

# An Overview of the Cognitive Seldon Toolkit and Proposed Methods of Interfacing with RIPPLE

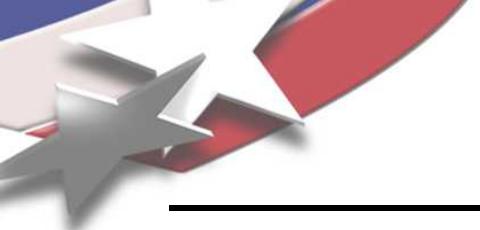
## Presentation to the 3<sup>rd</sup> System of Systems Conference

National Institute of Standards and Technology  
Gaithersburg, Maryland  
December 10, 2008

Eric Parker, Principal Investigator  
Sandia National Laboratories  
[eparke@sandia.gov](mailto:eparke@sandia.gov), 505-284-4325

Sridhar Natarajan, Principal Technical Manager  
Concurrent Technologies Corporation  
[nataraja@ctc.com](mailto:nataraja@ctc.com), 814-269-6220

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.



# Outline

---

- **Overview of Work**
  - **RIPPLE Overview**
  - **Cognitive Seldon Overview**
- **Results**
  - **Discovery Task (“Phase 1”)**
  - **Road Mapping Task (“Phase 2”)**
- **Discussion**



