IAO-Intel

An Ontology of Information Artifacts in the Intelligence Domain

Barry Smith
University at Buffalo NY, USA

Tatiana Malyuta
CUNY, NY, USA Data Tactics, McLean, VA

Ron Rudnicki
CUBRC, Buffalo NY, USA

William Mandrick
Data Tactics McLean, VA, USA

David Salmen
Data Tactics McLean, VA, USA

Peter Morosoff
E-Maps, Inc. Washington, DC, USA

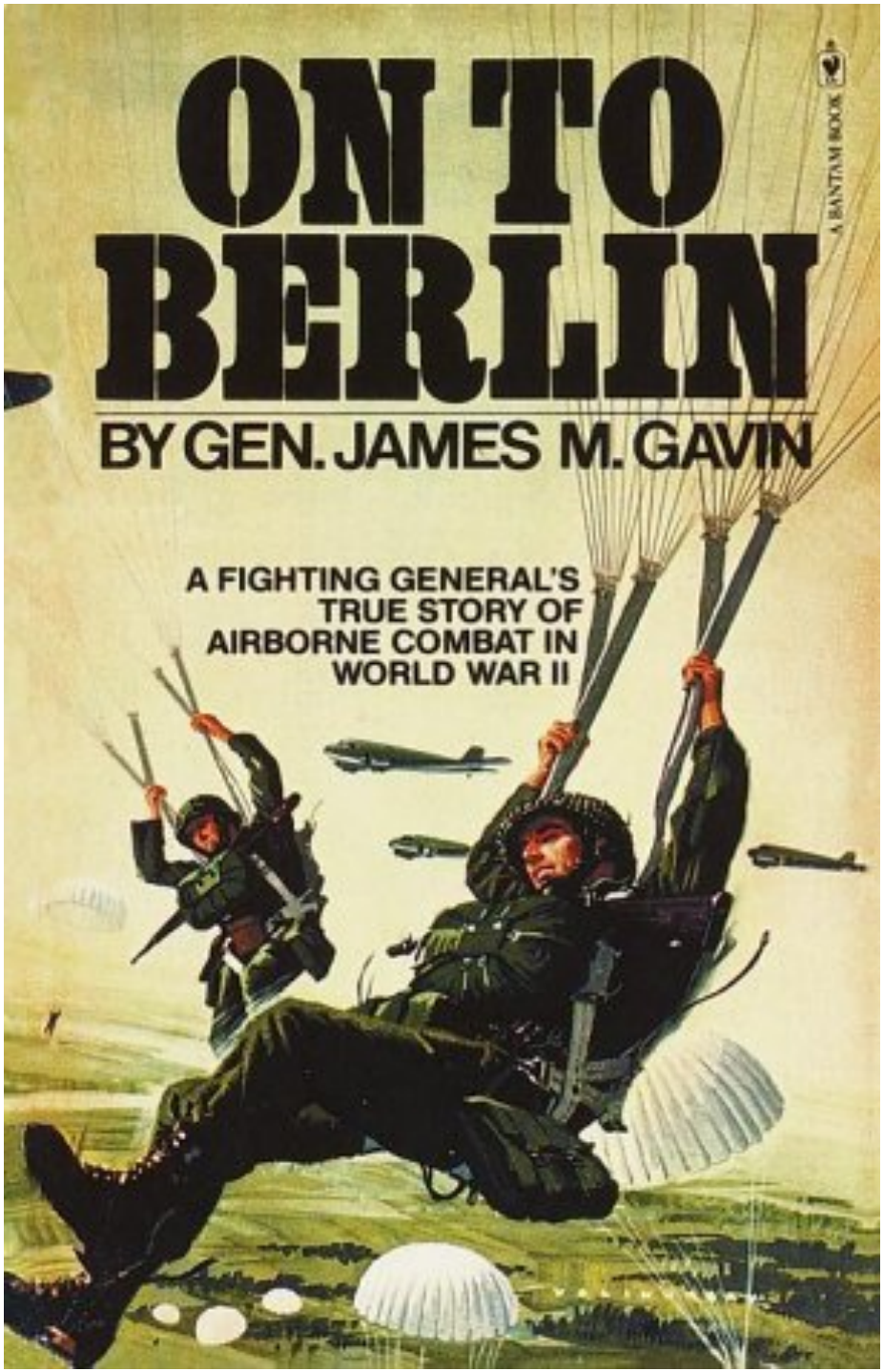
Danielle K. Duff
I2WD Aberdeen, MD, USA

James Schoening
I2WD Aberdeen, MD, USA

Kesny Parent
I2WD Aberdeen, MD, USA

Outline

- Background
  - Military Doctrine and standardization of terminology
  - History of the Information Artifact Ontology (IAO)
- Goals of the IAO
- Application of the IAO: explication and annotation
- Strategy of building and maintaining the IAO
- Organization of the IAO
  - IAO and BFO
  - Attributes of Information Artifacts
- Examples of use

