Out of the Ordinary: Finding Hidden Threats by Analyzing Unusual Behavior

A New Approach to "Connecting the Dots" in Intelligence

John Hollywood
RAND

Presentation to the NDIA Systems Engineering Conference
October 26, 2005

Introduction

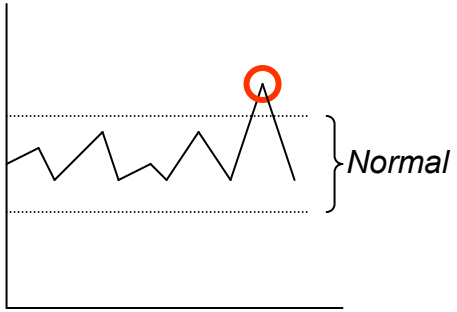
- Today's presentation is on the "Atypical Signal Analysis and Processing" (ASAP) Schema, an architecture for—
  - Identifying "unusual" observations worth further attention
    - E.g, violations of the status quo
  - Relating these observations to other data
  - Generating and testing hypotheses about these observations
- Purpose of research was to identify and characterize terrorist threats, but basic approach may have a number of other applications
  - Previously asked about applications for "tracking and identifying problem acquisition personnel early in the process"
  - Might also have applications for PM and SE
- From a RAND IR&D project – results documented *Out of the Ordinary: Finding Hidden Threats by Analyzing Unusual Behavior* (RAND MG-126-RC)
  - Focuses primarily on top-level architecture, with some discussion of data management and analysis algorithms

Motivation for "Out of the Ordinary" Research

- Immediate motivation was to improve ability to "connect the dots" – pieces of information that could be combined to produce understanding of a threat
- Focus on intelligence and homeland security data, especially on reports
  - Others were working on large-scale data mining; taking advantage of intelligence / HLS reports was unique opportunity
  - Also wanted to avoid privacy / legal issues
  - For acquisition, equivalent would be steady stream of reports and products coming out of PM and acquisition offices, plus any reported complaints
- Additional focus – look for "anti-patterns"
  - Knew we wanted to avoid "fighting the last war problem" (I.e. just looking for patterns matching previous attacks)
  - Also inspired by how proactive, successful problem solvers have "connected the dots" in the past…

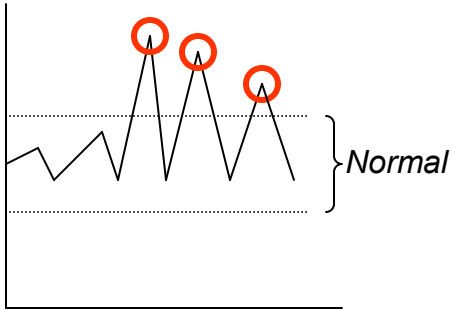
Proactive, Successful Problem Solvers Study the Atypical to "Connect the Dots"

1. Identify Atypical Behavior



Plant Managers' E-mail Traffic

2. Learn More About the Behavior



Plant Managers' E-mail Traffic

3. Understand the Meaning of the Behavior

- **Fact**: E-mail surge -> managers expected to go to a day-long meeting
- **Fact**: Quality goes down when managers are not supervising machines

4. Respond to the Meaning

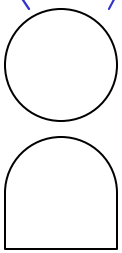
- **Response**: Don't make any important components while the meeting is going on

