BAA 05-44
Predictive Analysis for Naval Deployment Activities (PANDA)

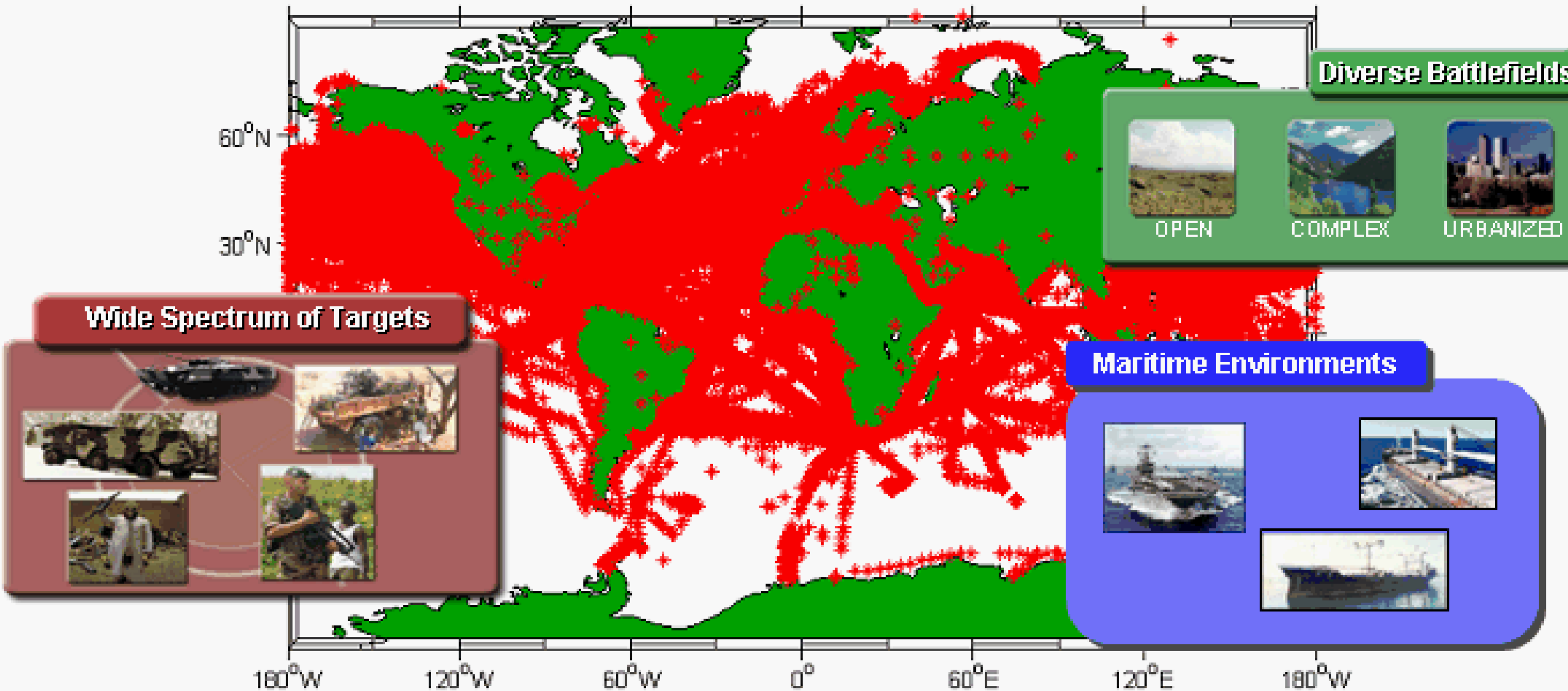
Briefing to Industry:
PANDA Overview

Dr. Kendra E. Moore
DARPA/IXO Program Manager

16 September 2005

PANDA Vision

Global-scale system providing anticipatory situation awareness on 50K+ entities
Motion-based change detection and adaptive context modeling
Wide-spread application; PANDA focus on maritime domain