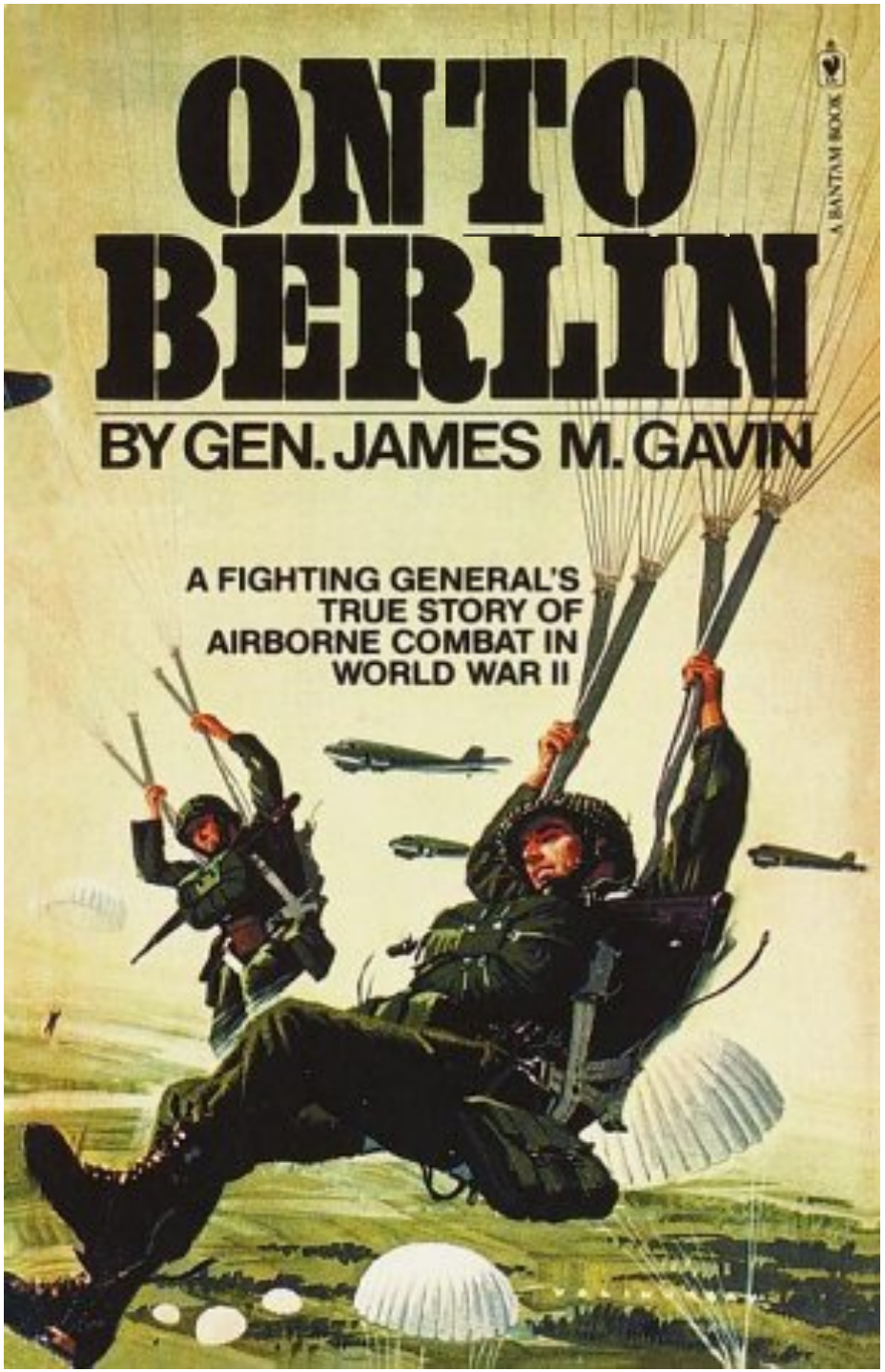

ON TO BERLIN
BY GEN. JAMES M. GAVIN
A FIGHTING GENERAL'S TRUE STORY OF AIRBORNE COMBAT IN WORLD WAR II
A BANTAM BOOK

General James Gavin, *On to Berlin: Battles of an Airborne Commander 1943-1946*

Observations on the contributions of standard terminology to the success of the D-Day invasion

'…. one of our most critical needs was to standardize the operating practices of our forces. … even simple terminology had to be agreed upon. … British flew in what they called "bomber stream" formations, We preferred troop-carrier group formations of 36 planes that flew in a V ... We referred to landing area as the "jump area," the British called it "drop zone," …

So: 'I went to work on a document to standardize airborne practices and I was able to publish the first memorandum on the subject, *Training Memorandum on the Employment of Airborne Forces*, late in 1943.'


ONTO BERLIN
A BANTAM BOOK
BY GEN. JAMES M. GAVIN
A FIGHTING GENERAL'S TRUE STORY OF AIRBORNE COMBAT IN WORLD WAR II

Military Doctrine and Standardization of Terminology

- Since the beginning of organized warfare
  - Beacon signals used by Chinese
  - Drill manual for the units of the Continental Army to respond uniformly to commands during the Revolutionary War
- DOD Dictionary of Military and Associated Terms (Joint Publication 1-02)
- New military dictionaries and terminology artifacts continue to be developed

Goals of Doctrinal Documents

- Advance consistency in communications
- Facilitate consistent interpretation of commands
- Compile lessons learned (outcomes assessment)
- Provide controlled vocabularies for official reporting
- Enhance discoverability and analysis of data.

Problems of Doctrinal Documents

- Little aid to computation
- Developed independently in divergent and non-principled ways
- As the result identical data may be classified and described entirely differently by different agencies, which causes failures of data integration initiatives
- New approaches are needed that can enable computational discovery, retrieval, integration and processing of data
- For example: Directive 8320.02 (last updated August 5, 2013) requires all authoritative DoD data sources to be registered in the DoD Data Services Environment (DSE)

Two kinds of data

1. Data about *entities in the world*
