model futures
www.modelfutures.com

presents:

IntegratedEA

STRATEGY • OPERATIONS • TECHNOLOGY

www: http://www.integrated-ea.com
HashTag: #IEA12
Twitter: @IntegratedEA

VEGA
A Finmeccanica Company

NITEWORKS

Abatis

TEAM ENSURE

de&s

Generic
YOUNG SPIRIT | SENIOR MIND

MooD
SMARTER DECISIONS

Reply
glue

mega

BAE SYSTEMS

The MUSTT of Metadata

Sean Barker
BAE Systems Advanced Technology Centre
sean.barker@BAESystems.com

BAE SYSTEMS

Message

- Meta data is Business Process Data
  - Used in decisions that control processes
  - So is defined by business processes

BAE SYSTEMS

BAE System on a Page

BAE Systems is a global defence and security company with approximately 100,000 employees worldwide.

The Company delivers a full range of products and services for air, land and naval forces, as well as advanced electronics, security, information technology solutions and support services.

Key Facts

- 2nd largest global defence company based on 2010 revenues*
- Global capability
- Customers in more than 100 countries
- 2010 sales of £22.4 billion

* Source: Defense News Annual Ranking, published July 2011

BAE SYSTEMS

Contents

- No such *thing* as meta data
- Meta Data is defined by process
- The MUSTT Principle
- MUSTT as Process Data
- Exception - Search terms

BAE SYSTEMS

No such thing as meta data

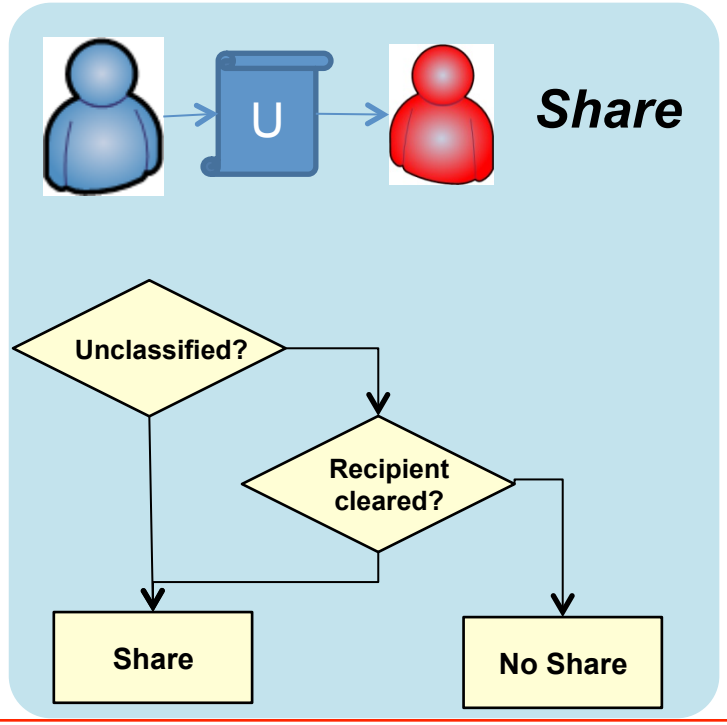
Claim: Meta Data is defined with respect to a process

