Do You Know Where Your Information Is? Information Asset Exploitation Across Large-Scale Distributed Enterprises

Sekar Chandersekaran, Institute for Defense Analyses

Andrew Trice, Institute for Defense Analyses

12th ICCRTS, June 2007

Outline

Challenges of Information Discovery

Information Discovery Example

Enterprise View and Overall Environment

Scenario Analysis

Emerging Solution for Information Discovery

Conclusions

Challenges of Information Discovery

That "old familiar tune":

Many systems, data repositories, documents…

Don't know or can't find what we have

Infrastructure, access policies aren't keeping up

COIs are supposed to help, but use different vocabularies

Example: Deployment Readiness

Mission Thread: Air Operations

A "Straightforward" Query:

    Find personnel summaries of all individuals with

        Skills A, B, C

        Experience in missions of type D, E

        Medical fitness level M

        Security profile X

        Availability within the next 10 days

Underlying complexity:

    Diverse domains of knowledge

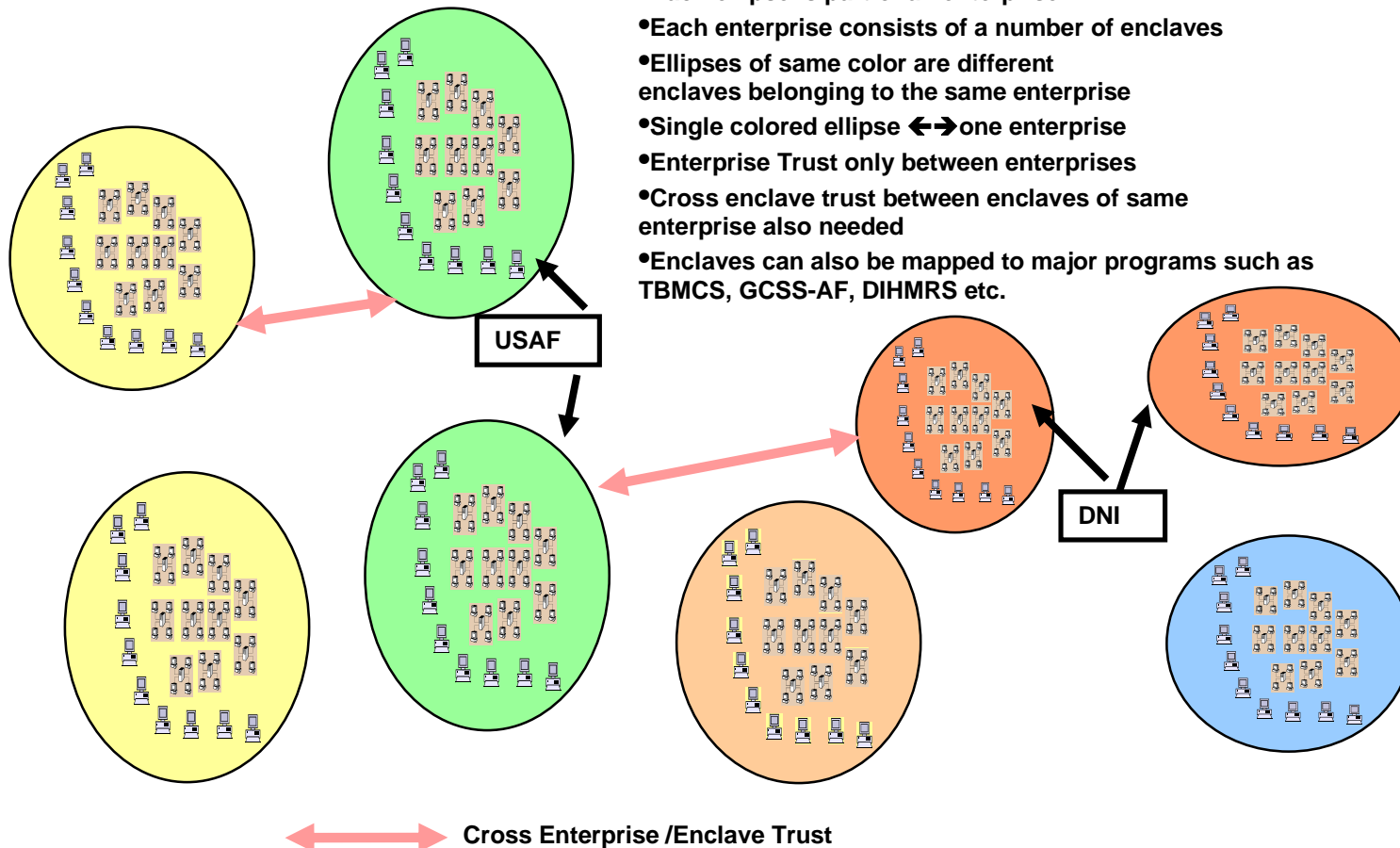   Knowledge exists in different enterprise domains or "enclaves"