2. Data about the *information artifacts in which these entities are represented* (= metadata)

Information Artifacts

artifact =def. an entity created through some deliberate act or acts by one or more human beings and which endures through time

information artifact =def. an artifact that can be the bearer of information

(a) a physical entity – a hard drive, a passport, a photographic image

(b) as an information content entity which can potentially exist in multiple (for example digital or printed) copies – a pdf file, a jpg file

Information content entities are entities which can be *about* something

IAO

- OBO Foundry – to aid the consistent description of biological data emanating from multiple heterogeneous sources
- IAO-Intel – to provide common resources for the consistent description of information artifacts of relevance to the intelligence community in a way that will allow discovery, integration and analysis

IAO Goal

- Is not designed to *replace* existing doctrinal or other standards created to guide human beings or computer applications in the creation and description of documents. Rather, its purpose is to allow the results of using such standards to generate the needed metadata in a uniform, non-redundant and algorithmically processable fashion
- To describe attributes of IAs, need to distinguish
  - The *particular* information artifact of interest, tied to some particular physical information bearer: the photographic image on this piece of paper retrieved from this enemy combatant
  - The *copyable* information content that is carried by the artifact in question. The photographic image may be printed out in multiple paper copies

IAO Application

- US Army's Distributed Common Ground System (DCGS-A) Standard Cloud (DSC) initiative as part of a strategy for the horizontal integration of warfighter intelligence data
- Two sorts of application to support computer-aided retrieval and analytics
  - *Explica-tion* of general terms used in source intelligence artifacts and in data models, terminologies and doctrinal publications which provide typo-logies of intelligence-related IAs to semantically enhance data in a way that enables computational integration and reasoning
  - *Annotation* of the instance-level information captured by such IAs to aid retrieval of information about specific persons, groups, events, documents, images, and so forth