1. Observe behavior, looking for status quo deviations

2. Big deviations (number or magnitude) are most significant

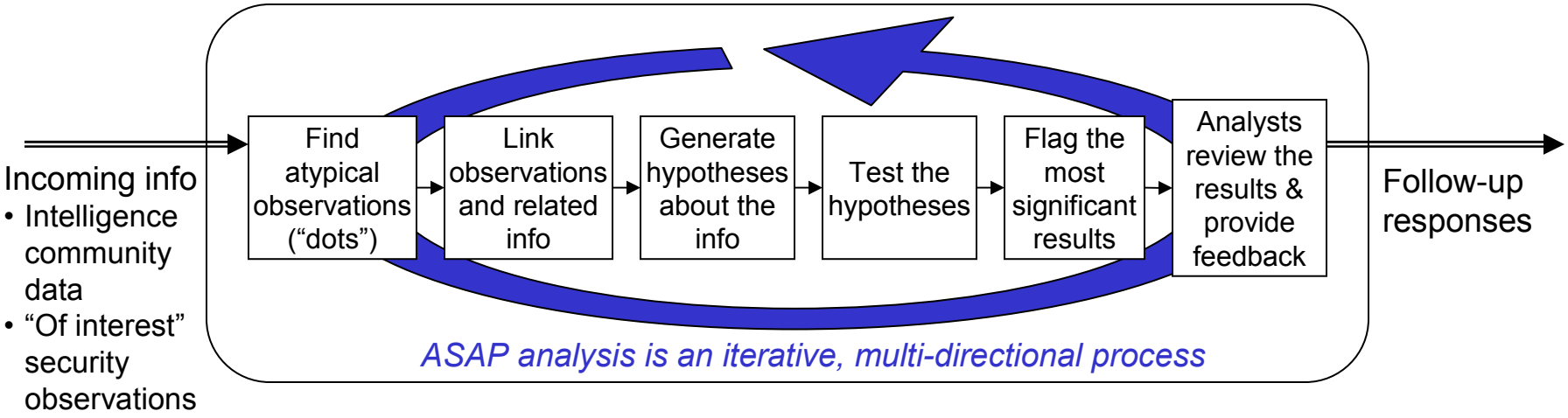
3. Search for related trends, info that helps explain deviations

4. Employ risk-avoidance and mitigation heuristics



Source: Ethnographic research by Prof. K. N. McKay

The ASAP Schema Is Modeled After Problem Solvers' Analysis of the "Out of the Ordinary"