    Potential access restrictions (e.g., privacy, personnel records)
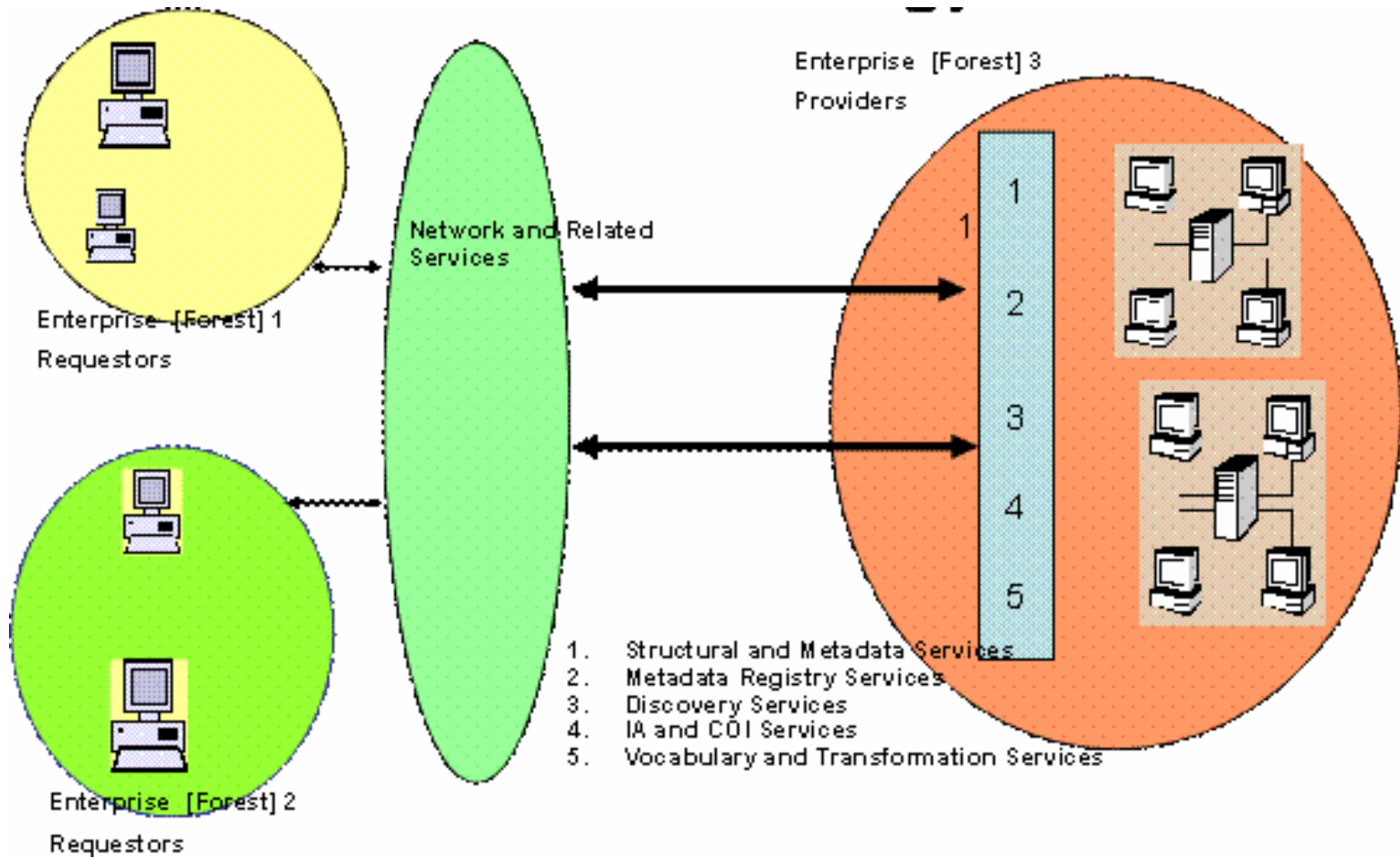
Multi-Enterprise Model

- **MULTI ENTERPRISE MODEL**
- **Each ellipse is part of an enterprise**
- **Each enterprise consists of a number of enclaves**
- **Ellipses of same color are different enclaves belonging to the same enterprise**
- **Single colored ellipse ←→ one enterprise**
- **Enterprise Trust only between enterprises**
- **Cross enclave trust between enclaves of same enterprise also needed**
- **Enclaves can also be mapped to major programs such as TBMCS, GCSS-AF, DIHMRS etc.**

USAF

DNI

Cross Enterprise /Enclave Trust

Model for Info Asset Exploitation

Enterprise [Forest] 3
Providers

Network and Related Services

Enterprise [Forest] 1
Requestors

Enterprise [Forest] 2
Requestors

1. Structural and Metadata Services
2. Metadata Registry Services
3. Discovery Services
4. IA and COI Services
5. Vocabulary and Transformation Services