Strategy of Building IAO-Intel

Downward population from the draft IAO reference ontology

IAO	IAO-Intel (examples)
Report	Intelligence Report (FM 6-99.2, 126)
Summary	Electronic Warfare Mission Summary (FM 6-99.2, 87)
Diagram	Network Analysis Diagram (from JP 2-01.3, II-51)
Overlay	Combined Information Overlay (JP 2-01.3, II 33)
Assess-ment	Assessment of Impact of Damage (FM 6-99.2, 53)
Estimate	Adversary Course of Action Estimate
List	List of High-Value Targets (JP 2-01.3, II 61)
Order	Airspace Control Order (FM 6-99.2, 17)
Matrix	Target Value Matrix (JP 2-01.3, II-63)
Template	Ground and Air Adversary Template (JP 2-01.3, II-57)

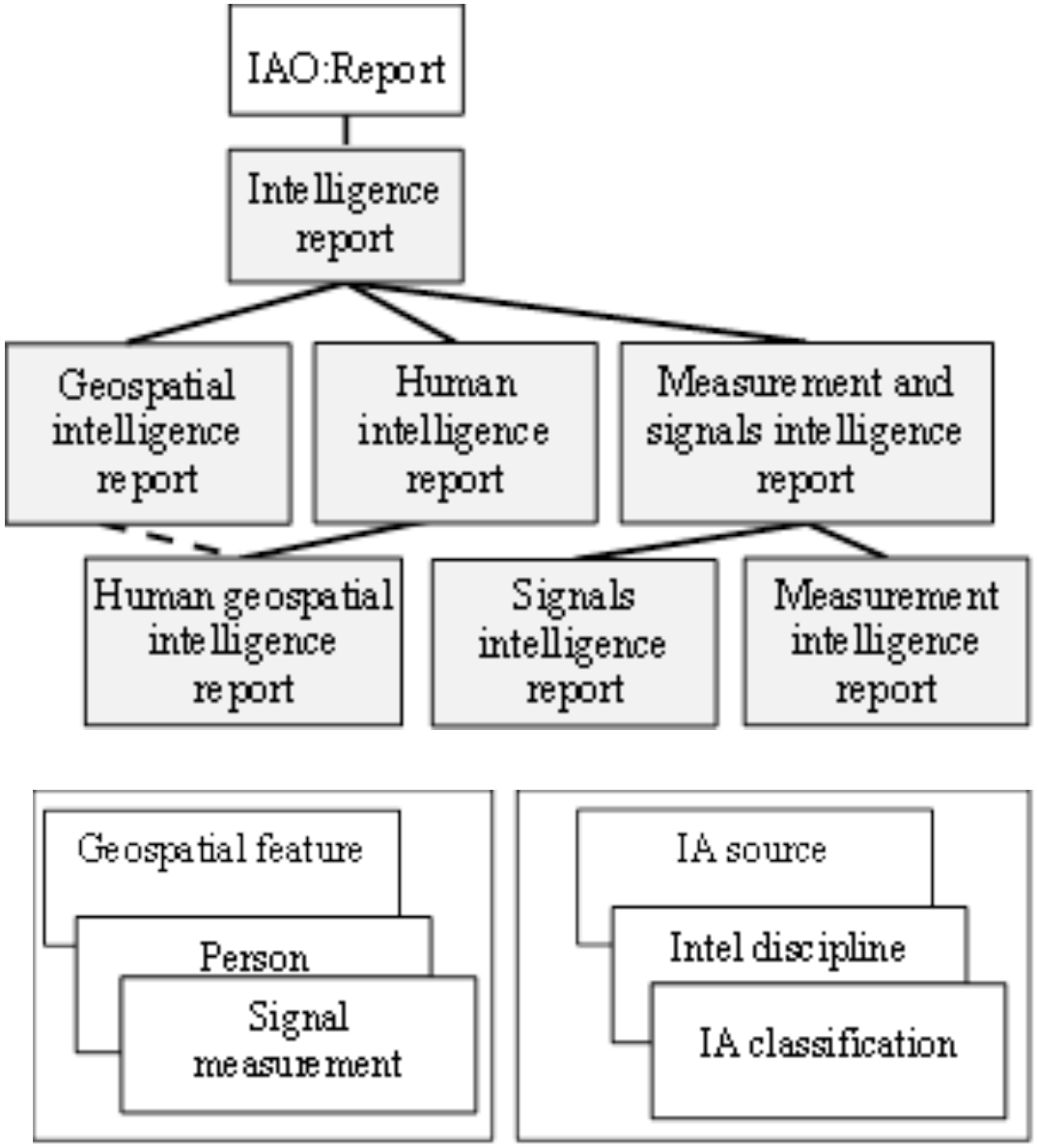
Strategy of Building IAO-Intel (cont.)