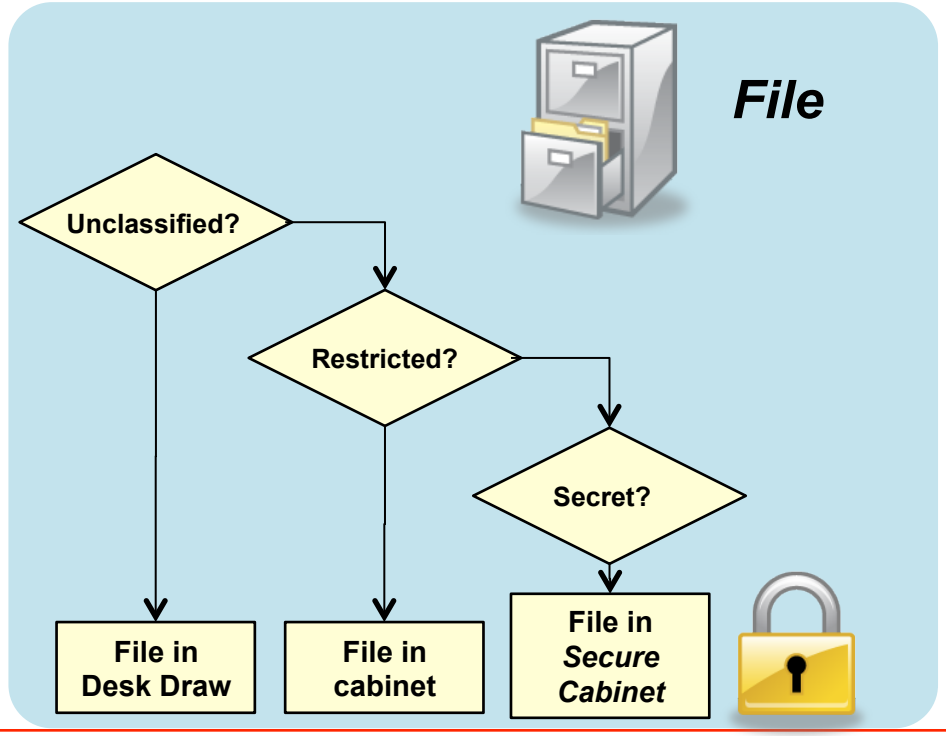
- Meta data is:
  - Relative to some data **artefact**
  - Controls the routes from creation to consumption
- For Example: Dublin Core meta data {...*Author..}*
  - Used to find books or papers
  - Contributes to *Trust* of the book
  - Is **not** meta data to an author payments system
- Meta data changes according to viewpoint

BAE SYSTEMS

Meta data matches Business Process

- Example: Security Classification
  - Defines the measures needed to protect a document
  - Example business processes: Share, File





BAE SYSTEMS

MUSTT Principle

MUSTT principle -

match meta data to the process

What Business Processes? What Data?

MUSTT - five key groups of process:

- Management
  - for keeping track of artefact - e.g. document Id
- Usage
  - who can do what to artefact - e.g. you can read it if you pay
- Search
  - clues to letting someone find the artefact, e.g. keywords
- Trust
  - how reliable and accurate is the artefact e.g. draft or final?
- Technical
  - Used by software - e.g. image size