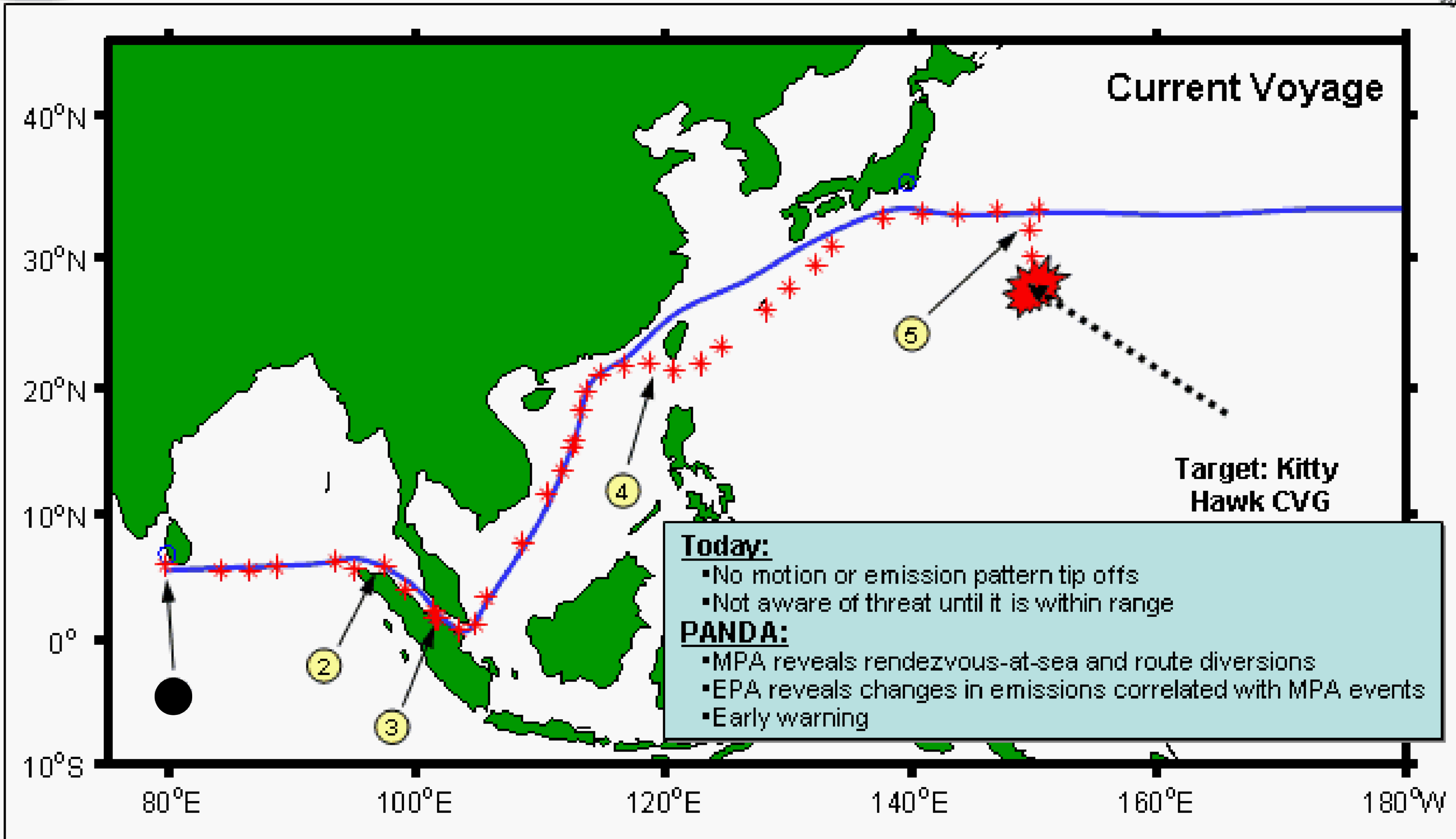


Today – even as scope has expanded, we use same old model: we look for what is already identified as "bad"
PANDA will enable us to automatically identify new threats