- Identify orthogonal dimensions of IA attributes and create Low-Level Ontology modules (LLOs)
  - Small, shallow, and structured following the principle of single inheritance
  - Used to
    - *Construct more complex terms* and *define* the terms of the IAO-Intel ontology and of its sister ontologies
    - *Explicate* the meanings of terms standardly used by different agencies
    - *Annotate* instance data

Strategy of Building IAO-Intel (cont.)

- Avoid combinatoric explosion of terms like 'Vehicle Inspection- Jurisdiction Authority-Text' (but allow Doctrinal terms like EEFI –'Essential Element Of Friendly Information') and allow compound annotations
  - Vehicle Inspection Report (aboutness)
  - EEFI (importance)
  - located-at, etc.
- Incremental development

IAO-Intel



The IAO-Intel terms on the top are defined by using terms from the ontologies on the bottom with the help of relations such as *is-about*, *created-by*, *derives-from* and so forth

Organization of IAO-Intel – ICE

- Information Content Entities (ICEs) are *about* something in reality (they have this something *as a subject*; they *represent*, or *mention* or *describe* this something; they *inform us about* this something). Aboutness may be identifiable from different perspectives. Thus one analyst may interpret a given ICE as being about the geography of a given encampment; another may view it as providing information about the morale of those encamped there.

Organization of IAO-Intel – IBE

- *Information Bearing Entity* (IBE) is a material entity that has been created to serve as a bearer of information. IBEs are either (1) self-sufficient material wholes, or (2) proper material parts of such wholes. Examples under (1) are: a hard drive, a paper printout (e.g., a report); and under (2): a specific sector on a hard drive, a single page of a paper printout.

Organization of IAO-Intel – IQE and ISE

- *Information Quality Entity* (IQE) is the pattern on an IBE in virtue of which it is a bearer of some information
- *Information Structure Entity* (ISE) is a structural part of an ICE, for example an empty cell in a spread-sheet; a blank Microsoft Word file. ISEs thus capture part of what is involved when we talk about the 'format' of an IA.