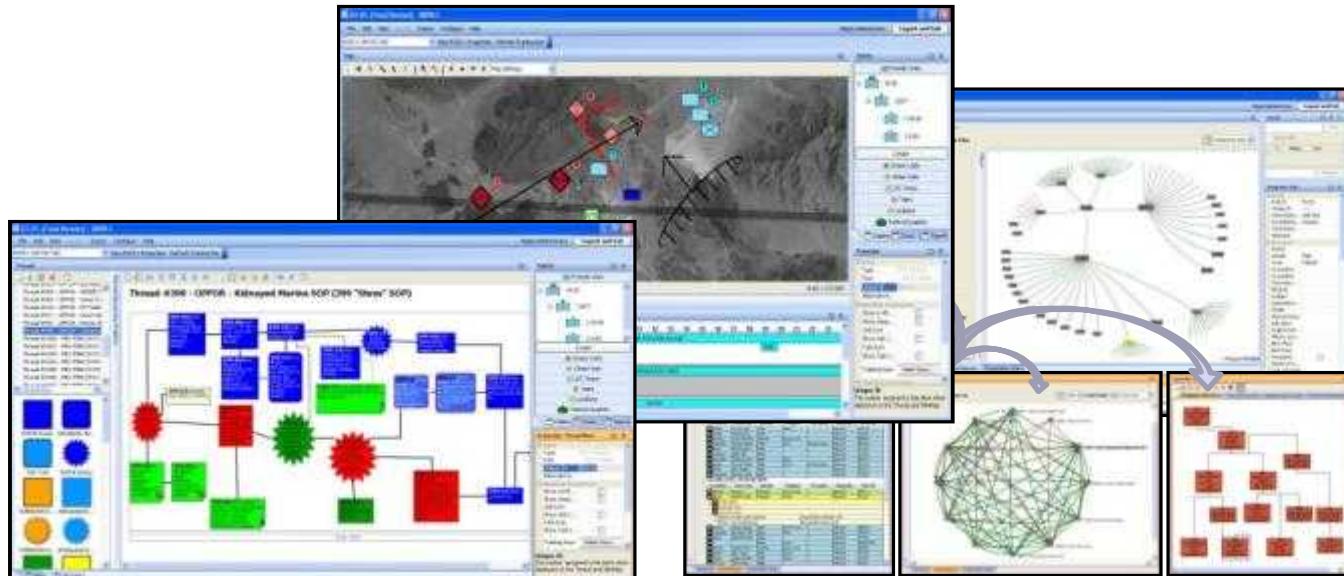
# Overview of Work

---

- **System of Systems Interoperability**
  - 10+ years of SANDIA's Research in Cognitive Modeling – "Cognitive Seldon"
  - 7+ years of DARPA initiated and Army/Marine Corps Funded Work – "RIPPLE"
    - Reactive Information Propagation and Planning for Lifelike Exercises
- **Scope:**
  - Phase 1: designed to be a limited-duration discovery and scoping activity. The main goal was to determine if integration of Cognitive Seldon and RIPPLE was
    - Desirable
    - Feasible
- **Approach:**
  - Established common understanding of vocabulary
  - Exchanged understanding of existing RIPPLE and SNL social simulation capabilities.
  - Conducted In-depth code exploration: collaboration between SNL and CTC developers/programmers to allow both SNL and CTC to determine
    - System "Hooks" and development necessary to link codes
    - How much the current codes would have to be adapted to provide those links.
- **Accomplishments**
  - Determined a service-based linkage between Cognitive Seldon and RIPPLE
    - Should allow both CTC and SNL to Publish/Subscribe data
    - Data to be exchanged as lists of comma separated values
    - Code can be exchanged using the CTC SharePoint site
  - Produced Phase 2 proposal based on three potential use cases developed with input from the Government client (Marine Corps).

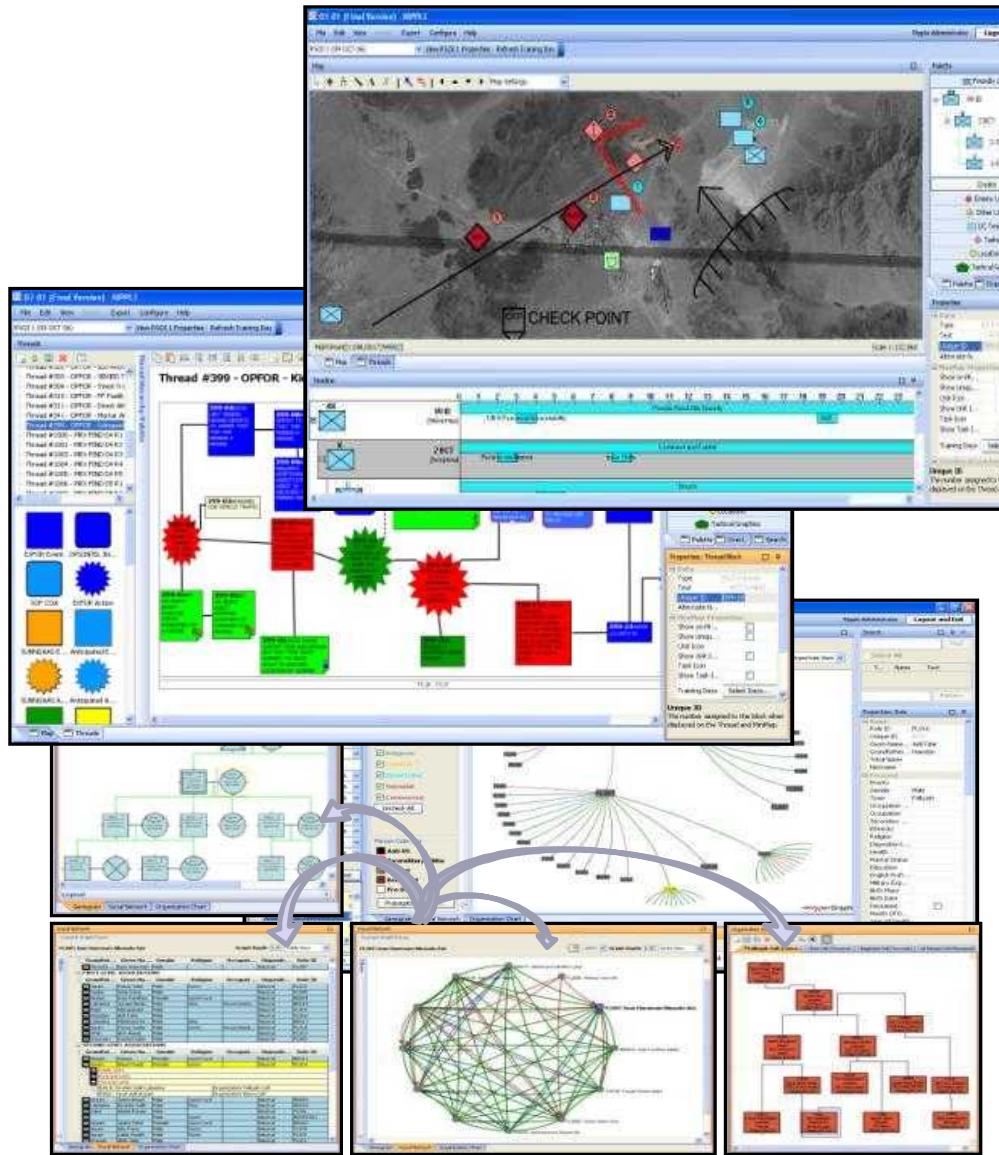
## Reactive Information Propagation and Planning for Lifelike Exercises