BAE SYSTEMS

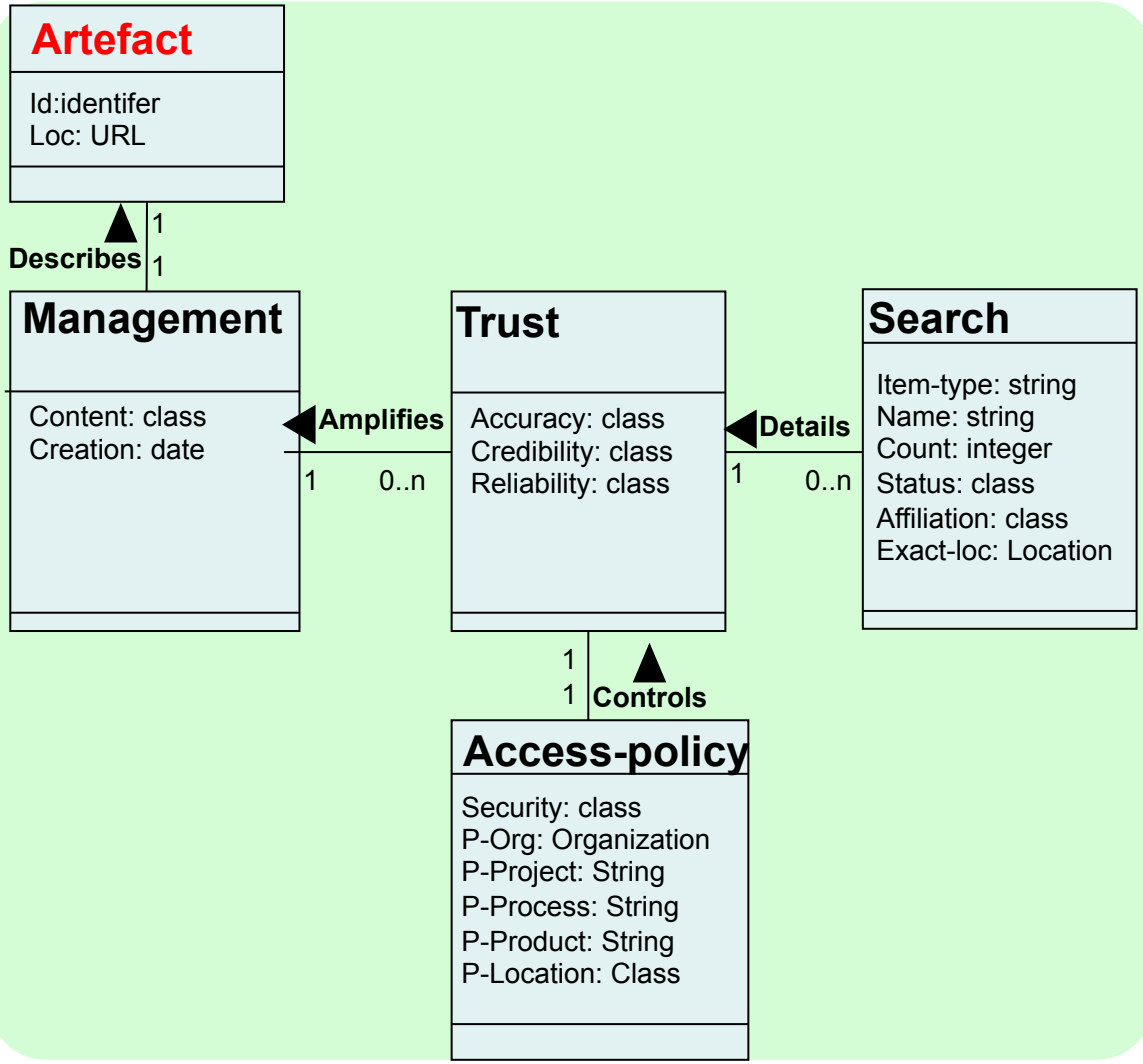
MUSTT in ISRMDUM

Query across ISR sources:
- Document repositories
- Databases
- Website...