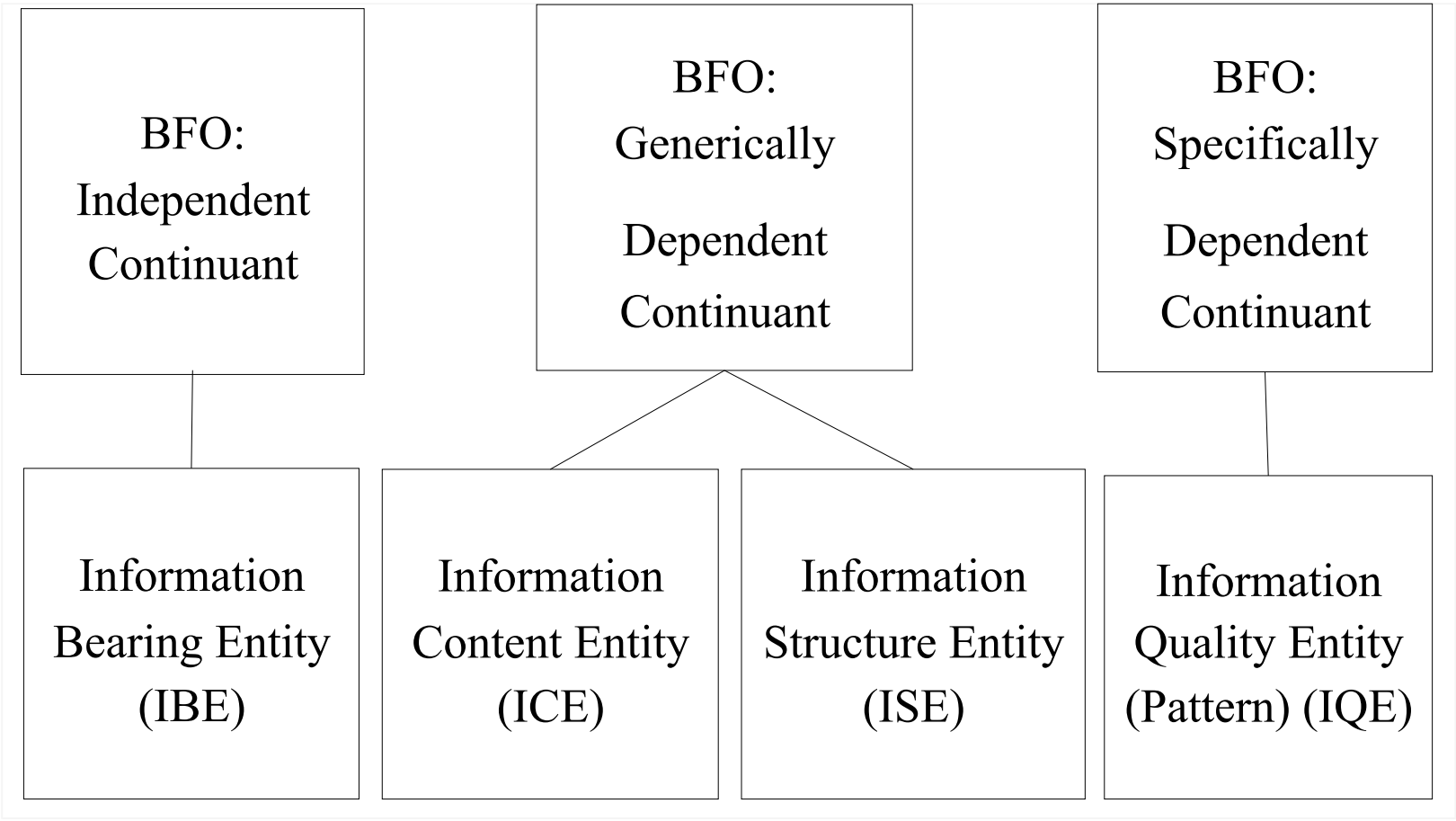
Organization of IAO-Intel – IA

- IA refers either
  - Some combination of ICEs and ISEs (roughly: the IA as body of copyable information content); or
  - Some concreti-zation of ICEs and ISEs in some IBE in which some IQE inheres (the information artifact is: this content here and now, on this specific computer screen or this printed page). Different information artifact types will differ in different ways along these dimensions, as illustrated in Table 2.

Example of IAO-Intel Organization

Information Artifact	IBE	ISE	ICE
MS Word file (.doc, .docx)	Hard drive (magnetized sector)	MS Word format	Varies
KML file	Hard drive (magnetized sector)	KML	Map overlay
JPEG file (.jpg)	Hard drive (magnetized sector)	JPEG format	Image
Email file	Hard drive (magnetized sector)	Internet Message Format (e.g., RFC 5322 compliant)	Message
USMTF Message file	A specific government network	USMTF Format	Message
Passport	Paper document; (may include photographs, RFID tags)	ID formats, security marking formats …	Name, Personal data, Passport number, Visas …
Title Deed	Official paper document	Varies	Varies
Report	Varies	Varies	Varies
Overlay Sheet (e.g. Map Overlay Sheet)	Acetate sheet	MIL-STD-2525 Symbols; FM 101-1-5 Operational Terms and Graphics	Map overlay

IAO and BFO


BFO:
Independent
Continuant
BFO:
Generically
Dependent
Continuant
BFO:
Specifically
Dependent
Continuant
Information
Bearing Entity
(IBE)
Information
Content Entity
(ICE)
Information
Structure Entity
(ISE)
Information
Quality Entity
(Pattern) (IQE)

IAO and BFO (cont.)