- Holistic exercise planning and control application for full-spectrum field training from squad to brigade level.
- Combines best training practices from Army and Marines to propagate training expertise to home station and remote sites.
- Capable of client-server network or stand-alone laptop operations. Also capable of supporting mobile device (PDA) and web access.
- Five modules: Planning, Threading, Role Player Management, Reports, Assessment. All modules share an integrated data model so one change “ripples” across all modules.
- One stop shop for all training exercise needs.

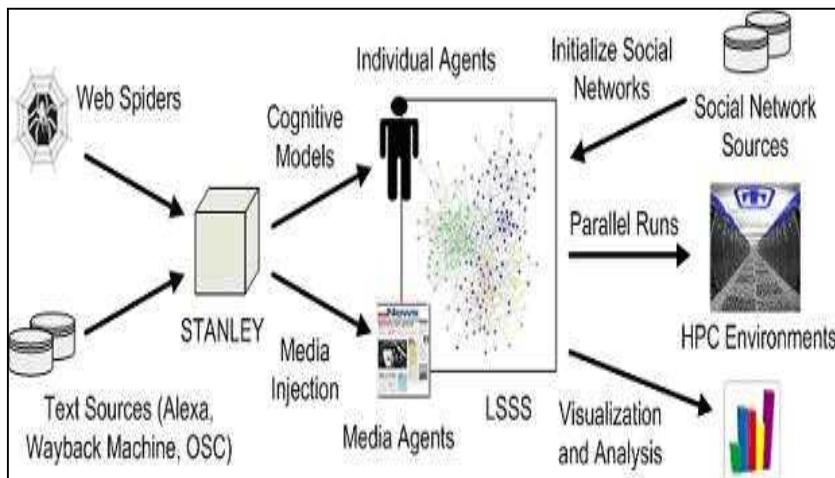


# RIPPLE: Five Modules

- Planning View
  - Plan/record events on map and auto-populate synchronization matrix
- Role Player Management
  - Track relationships between role players and organizations
- Threads and Injects
  - Plot chains of events involving complex human interactions
- Customizable Reports
  - MS Office compatible
  - Constructed to fit each system
- Assessment
  - Training objective development through AAR support
- All modules share an integrated data model so one change “ripples” across all modules.



# Large Scale Social Simulation for Human Behavior Modeling – SNL Internal R&D



## Purpose, Goals and Approach

- Problem: The importance of thoughts & opinions of individuals & groups means greater understanding is required for how information is disseminated across a society to form and transform thoughts and opinions.
- Goal: Aid analysts & policy makers understand how propagation of information may be perceived
- Approach:
  - Use **STANLEY** to collect & process text
  - Translate media information into cognitive models
  - Integrate cognitive models into Seldon
  - Parallelize the combined system

