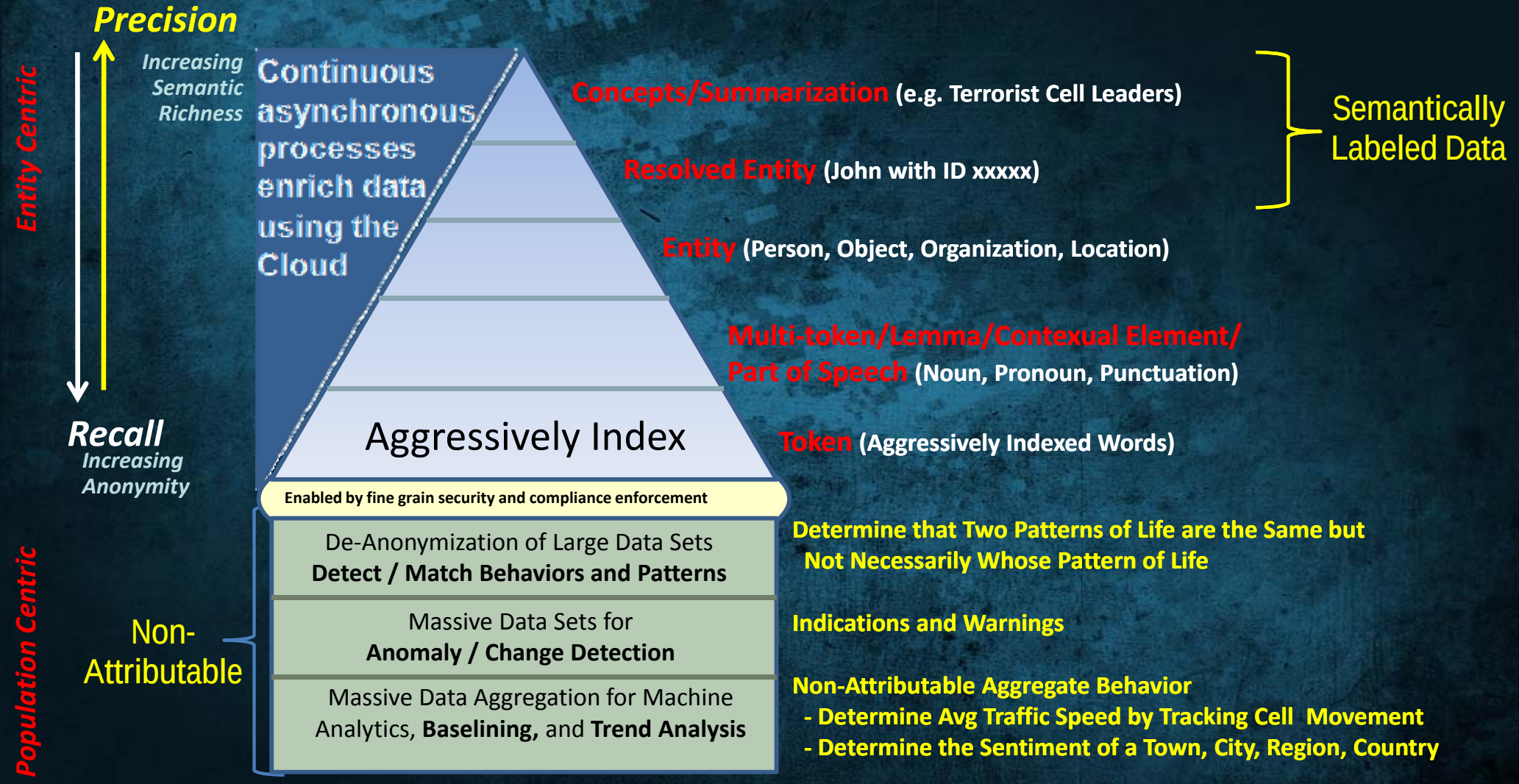
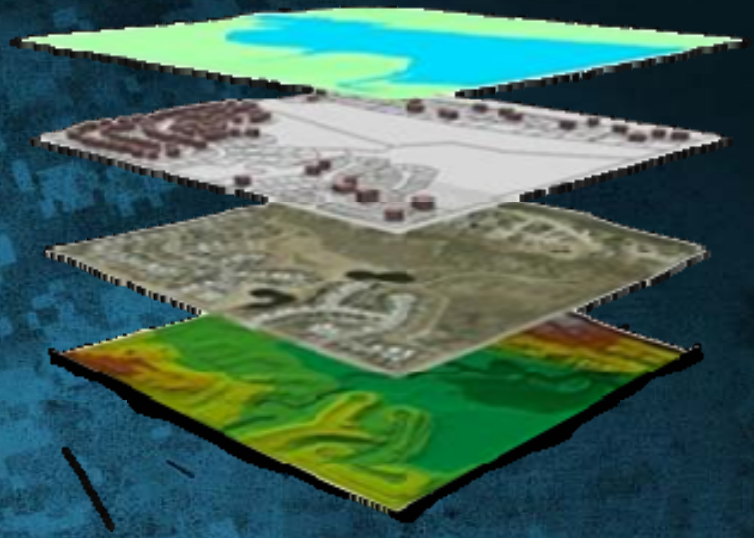
Dr. Russell D. Richardson, G2/INSCOM Science Advisor