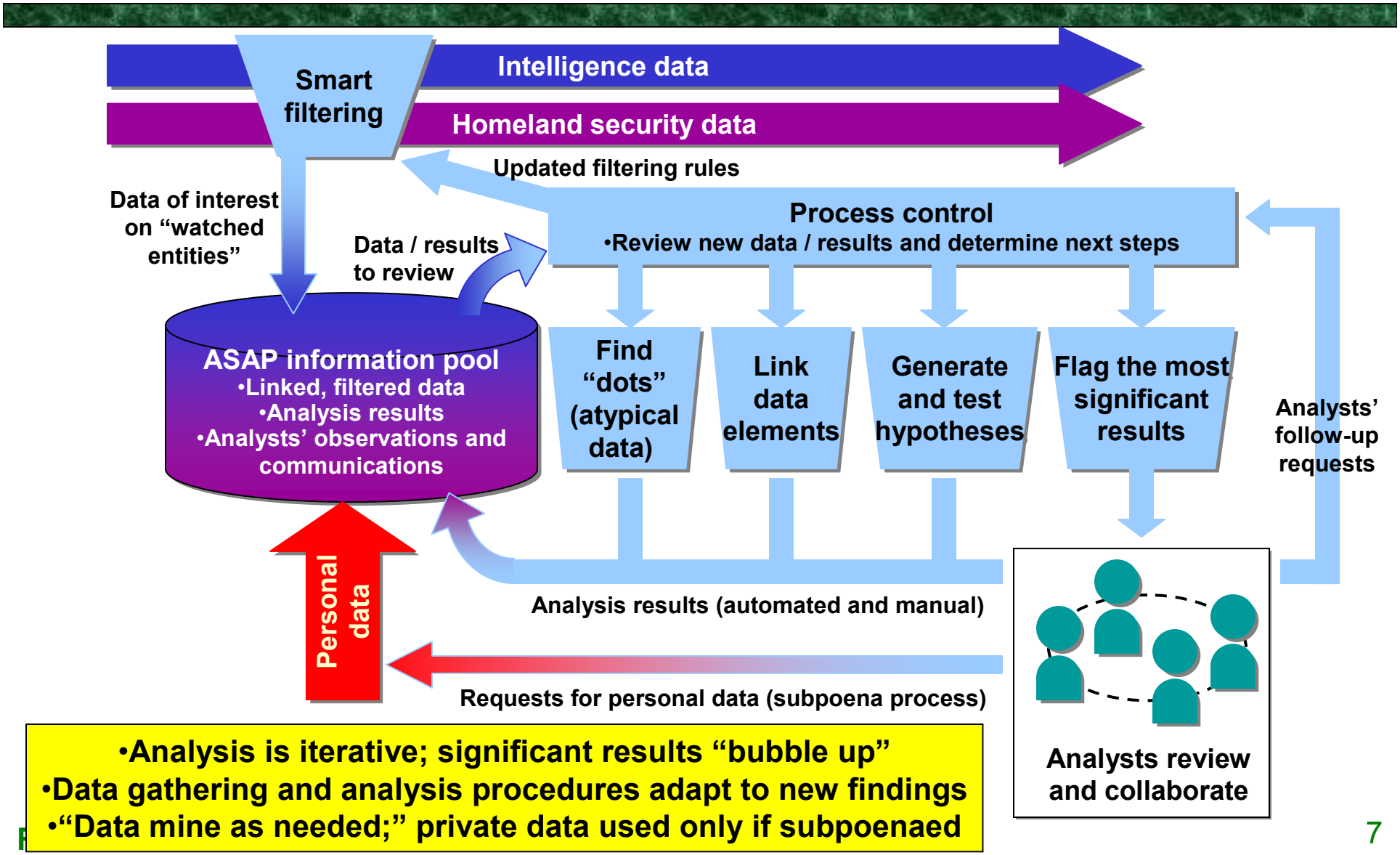


Improve awareness and collaboration for networks of intelligence collectors and analysts at all levels— national, strategic, operational and tactical

Observations on Research and Development Needs

- Initial research focus was going to be on algorithms development, but numerous tools carrying out specific tasks to "connect the dots" already exist
  - Text to structured data tools
  - Network analysis tools / link analysis tools
  - Data mining tools
  - Collaboration / groupware suites
- What was missing was putting them together into a single architecture
  - Led to focus on developing ASAP *architecture*, rather than specific algorithms or tools
- What is also needed is a basic suite of analysis and collaboration tools that can be deployed quickly and cheaply, and can be customized and tailored to the situation at hand
- "Artificial intelligence" research on developing models of the status quo, and developing algorithms to find outliers, also needed