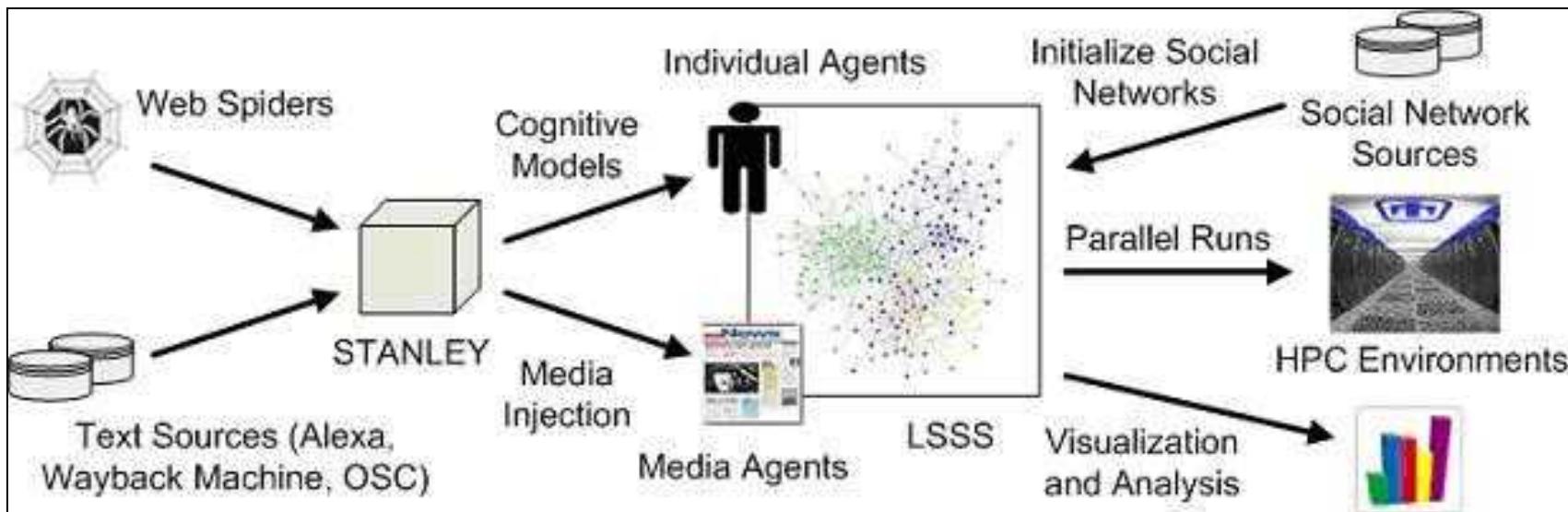
## Key Accomplishments

- Development of an HPC compatible cognitive model
- Parallelization of Seldon & Deployment on cluster
- Integrating two dissimilar technologies (cognitive models & agent-based simulation)
- Running large simulation (100K agents) using heavy-weight (due to cognitive models) realistic agents
- Development of HPC compatible tools for tracking information flow
- Incorporation of the 5-factor personality model
- Modeling test case underway: shift in public opinion in Jordan after November 2005 bombings

## Significance of Results

- The capabilities developed by this project can help not only analysts, but also have applications in Iraq and Afghanistan
  - **Course of action analysis (determine the best way to release news of an event or action)**
  - **Training – help new Commanders get a feel for how an area they are assigned to may react to news of specific events**
- What is the impact/legacy of this project?
  - **This project provides a testbed for other projects needing to use or seeking to study realistic social networks.**
  - **Further work may include developing IP by**
    - **Adding cultural factors**
    - **Adding economic agents**
    - **Follow-on projects**