Scenario Analysis Example

No.	Step	Output	Preconditions	User-Level Questions	System-Level Questions	Supporting Services (Auth., Athzn needed for all)
1	User logs on / authenticates	User's "landing page" or starting page in application	-User has appropriate credentials	What range of authentication methods will be supported?	How to handle Windows forest account vs. non-Windows accounts?	Credential, Audit, Timestamping, Device Services
2	User connects to COI list and discovery "portal"	List of relevant COI's	User is member of at least 1 COI; Assumptions (Stage 1) COIs in 1 enclave (Stage 2) COIs in multi-enclaves	How does user know and select COI's?	How has system determined and validated user COI membership? Who creates COIs and members?	COI Retrieval and Validation, Data
3	User requests list of available tailorable queries	List of tailorable queries	User has validated system about COI membership, If User environment not a web page, app needs to be launched	How does user know about / access web page or application containing queries? Are the query pages common across COI's?	What is the basis for selecting query pages? Is the query related to the metadata structure? Who has constructed and stored the queries?	Metadata, Data, Discovery, Naming Services

Emerging Solution for Info Discovery

Authoritative, aligned vocabularies identified by COIs

Authoritative sources for information assets, COTS indexing tools

Runtime environment:
- Information Assets
- Metadata Registry
- Metadata Catalog
- Service Registry
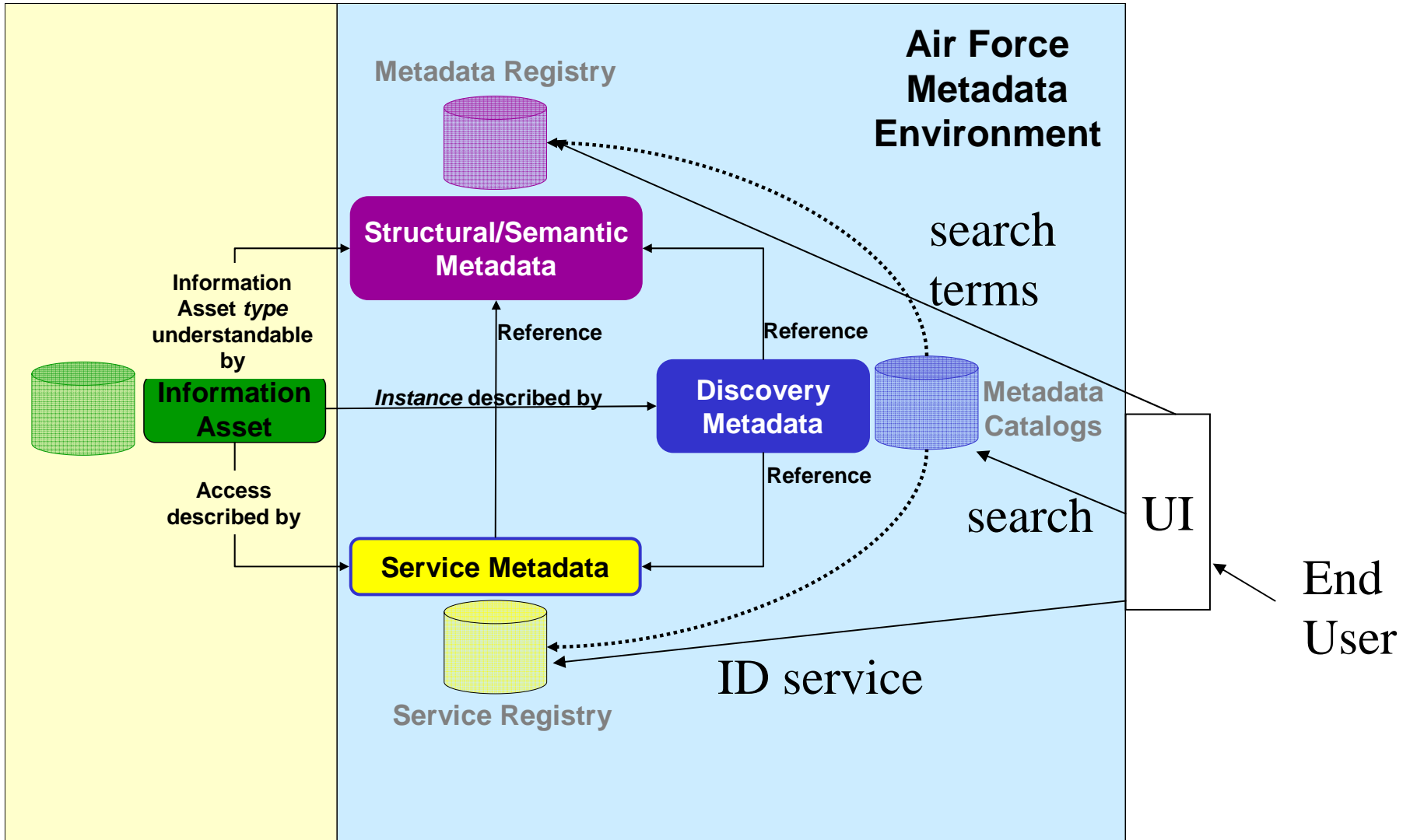
Federated enterprise environment
- Enclave model
- COTS tools
- Industry standards for infrastructure (e.g., OASIS, WS-*)

Metadata Environment Components

INFORMATION TECHNOLOGY & SYSTEMS DIVISION

9

Conclusion

Goals are

Increased effectiveness of discovery,

Leverage information across entire enterprise

Cost savings (remove redundancy, use COTS tools)

Environment is under pilot development (single enclave), moving towards acquisition phase