Complex Search Model
- based on JC3IEDM
- Search community:
  - Field HQ

Semantic technology
- Automatic query decomposition

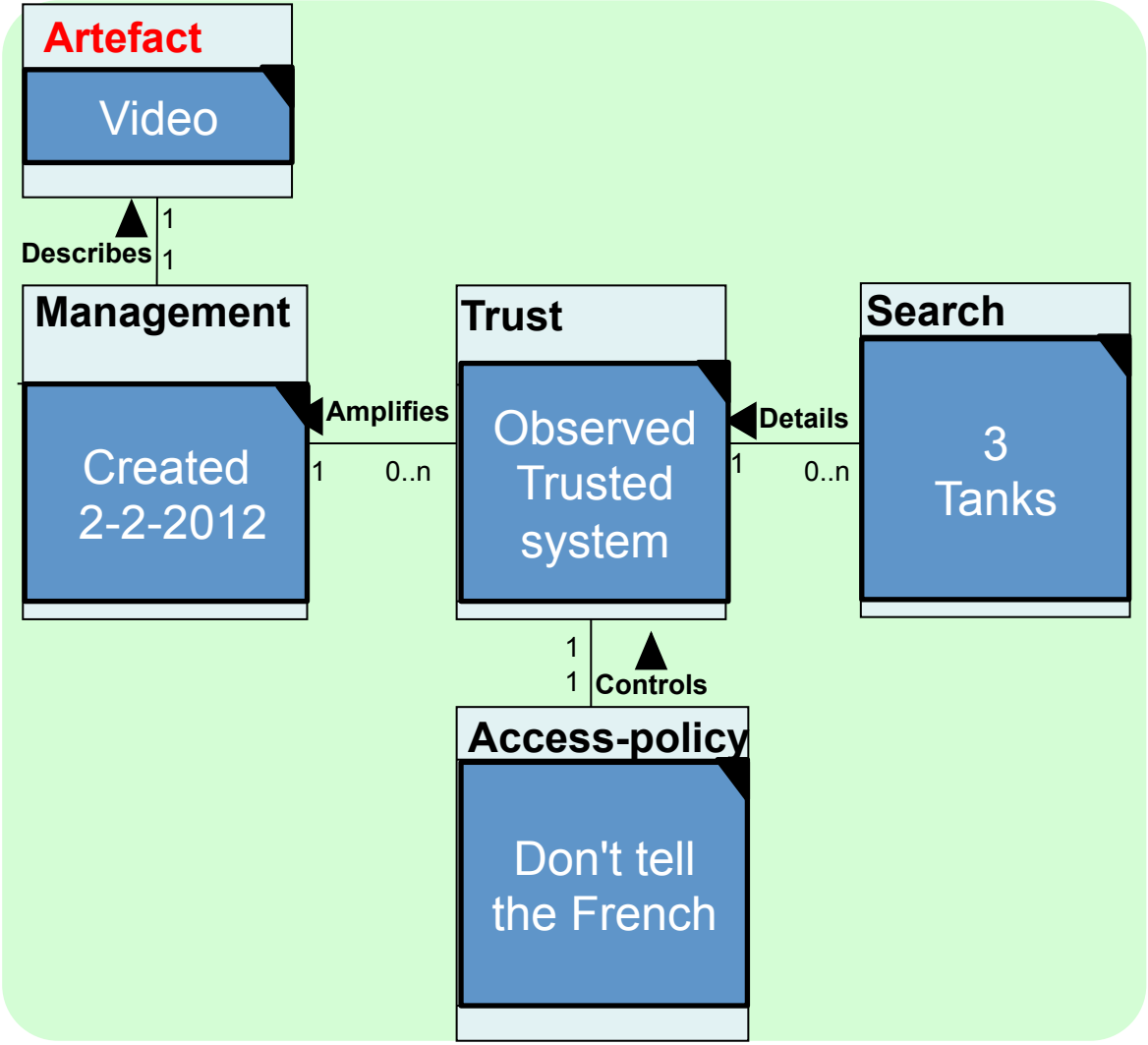
Common meta data based on MUSTT



BAE SYSTEMS

MUSTT in ISRMDUM

Example Data



BAE SYSTEMS

Contents

- No such *thing* as meta data
- Meta Data is defined by process
- The MUSTT Principle
- MUSTT as Process Data
- Exception - Search terms

BAE SYSTEMS

Usage Control

Usage = What can this person do with the artefact?

- Security & Access Control
- IPR, Copyright
- Contractual rights

E.g. Access Control

- What credentials that are needed?
  - Is person in organization? Does person work on project?
  - Is person cleared to the right level? In the right building?
- What Actions are permitted?
  - e.g. read, read-5-times, read-unclassified-extract
- What Consequences follow?
  - Access is logged
  - Payment each time the artefact is accessed

BAE SYSTEMS

Trust

Trust data to assess the quality of the artefact

- Example: Intelligence Data (JC3IEDM)
  - Is the report a rumour or what was actually seen?
  - Is the informant a reliable witness?
  - Are there other correlating sources?
- Example: Document
  - Is it a draft or the final issue?
  - What organization prepared it?