# Cognitive Seldon Architecture

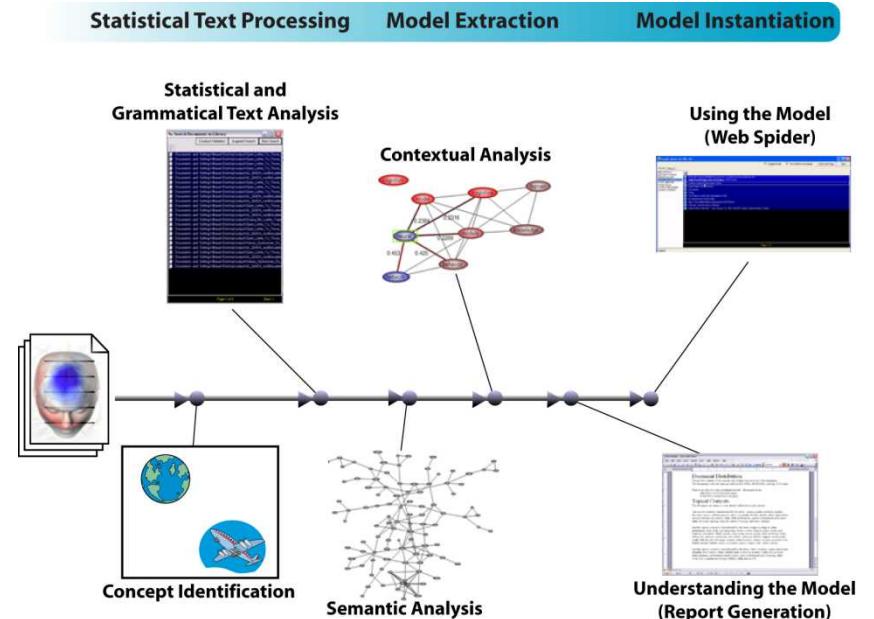


## Our Approach

- Use STANLEY to collect and process internet-based media text (i.e. newspaper articles, blogs)
- Translate that media information into the Cognitive Foundry's cognitive models using the Pattern Analyzer function of STANLEY
- Integrate the cognitive models into the Seldon social network simulation
- Parallelize the combined system to study how realistically-sized populations respond to information spread from media sources.
- Run Cognitive Seldon simulations

# STANLEY Library

- STANLEY is the Sandia text analysis extensible library
  - Text analysis library designed for building and utilizing models of individuals from text that they produce or access
- Main Features
  - Reusable library that programmers can implement into new applications.
  - Flexible implementation of text analysis algorithms
  - Identification of key concepts, relationships, and contexts from text documents.
  - Web spider for evaluating new information based on a previous analysis.
  - Generation of reports in MS Word describing key characteristics of an individual based on the text analysis.

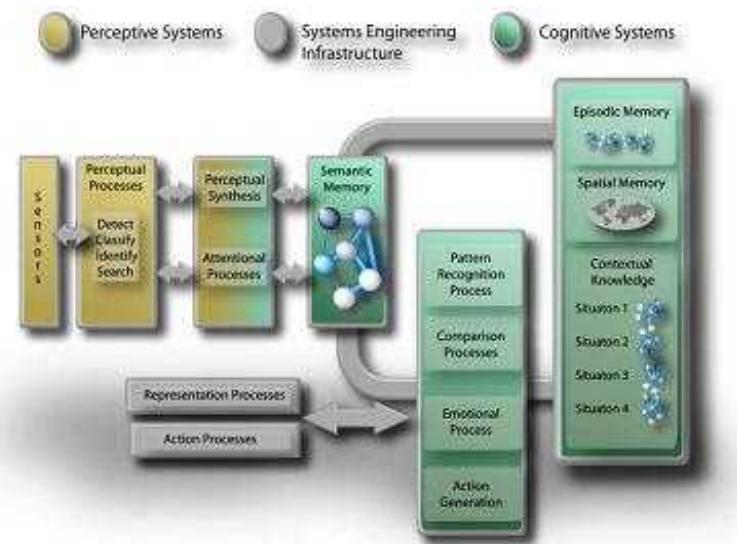
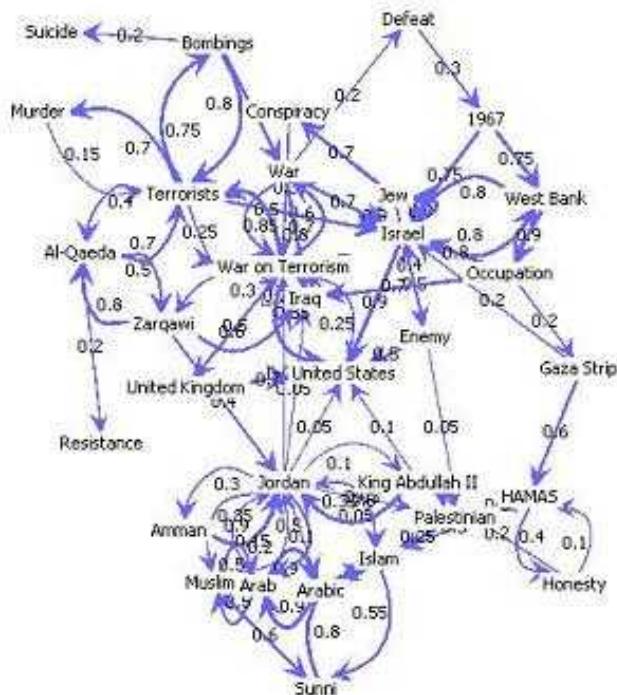