- IBEs are *independent continuants*. An IBE is a physical entity that is created or modified to serve as bearer of certain patterned arrangements – for example of ink or other chemicals, of electromagnetic excitations.
- IQEs are *specifically dependent continuants* – entities which require some specific physical bearer but which are not themselves physical. An IQE is a quality of an IBE which exists in virtue of such patterned arrangements and which is interpretable as an ICE or ISE.

IAO and BFO (cont.)

- ICEs and ISEs are *generically dependent continuants* (GDCs); they are entities – such as a pdf file or an email – which can be copied from one physical bearer to another and thus may exist simultaneously in multiple different IQEs, which are called 'concretizations' of the corresponding GDC.

IAO and BFO (cont.)

- BFO *relations* between ICEs, ISEs, IQEs and IBEs can be set forth as follows:
  - ICE *generically-depends-on* IBE
  - ISE *generically-depends-on* IBE
  - IQE *specifically-depends-on* IBE
  - ICE *concretized-by* IQE
  - ISE *concretized-by* IQE
- IAO contains in addition relations which allow to formulate metadata concerning attributes of IAs such as author, creation date, classification status, and so forth

Attributes of IAs

- Information artifacts have attributes along a number of distinct dimensions, treated in LLO modules
- Terms in these modules will be applied to explicate information relating to IAs of different types, and to annotate data pertaining to IA instances with the help of relations mentioned above

Attributes of IAs (cont.)

- Some dimensions of IA attributes are common to all areas, both military and non-military
  - *Purpose*
  - *Life-cycle Stage* (draft, finished version, revision)
  - *Language*,
  - *Format*
  - *Provenance*
  - *Source* (person, organization)

Attributes of IAs (cont.)

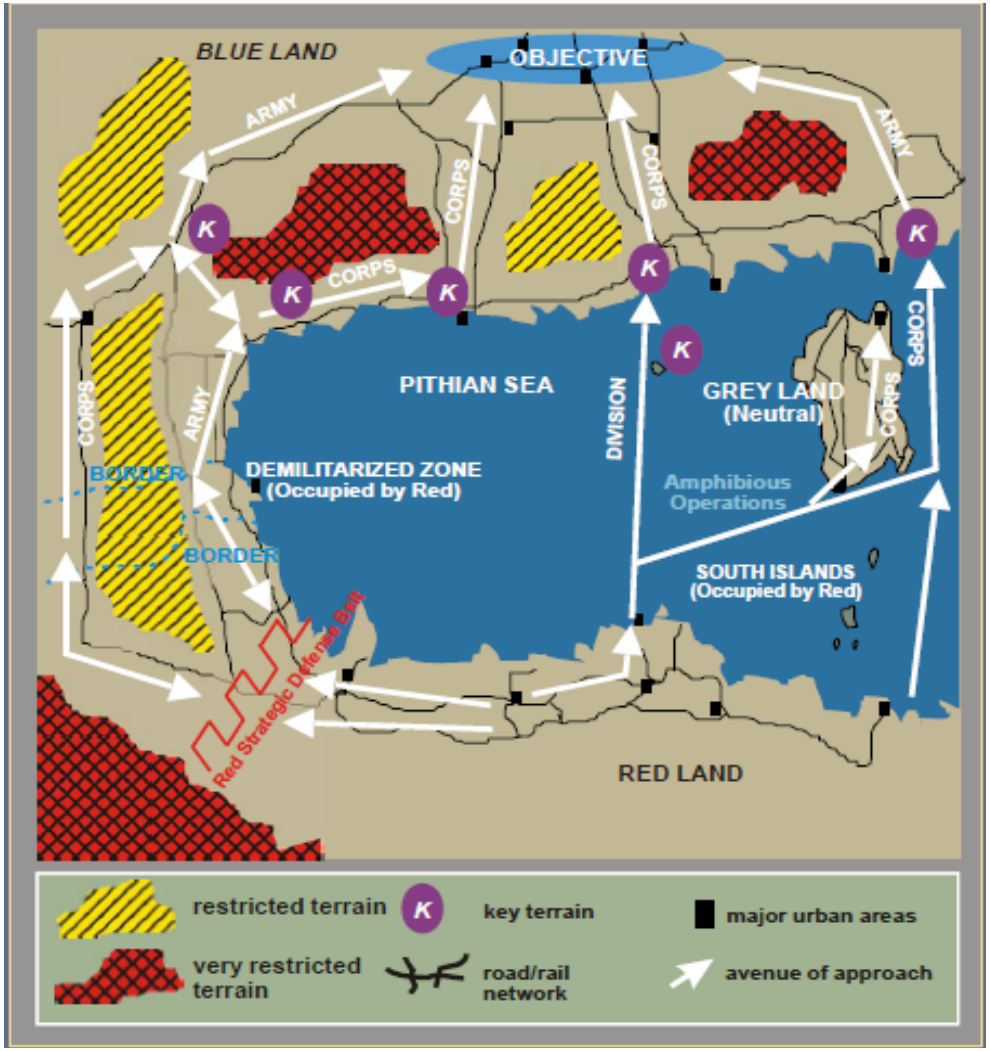
- Purpose
  - Descriptive purpose: scientific paper, newspaper article, after-action report
  - Prescriptive purpose: legal code, license, statement of rules of engagement
  - Directive purpose (of specifying a plan or method for achieving something): instruction, manual, protocol
  - Designative purpose: a registry of members of an organization, a phone book, a database linking proper names of persons with their social security numbers
- Purposes specific to IAO-Intel
  - Inform-ing the commander,
  - Providing targeting support
  - Intelligence preparation of the battlefield.