The ASAP Architecture for Automated Support to Counter-Terrorism Analysis

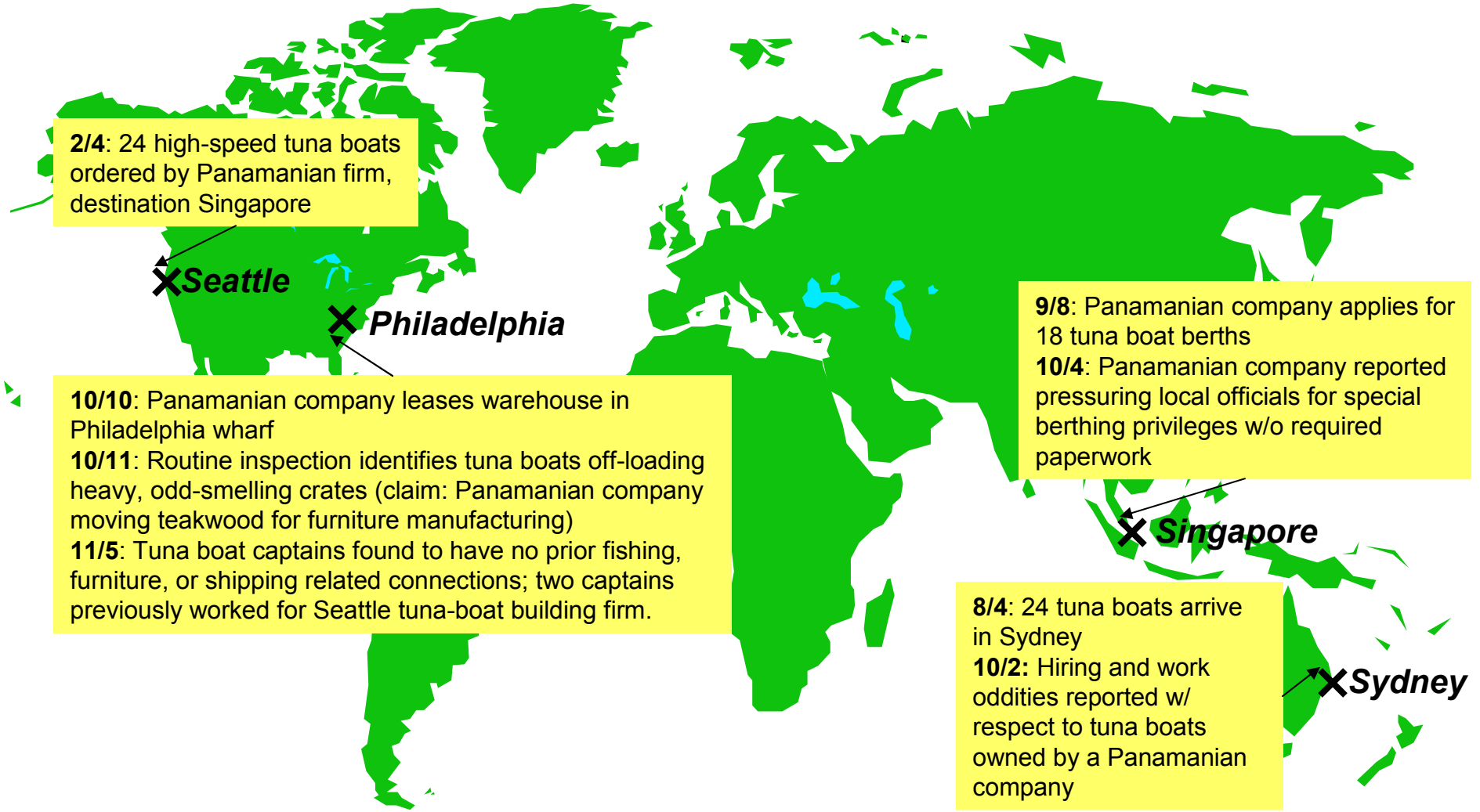


- Analysis is iterative; significant results "bubble up"
- Data gathering and analysis procedures adapt to new findings
- "Data mine as needed;" private data used only if subpoenaed

