**
- **Security has to be considered in the beginning**

Data Mining as a Threat to Privacy

0 Data mining gives us “facts” that are not obvious to human analysts of the data

0 Can general trends across individuals be determined without revealing information about individuals?

0 Possible threats:

  - Combine collections of data and infer information that is private

    = Disease information from prescription data

    = Military Action from Pizza delivery to pentagon

0 Need to protect the associations and correlations between the data that are sensitive or private

Some Privacy Problems and Potential Solutions

0 **Problem: Privacy violations that result due to data mining**

- **Potential solution: Privacy-preserving data mining**

0 **Problem: Privacy violations that result due to the Inference problem**

- **Inference is the process of deducing sensitive information from the legitimate responses received to user queries**
- **Potential solution: Privacy Constraint Processing**

0 **Problem: Privacy violations due to un-encrypted data**

- **Potential solution: Encryption at different levels**

0 **Problem: Privacy violation due to poor system design**

- **Potential solution: Develop methodology for designing privacy-enhanced systems**

Privacy Preserving Data Mining

0 Prevent useful results from mining

- Introduce “cover stories” to give “false” results
- Only make a sample of data available so that an adversary is unable to come up with useful rules and predictive functions

0 Randomization

- Introduce random values into the data and/or results without significantly affecting the data mining results
- Give range of values for results instead of exact values

0 Perturbation

- Perturb the data in such a way that one can still carry out useful mining but not cover the exact sensitive correlations

0 Secure Multi-party Computation

- Each party knows its own inputs; encryption techniques used to compute final results

Privacy for Web Services

0 W3C Web Services Architecture Requirements:

- the WSA must enable privacy policy statements to be expressed about web services
- advertised web service privacy policies must be expressed in P3P
- the WSA must enable a consumer to access a web service’s advertised privacy policy statement
- the WSA must enable delegation and propagation of privacy policy
- web services must not be precluded from supporting interactions where one or more parties of the interaction are anonymous

0 Can we handle privacy policies similar to the way we are handing security policies?

Status and Directions

- Knowledge management has exploded due to the web
- Knowledge Management has different dimensions
  - Technology, Business
  - Goal is to take advantage of knowledge in a corporation for reuse
  - Objects will play a key role in technology
- Tools are emerging
- Need effective partnerships between business leaders, technologists and policy makers
- Incorporating security is the next big challenge
  - How much value are we losing by incorporating security into knowledge management/sharing practices?

Other Ideas and Directions?

0 **Please contact**

- **Dr. Bhavani Thuraisingham**
The National Science Foundation
Suite 1115
4201 Wilson Blvd
Arlington, VA 22230
Phone: 703-292-8930
Fax 703-292-9073
email: bthurais@nsf.gov
- **Alternate email: thura@mitre.org**