## • Example Applications

- Any application where one wants to integrate Sandia's cognition text analysis capability into some new application.
- Finding new relevant documents
- Understanding how a person or source has changed across time.

# Cognitive Foundry

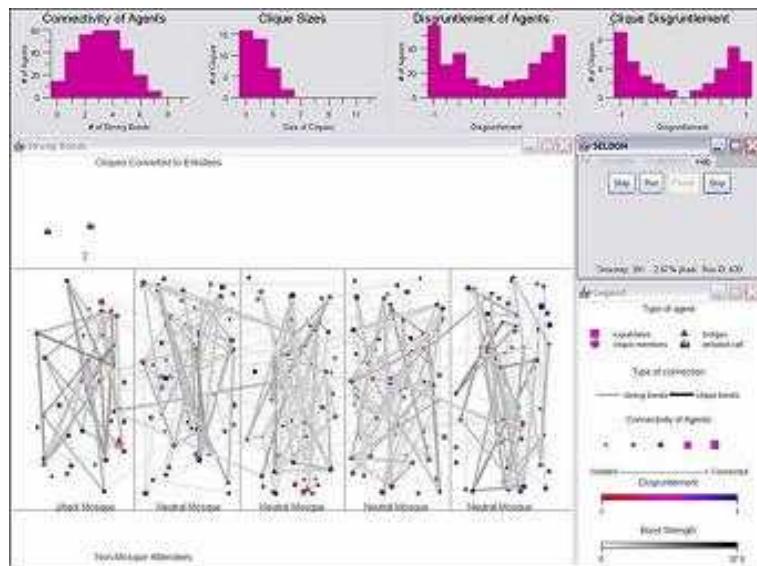
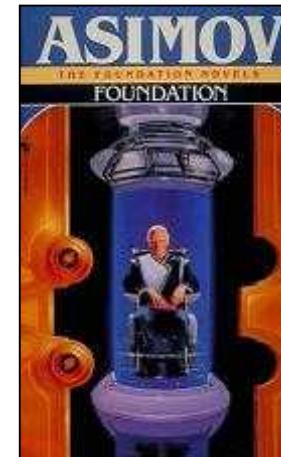
- The Cognitive Foundry is used to create psychologically plausible cognitive models
- The cognitive model is a semantic graph of concepts and their relationships



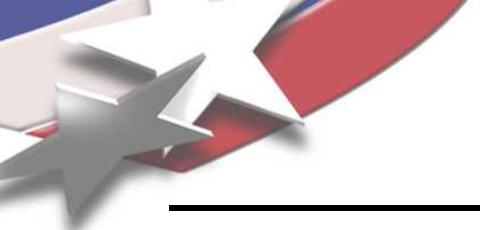
- Information spreads throughout the graph when activated
- Cognitive models are created automatically from text-based documents
- A lightweight version of the foundry facilitates running in a parallel environment

# Seldon

- An agent-based social simulation toolkit that combines technology and concepts from a variety of different fields
- Name comes from Hari Seldon in Isaac Asimov's *Foundation* stories
- Has been used to study inner-city gang and terrorist recruitment



- An agent can be an individual or a collective (i.e. school, mosque)
- Seldon allows for dynamic social network evolution
  - Agents interact according to exchangeable rule sets
  - The interactions build relationships and form social networks