Operational Problem

Example: Merchant Vessel Attack on U.S. Navy CVG

PANDA



Departs Colombo 9-28-05 12:00

Hijacked 9-30-05 23:07

Rendezvous-at-sea 10-01-05 01:22 2 Hours

Route Diversion

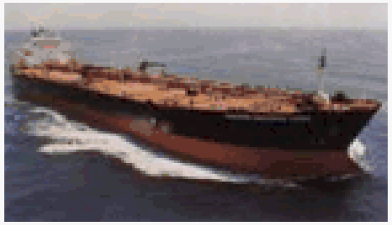
2nd Route Diversion

DARPA

Patterns Do Exist: Proof of Concept

- **Challenge**
  - Determine the percent of vessels that have motion-based patterns as evidenced in their tracks
  - 95% confidence level
- **Quick Look Study**
  - Supported by ONI
  - *Methodology*
    - Randomly selected 15 vessels
    - 12 months of track data
    - Visually reviewed ship tracks to determine if a pattern was present
  - Results: 13/15 of vessels have human-discernable motion patterns

- **In-Depth Study completed by CCRI**
  - *Methodology*
    - Population ≅ 50K vessels
    - Sample size to achieve 95% confidence level = 381 vessels
    - Randomly selected 422 vessels
    - Visually reviewed ship tracks to determine if a pattern was present
  - Results: **75-85%** of vessels have human discernable motion patterns
  - This study will be covered more in depth in the afternoon session



PANDA Concept

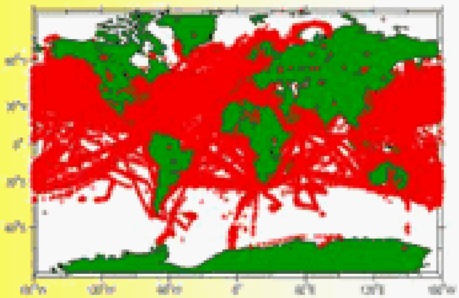
Maritime Situational Awareness

PANDA

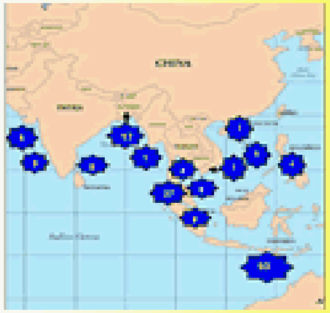
- **PANDA system will provide anticipatory situation awareness for 50K+ vessels**
- PANDA will go beyond the vision of merely tracking to perform motion-based change detection and infer intent
- PANDA will
  - Learn motion-based activity patterns from long-duration tracks and correlated data
  - Detect deviations from these patterns
  - Learn and apply context models to predict potential threats even as patterns evolve

CLUTTERED operational environment prevents effective data analysis…

…Especially in geographic regions that pose very HIGH RISK.



Sample "Ocean Watch" Data for One Month



Piracy & Hijacking (2004)

Operational Challenges

- Broad and diffuse threat
- Highly cluttered environment (126K surface vessels > 100 tons; >50K merchant vessels)
- Highly variable environment (behaviors and updates)
- "Stovepiped" data sources prevent cross-organization analysis
- Use of global and local sensors complicates data fusion

Operational Needs

- Exploit all-source data to maintain a complete and continuous surface picture
- Classify vessels in terms of what they are doing in addition to where they are
- Predict where vessels will be 6, 12, and 24 hours out to support Naval operations
- Detect potential threats to support interdiction activities

DARPA

Military Participation & Validation

PANDA

Data/Evaluation & Experimentation

MAST/DSOC

ONI

TMFC

Demonstration

ONR → NNWC
(NUWC & NWDC)