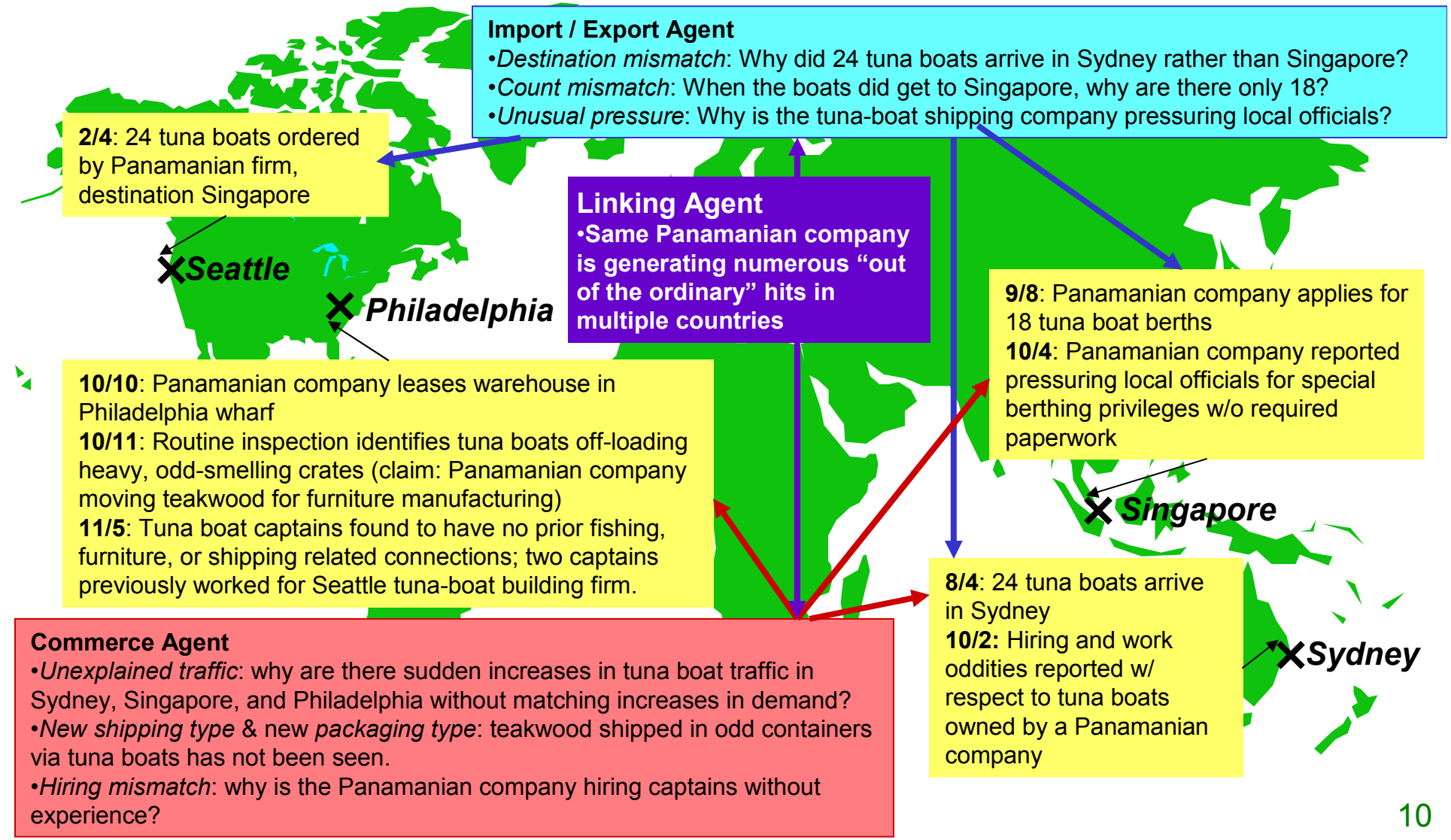
Major Features of the ASAP Architecture

- Identifies phenomena that are “out of the ordinary,” even if they do not match previous patterns for suspicious behavior
- Focuses on information already collected from intelligence and other government monitoring systems
  - In the acquisition / SE context, this would be acquisition and SE-related products, and any incoming complaints
- Employs rules allowing data to be analyzed in the context of existing knowledge
- Employs rules that are dynamic – significant developments automatically increase the priority of related information
- Explicitly deals with uncertainty by formulating and testing competing hypotheses
- Initiates collaboration of personnel needed to “connect the dots”
  - Analysts’ reviews are considered data just as much as source information

An Example Scenario: "Something Bad Happened on November 9th"



Applying the ASAP Schema to the 11/9 Scenario



What Can Be Done in the Near Term To Bring About ASAP Schema Benefits? (1)

Basic Collaboration Support:

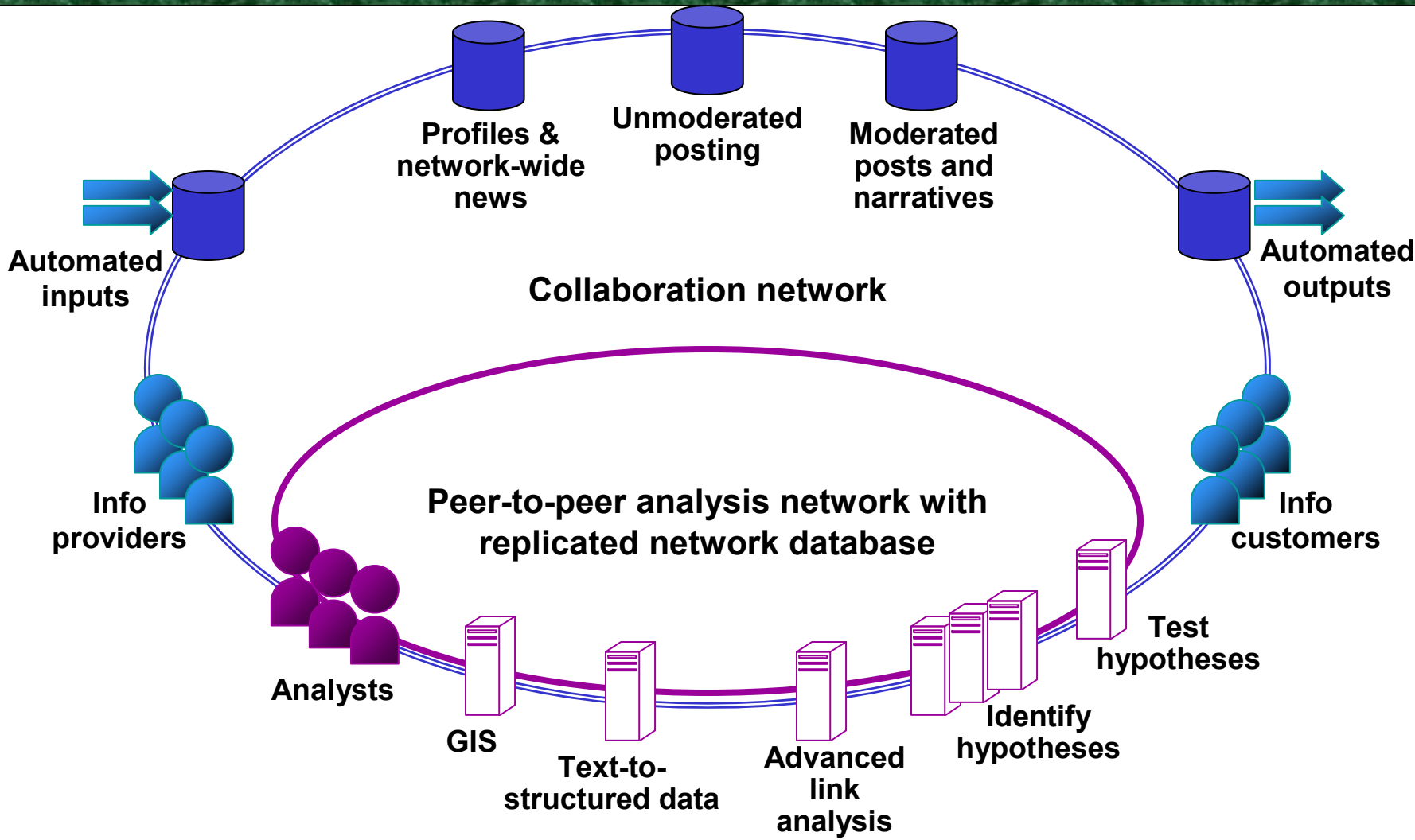
- Core functionality – improve collaboration
  - Create and distribute 2-4 page profiles of ordinary behavior, and examples of "significant" violations
    - Tight initial focus on training
    - Update regularly – "be on the look out for…"
  - Create a family of posting boards for the Community
    - Have unmoderated boards for posting "curious" or "unusual" information
    - Have moderated board for posting of "key" information on various potential problems
    - Have weblogs / wikis for presenting narratives
- Extended functionality – support analysis of posting boards
  - Allow Google-like search engines for posts
  - Use simple heuristics on posts, looking for connections and patterns across the posted messages

What Can Be Done in the Near Term To Bring About ASAP Schema Benefits? (2)

Analysis Support:

- Focus on evolutionary build-out of analysis tools, starting with existing products
- Core functionality – create and manage "the dots":
  - Network database / analysis tool that stores data in a replication-friendly and update-friendly format
  - Peer-to-peer system for performing the distribution and updating
    - Ideally, includes project management / workflow tools as well
- Extended functionality – automatically populate and analyze "the dots":
  - Free text / data to structured data tool
  - Advanced link analysis tool
  - Hypothesis generation tools (pattern-matching and pattern-violation agents)
  - Hypothesis tracking and testing tools

What Might a Near-Term ASAP Network Look Like?



How Might The ASAP Schema Be Applied to Defense Acquisition?

- Generate simple models of the "status quo" for acquisition programs and acquisition jobs
  - Describe major steps that the program or person performing the job should go through, and what each step should look like
    - Acquisition Culture, Successful Program Implementations discussions may be useful here
    - Exercise: "this quarter, the program does nothing problematic. What kind of things occur?"
- ID and prioritize diagnostic measures and metrics that can determine whether a program or acquisition job is "likely staying on track"
- Set up near-term implementation described on previous slides
  - Start by distributing models and diagnostic metrics, identifying key personnel who should be involved in the "ASAP for Acquisition" network, and identifying key sources of acquisition data
- Begin building a library of decision rules on how to flag information for further review, and what sorts of hypotheses to generate

For More Information…

- *Out of the Ordinary: Finding Hidden Threats by Analyzing Unusual Behavior* available at:
  - http://www.rand.org/publications/MG/MG126/
- POC: John_Hollywood@rand.org