Semantic Enrichment of the Data

Solving the Precision / Recall Conundrum with Semantic Enrichment of the Data



Extending Cloud-Enabled Advanced Analytics



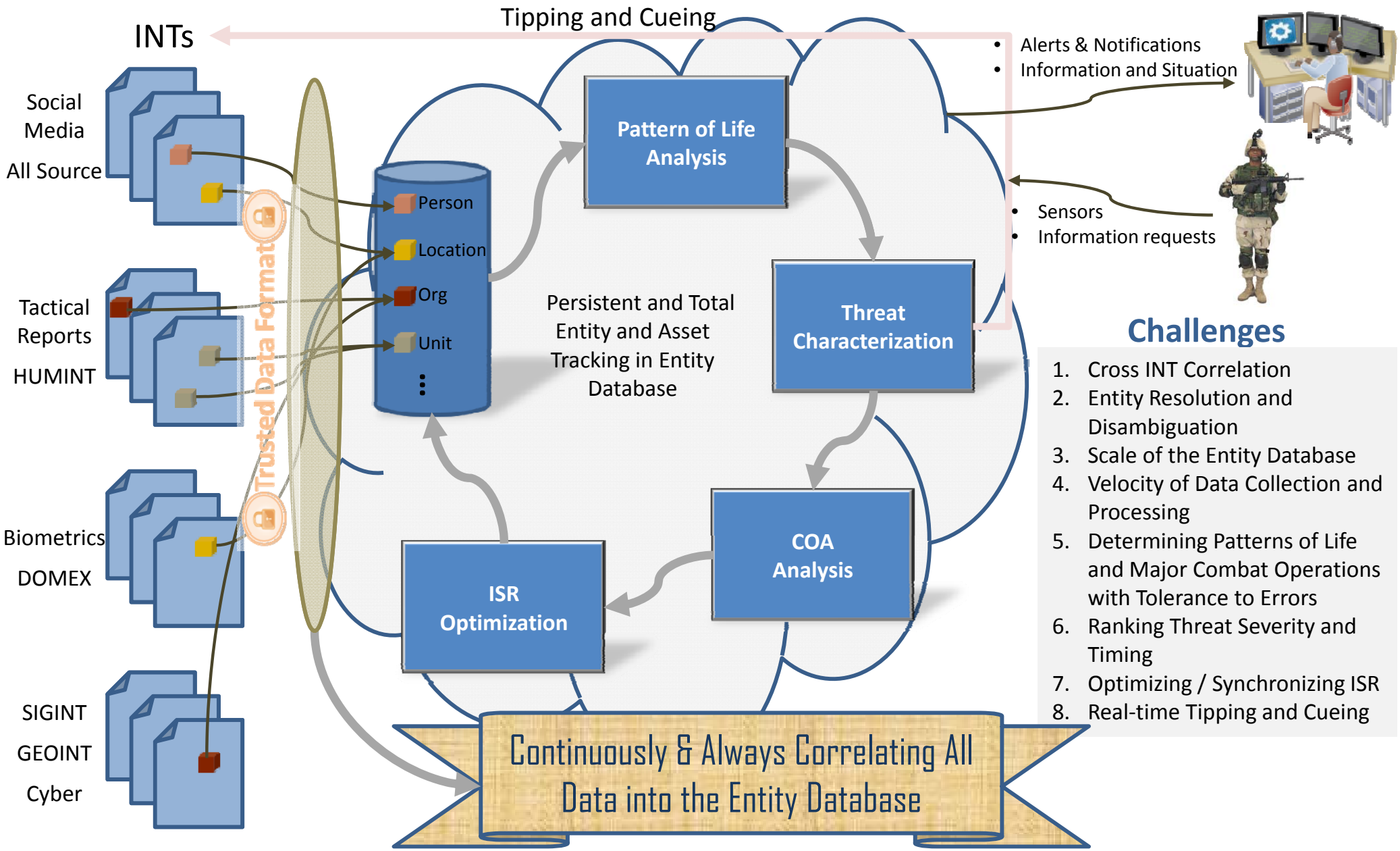
Continuous asynchronous processes enrich data using the Cloud