- Special Consideration: Evidential Weight
  - Evidence that the artefact is what it purports to be. e.g.:
    - Digital Signature
    - Chain of custody

BAE SYSTEMS

Exception Search

Search is not precise

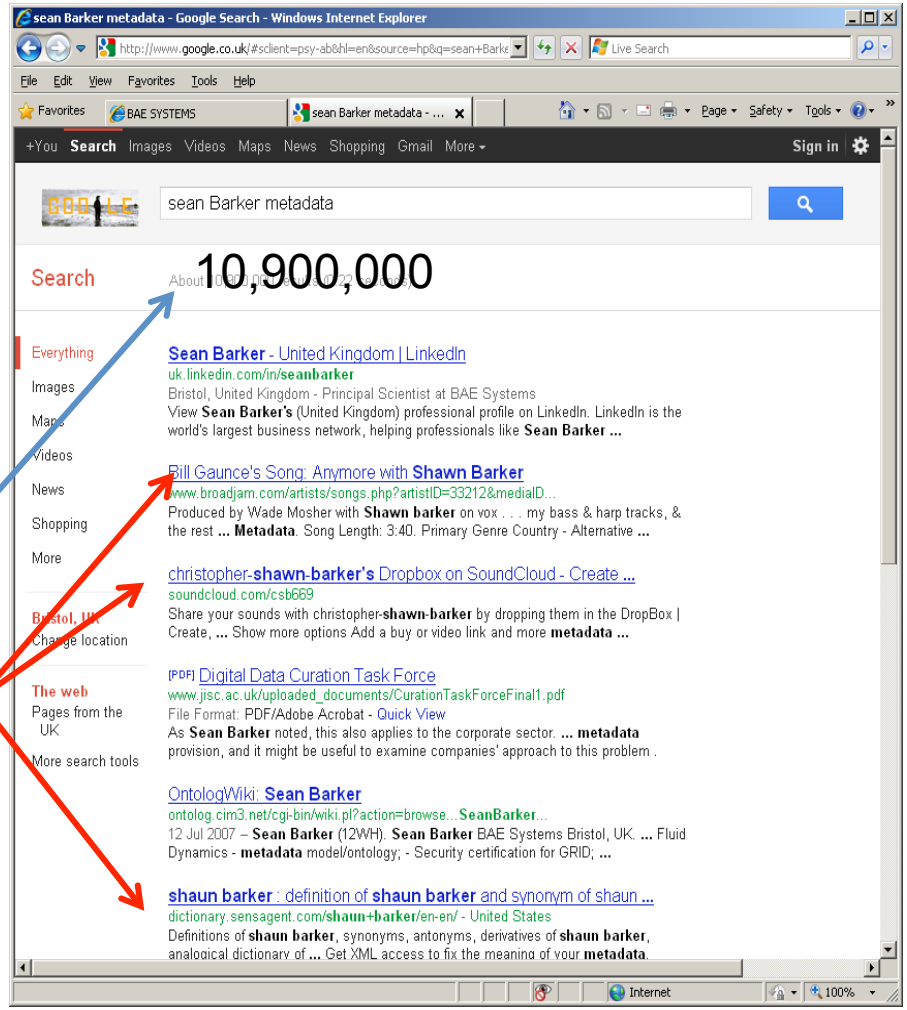
Search creates two Sets

ACCEPT - possible matches

REJECT - not matches

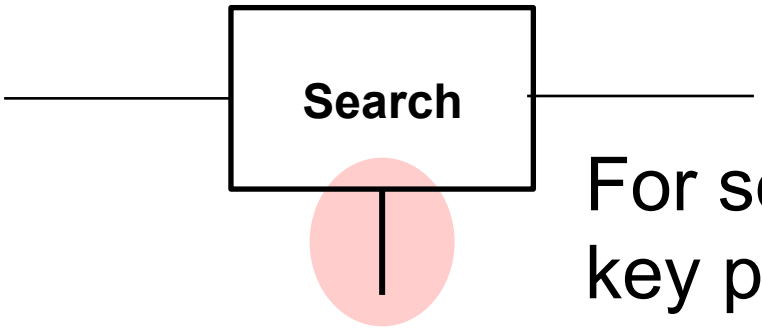
BUT

- ACCEPT Set is too large to check every answer
- ACCEPT Set contains many non-matches
- REJECT may contain valid answers

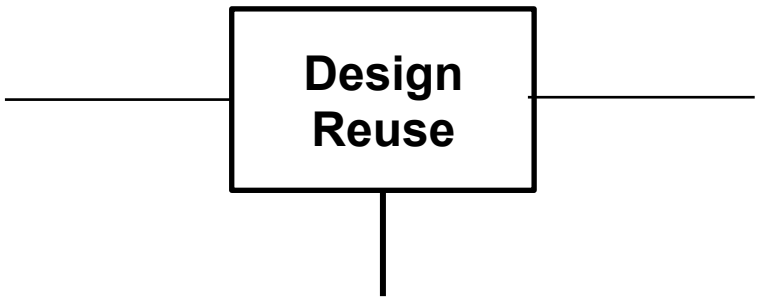


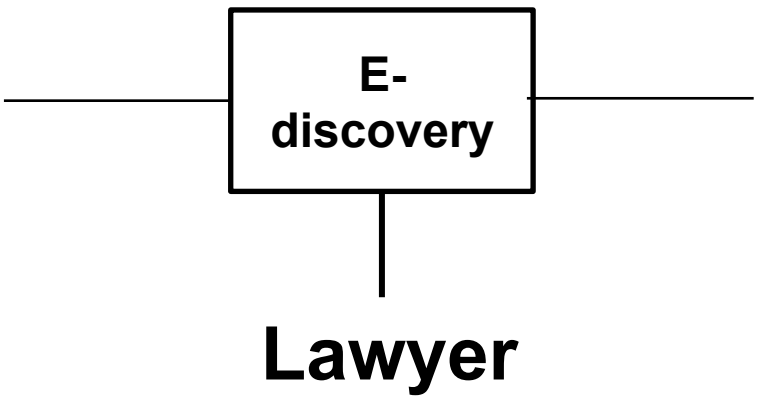
BAE SYSTEMS

Search Process



For search meta data, the key parameter is the searcher





Search terms are defined by the designated communities of users

BAE SYSTEMS

Exception - Search

View Search through Game Theory

Match in ACCEPT **GAIN!**	non-match in ACCEPT **LOSE**
Match in REJECT **LOSE**	non-match in REJECT **gain**

Win overall if

P(gain)*gain > P(loss)*cost-of-loss

Probability Factors:

- \+ Terms are common between creator and user
- \- Some users use a different term for the same thing
- \- Some creators use term in a different way
- \- Many terms are inherently ambiguous

BAE SYSTEMS

Search Precision

Natural language is imprecise

Search involves both selection and ranking

Ranking - the probability an artefact fits into the ACCEPT set

- Semantic search is precise
  - Search as selection only
  - removes homonym ambiguity: e.g. Tank
  - removes role ambiguity
    - Which is the biographer?
  - Can use class hierarchies

A.N. Wilson Iris Murdoch

However - user (natural) language ≠ formal semantics

Users are vague, imprecise, inconsistent even when using "semantic" terms

BAE SYSTEMS

Conclusions and summary

- Meta data is about business processes
- The MUSTT principle helps identify relevant processes and matching data
- The key question:
  - Do the data elements help choose between one process step or another?
- MUSTT means meta data is more focussed and more precise
  - Avoids cost of redundant meta data
  - Avoids losses of inconsistent process decisions