NAVSEA
WARFARE CENTERS
NEWPORT

Acquisition

ONI

ONR → PEO C4I&S

NGA

Execution

SPAWAR

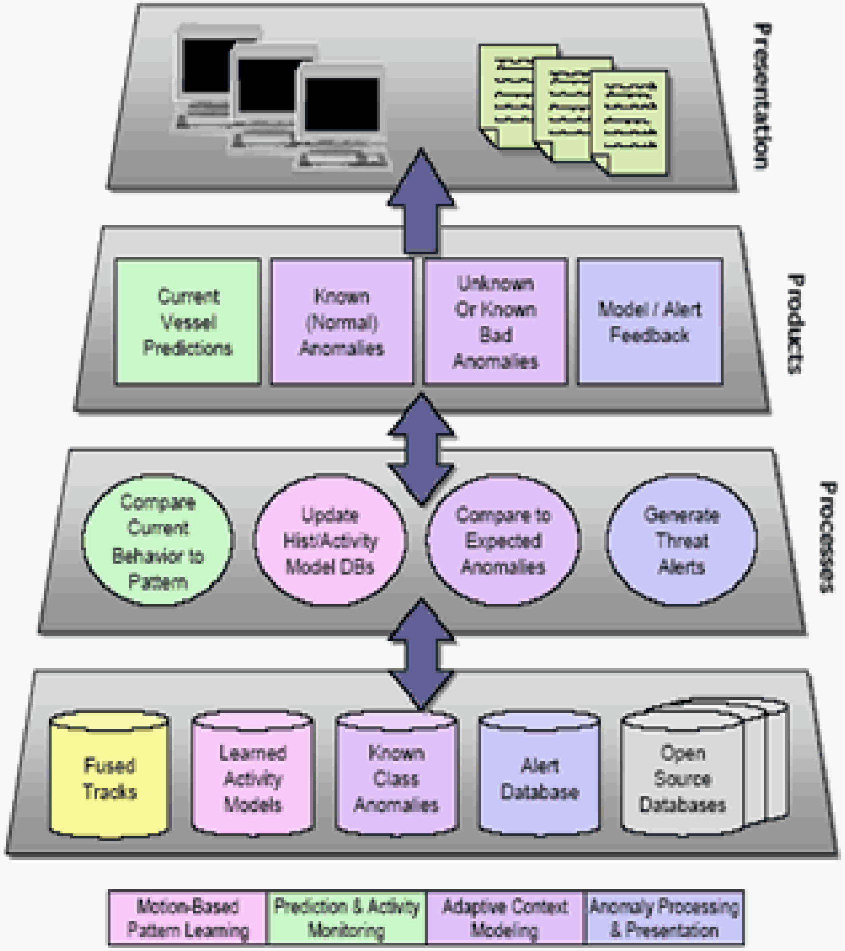
SPAWAR
Systems Center
San Diego

NRL

ONR

DARPA

Key System Elements



- Motion-Based Pattern Learning
  - Learn activity patterns from long-duration tracks and correlated activity data
  - Technologies:
    - Hidden Markov Models
    - Neural Processing
    - Machine Learning
- Prediction & Activity Monitoring
  - Leverage models of normal behavior to predict future behavior and detect anomalous behavior
  - Technologies
    - Boundary Detection
    - Bayesian Methods
    - Conceptual Spaces
- Adaptive Context Modeling
  - Represent data and business rules for resolving expected anomalies
  - Technologies
    - Case-based reasoning
    - Conceptual Spaces
    - Ontologies
- Anomaly Processing & Presentation
  - Assimilate and convey alerts using known and predicted tracking information integrated with context
  - Technologies
    - Agent-based technologies (Fast C^2AP)

Motion-Based Pattern Learning

- **Purpose**
  - Learn activity patterns from long-duration tracks and associated activity data
- **Challenges**
  - Multiple dimensions
  - Individual vessel and class-level (each vessel can be in more than one class)
  - Rapidly learn & relearn on few exemplars
  - Local (densely sampled) and global (sparsely sampled) learning; sample from overlapping data sets
- **Existing Work**
  - Transactional profiling (e.g., normal spending patterns): lots of data, but low dimensionality; assumed statistical independence; lots of training data
  - Motion pattern learning (self-organizing maps, neural computing): current focus has been on individual ports
- **Potential Metrics**
  - Model accuracy as a function of training set size
  - Coverage of modeling approach (range of vessels and types of patterns that can be learned with a specific approach)
  - Speed of learning