Attributes of IAs Specific to Intelligence IAs

Role in the Intelligence Process (JP 3-0, III-11) Priority Intelligence Requirement (PIR) Commander's Critical Information Requirement (CCIR) Essential Element of Information (EEI) Essential Element of Friendly Information (EEFI)	
Confidence Level (JP 2.0, Appendix A)	
Highly Likely Likely Even Chance	Unlikely Highly Unlikely
Discipline (JP 2.0, I-5) Legal Ideology Religion Propaganda	Intelligence Signal Human Rumor intelligence Web intelligence
Intelligence Excellence (JP 2.0, II-6)	
Anticipatory Timely Accurate Usable	Complete Relevant Objective Available

Use of IAO-Intel – Example 1

- **IA #1** - Modified Combined Obstacle Overlay (MCOO) - a joint intelligence preparation of the operational environment product used to portray the militarily significant aspects of the operational environment, such as obstacles restricting military movement, key geography, and military objectives.
- **IA #2** – a plan in accordance with which the IA #2 was prepared



Use of IAO-Intel – Example 1 (cont.)

- Annotations to the attributes of IA #1
  - ICE: MCOO
  - IBE: Acetate Sheet
  - *uses-symbology* MIL-STD-2525C
  - *authored-by* person #4644
  - *part-of* plan IA#2
- Annotations relating to the aboutness of IA#1
  - Avenue of Approach
  - Strategic Defense Belt
  - Amphibious Operations
  - Objective

Use of IAO-Intel – Example 2

- A collection of documents prepared according to FM 6-99.2 of types:
  - Intelligence Report [INTREP]
  - Intelligence Summary [INTSUM]
  - Logistics Situation Report [LOGSITREP]
  - Operations Summary [OPSUM]
  - Patrol Report [PATROLREP]
  - Reconnaissance Exploitation Report [RECCEXREP]
  - SAEDA Report [SAEDAREP]

Use of IAO-Intel – Example 2 (cont.)

- We need to computationally cross-reference these with comparable sets of documents prepared by other commands.
- FM 6-99.2 provides definitions of the mentioned report types, but does not take the step of formulating these definitions computationally.
- IAO-Intel addresses this problem by providing a common, algorithmically useful, set of ontology terms that is designed to allow consistent explication of these and related types as they appear in different doctrinal resources.
- The results can then be used for computer-aided aggregation of the data represented using corresponding IA types, cross-checking of mismatches, and so forth.

DOD DSE

Semantic Technology not Enough

Conclusion

Acknowledgements

Work on IAO-Intel was supported by I2WD.

Thanks are due also to Mathias Brochhausen, Werner Ceusters, Mélanie Courtot, Janna Hastings, James Malone, Bjoern Peters, Jonathan Rees, and Alan Ruttenberg for their work on IAO.