Aggressively Index

Statistical Analysis of "Big Data" Enables Change Detection + Population-Centric Analysis

All-Source Analytics for 'Big Data' with Advances in Geospatial Indexing, Voice Index and Search, Biometric Entity Management, Motion Imagery Tracks, Multi-INT Visualization, Collection Management, Support for Mobile Devices, Powerful Compute Platforms, Multi-Level Security, IC Shared Software, ….

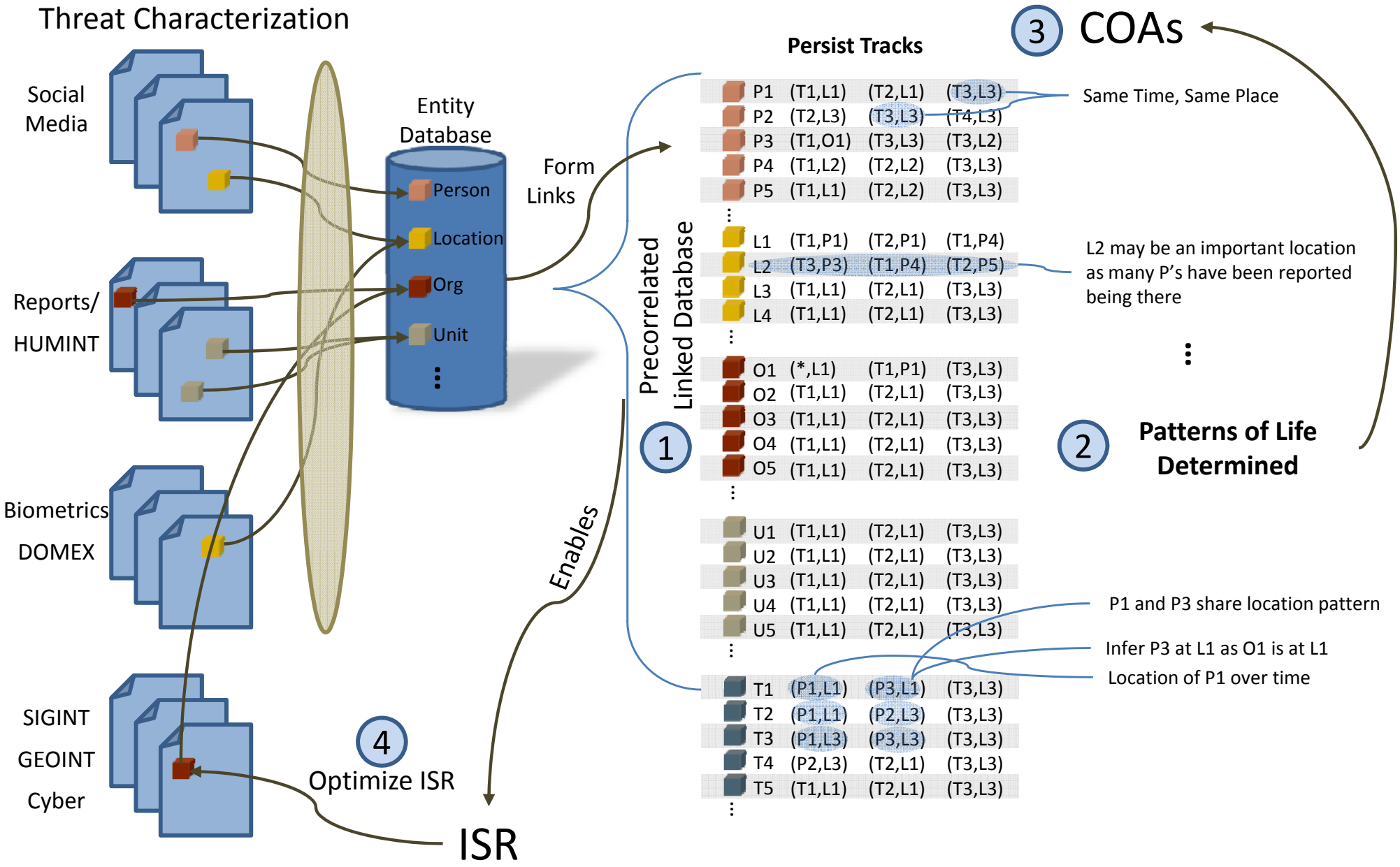
Fusion Challenges



Challenges

1. Cross INT Correlation
2. Entity Resolution and Disambiguation
3. Scale of the Entity Database
4. Velocity of Data Collection and Processing
5. Determining Patterns of Life and Major Combat Operations with Tolerance to Errors
6. Ranking Threat Severity and Timing
7. Optimizing / Synchronizing ISR
8. Real-time Tipping and Cueing

Need to Work Entity Tracking



The DCGS-A ICITE Cloud (aka Red Disk) at-Scale Entity Database

Data Interoperability

Enterprise Context

Decreases Information Burden

Full Spectrum Analytic Awareness

Sensors

Sources

- SIGINT
- MTI
- WAMI
- FMV
- TED
- Fires
- Harmony
- USMFT
- Collection Mgmt
- Open Source
- Mission Command
- Audio
- etc

Velocity & Content

RTAAP

NIFI / Storm

Real-Time Advanced Analytics Pipeline … updating analytics earliest as possible.

- EDH
- Provenance
- TDF
- Metadata tagging
- Geo temporal extraction
- Entity extraction and nomination
- Artifact enrichment
- Security Labeling
- Metrics
- more

Artifacts, Terms Statements

Analytics and indexes updated wrt to each artifact

UCD

M/R

- Enrichment
- Analytics and indexes wrt to the entire corpus,
- Bulk and incremental updates
- Data sharing

Data Access Process - User's authorizations and roles are matched to data security labels

DIAS

User's Authorizations

Community Partners

Analyst's conclusions enrich the UCD

- Maximally correlated data
- Contextual-based navigation
- All data under a common representation to enable assess to all layers
- Cross corpus analytics without custom code
- Signatures are analyzed in real time
- Logical representation needs to be consistent in order for the physical model and view to be dynamic

- Classes of relationships determined at various points
- No all relationships need to be explicitly expressed – correlations enable this
- All analytic disciplines together with correlated data among all disciplines

Analytic models are represented in the UCD