Prediction & Activity Monitoring

- **Purpose**
  - Apply models of normal behavior to current motion-based activities
  - Predict future behavior and detect anomalous behavior
- **Challenges**
  - Multiple dimensions
  - Multiple resolution prediction (local and global prediction)
  - Context-specific anomalies:
    - Size of anomaly a function of geography and presence of other vessels
    - Type of anomaly a function of location in motion-based pattern
- **Existing Work**
  - Activity monitoring in stock & cell phone fraud: looking for known patterns of bad behavior vice deviations from normal behavior
  - Pattern recognition
- **Potential Metrics**
  - Prediction accuracy
  - Anomaly detection accuracy
  - False alarm rates

Adaptive Context Modeling

- **Purpose**
  - Efficient & flexible representation of data and business rules for resolving expected anomalies
- **Challenges**
  - Multiple context/multiple dimensions
  - Constantly evolving business rules (tax laws, spot markets, regulations)
  - Need to update existing cases to maintain a manageable set of cases
- **Existing Work**
  - Ontologies: computationally inefficient, difficult to evolve, requires on-going knowledge engineering
  - Case-based reasoning: large number of pre-resolved cases, typically small focus areas, fast
- **Potential Metrics**
  - Accuracy of learning new cases as a function of the number of new anomalies and their resolution
  - Ease of incorporating new models
  - Speed of searching existing case base as a function of the size of the cases
  - Accuracy in matching new cases to existing cases or identifying a new case as truly anomalous

Anomaly Processing & Presentation

- **Purpose**
  - Utilize context models and search agents to process and resolve anomalies as model change events or possible threats
- **Challenges**
  - Efficient matching to alternative cases and selection of resolution plan
  - Efficient use of resources for conducting searches
  - Clear and concise presentation / explanation of alerts
  - On-line capture of feedback to Context Modeling
- **Existing Work**
  - High-level reasoning methods applied against ontologies, case bases, and knowledge bases
  - Advanced display technologies
  - Agent-based technologies
  - Services-oriented computing architectures
- **Potential Metrics**
  - Ease and speed of use within and across PANDA nodes

Systems Engineering & Integration

- **Purpose**
  - Manage and support technical components with framework for interactions
  - Tie technical components together into a real-time system
  - Interface with external systems at geographically separated government test sites
  - Lead PANDA transition activities
- **Challenges**
  - Demonstrate increasing scalability, utility, and computational efficiency as the program progresses
  - Function at multiple sites, with different data streams, security requirements, and foci for situation awareness
  - Achieve efficient and secure exchange of data and models across nodes
  - Support rapid addition and removal of data sources and working with legacy and new presentation systems
- **Potential Metrics**
  - Processing Latency (time to process that request and deliver an answer to the user)
  - Usability (user training time)
  - Maintainability and Extensibility (time to revise, debug, and upgrade major components)
  - Installation (time to install system in a new operational location)

Performance Evaluation

- Purpose
  - Provide objective evaluation of progress towards system goals
  - Generate insight for component developers regarding their respective impacts on performance
- Challenges
  - Interface with government facility personnel and assist in the execution of tests and evaluations
  - Develop a performance evaluation plan for the execution of these tests
  - Support the SEI contractor and technology developers in developing their test plans
  - Review proposed metrics and evaluation plans
  - Review self evaluation results
  - Provide feedback
- Potential Metrics
  - Scale (process incrementally increasing numbers of vessels and information about those vessels)
  - Detection (detect increasing data anomalies and increasingly complex anomalies)
  - False Alarms (reduce false alarm events generated as vessels and detections increase)
  - Classification (increase accuracy of anomaly classification)
  - System Abilities (discover and learn correlated activities, learn new models, provide timely alerts, leverage future detection/tracking capabilities to expand system to smaller vessels, and harden system for full-scale transition to military customers)

High-Level Metrics

PANDA

Phase	Challenges	Evaluation Criteria
1 *18 months*	•Demonstrate that motion & emission patterns can be learned •Discover & learn correlated activities (tankers, single-owner liners) •Install three individual nodes	•Discover & learn individual patterns for 1000 vessels •Detect and classify 70% of anomalies •Learn patterns on <20 exemplars •False Alarms: <1%
2 *12 months*	•Rapidly relearn models in response to sudden changes and generate timely alerts •Integrate with two nodes simultaneously •SeaTrial Demonstration	•Discover & learn individual patterns for 10,000 vessels •Detect and classify 80% of anomalies •Learn patterns on <10 exemplars •False Alarms: <1%
3 *12 months*	•Learn patterns on small number of exemplars •Integrate with three nodes simultaneously •SeaTrial Demonstration	•Discover & learn individual patterns for 100,000 vessels •Detect and classify 90% of anomalies •Learn patterns on <6 exemplars •False Alarms: <1%
4 *6 months*	•Incorporate smaller vessels •Harden, scale and transition system •SeaTrial Demonstration	•Discover & learn individual patterns to global traffic •Detect and classify 95% of anomalies •Learn patterns on <6 exemplars •False Alarms: <1%

DARPA

Evaluation Criteria (1 of 2)

- **Overall technical and scientific merit, innovation, and depth**
  - Understanding of the current state of the art in the technologies proposed
  - Soundness of technology approach at the component level and systems level
  - Degree of innovation and potential for revolutionary advance
  - Justification of design choices as compared to alternative techniques
- **Relevance to PANDA mission objectives and program concepts**
  - Understanding of the depth and breadth of PANDA missions and propose an innovative solution according to PANDA objectives
  - Understanding of PANDA program goals and metrics
  - Familiarity with current and previous work on related systems
  - Awareness of current and projected support requirements to analysts and operational commanders
  - TDs: commitment to comply with decisions of the SE&I contractor, including collaboration with other TDs
  - SE&I: plan for coordinating the activities of the TDs and for working with the PE
  - Precision and coverage of the proposed effort's metrics and evaluation plan

Evaluation Criteria (2 of 2)

- Personnel and corporate capabilities and experience
  - Qualifications of proposed technical personnel
  - Availability of personnel for the duration of the contract
  - Proposer's experience related to the proposed technology area
  - Ability to collaborate in off-site integration and experimentation
  - Adequacy of proposed facilities
  - Adequacy of security plan
- Proposed schedule and metrics per Phase
  - Consistency of detailed schedule with overall PANDA program plan and goals
  - Consistency of specific proposed metrics with overall PANDA evaluation criteria
  - Likelihood that proposed detailed schedule and specific metrics contribute to achieving increasingly aggressive program performance goals while reducing risk
- Project and technology transfer plans and potential transformative impact on military operations
- Cost realism and value of proposed work to the Government
  - The total cost relative to benefit
  - The realism of cost levels for facilities and staff
  - The cost-effective use of existing equipment and software
  - Competitive costs on procurements

Metrics and Schedule

- The metrics and schedules outlined here are high-level and must be followed for award consideration
  - Note: These metrics have been corrected since the BTI. This is the correct version.
- More detailed metrics and schedules are expected in proposals
  - Detailed metrics to better evaluate progress
  - Detailed schedules with intermediate milestones
- Proposals should include an explanation of how detailed metrics and schedules will align and integrate with the program's high level goals

Seedling Activity: Trident Warrior

- Prototype Tools and Display Configurations
  - Facilitate visualization and understanding of information linkages
  - Examine existing tools and software, develop tool and display storyboards, and acquire feedback from SMEs and users
  - Will provide user-interaction and highlight concepts to facilitate collection of user feedback
- Scenario Events
  - Develop vignettes designed to highlight PANDA concepts
  - Events will be appropriate for incorporation within the larger TW06 scenario framework
- Preparation for full integration with TW07
- Performer: Pacific Science and Engineering (PSE)

Seedling Activity: Patterns in Tracks

- In-Depth Study to prove pattern existence
  - Results: **75-85%** of vessels have human discernable motion patterns
  - This study will be covered more in depth in the afternoon session
- Continuing work to develop a tagged set of at least 1000 tracks to be leveraged by PANDA
- Performer: Commonwealth Computer Research Inc (CCRI)

Questions?

- Please fill out a Q&A card
  - Your question will be answered all together during the Q&A portion of the morning

  AND/OR

  - Your question will be posted in the PANDA FAQ online
- Please review the online FAQ periodically, as it will continue to be updated