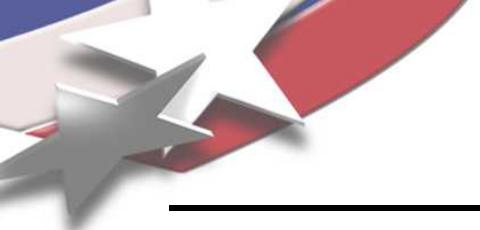


# Cognitive Seldon

---

- Cognitive Seldon integrates Seldon and the Cognitive Foundry
  - Each individual agent has a cognitive model that's used to process new information
- Goals:
  - Use cognitive models for a more realistic representation
  - Study how media affects societies
    - Cognitive models are generated automatically from text corpora
    - Media agents inject additional cognitive information during the simulation
  - Run large-scale simulations ( $O(10^6)$  agents) that represent entire populations – parallelization!





# Discovery Task: Findings

---

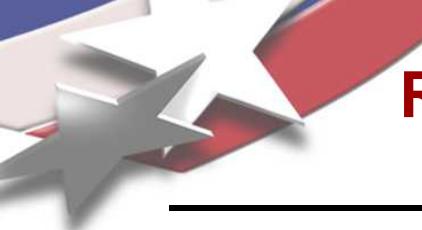
- As a result of reviewing documentation, examining code, and exchanging/answering lists of questions we determined that linking the Cognitive Seldon and RIPPLE packages was viable.
  - Use cases involve sequential exchange of information
  - A service-based linkage (such as a federation) is sufficient, rather than a hard-coded link
    - Should allow both CTC and SNL to Publish/Subscribe necessary data
  - Data to be exchanged as lists of comma separated values
- The lists of functional gaps and necessary changes produced for both RIPPLE and Cognitive Seldon should be addressed to facilitate interactions & follow-on projects



# RIPPLE Functional Gaps/Extensions to support end users that can be filled by Cognitive Seldon

---

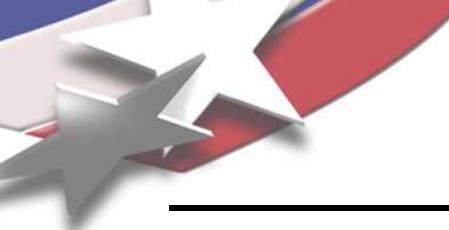
- Role information has to be entered by hand by experts setting up training
  - Very time-intensive process
  - Requires a substantial amount of background information and knowledge
  - If Role information could be generated automatically from documents written by and about the population of interest, it would
    - Reduce the amount of time required to setup the scenarios
    - Allow for more time to actually perform the exercises and analyze results.
- Rotation validation in RIPPLE is left entirely to the user
  - When live exercises can't be performed due to time or scheduling constraints, the scenario planner has to make a judgment call on how the rotation should proceed
  - Rotations would be much more realistic if scenario planners had a tool that could evaluate different options and inform them which choice was most likely to be correct.
- There are no tools in RIPPLE to support/validate an after-action review
  - Once users decide on a particular course of action, they have no means to see how their results would have changed had they made different decisions
  - It would be helpful to have a tool that compared the user's choices with alternatives choices



# RIPPLE Technical Gaps related to supporting cognitive/behavioral M&S

---

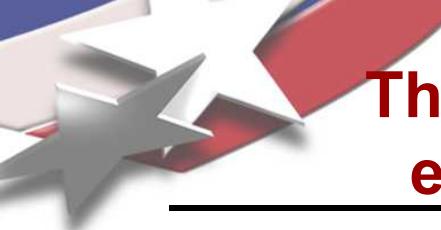
- Visualization – needs to be updated to reflect the strength of the links in the social model
- RIPPLE injects / media injects and agent actions
  - Unstructured format – needs to be converted to a vector of concepts
  - These are the attributes that need to be added to RIPPLE
    - There are regular expressions, but we probably need more
    - For Role data RIPPLE needs to support the data for Seldon parameters as well as information for cognitive models
    - Give weights to shapes – i.e. assign default weights to specific action shapes in the threads
    - Decision points provide different paths – who is going to parse the threads to produce the individual paths?
      - These would form batch jobs of Cognitive Seldon simulations
  - Need evaluation criteria / metrics (attitude, etc.)



# SNL CS&T tool Technical Gaps related to supporting RIPPLE

---

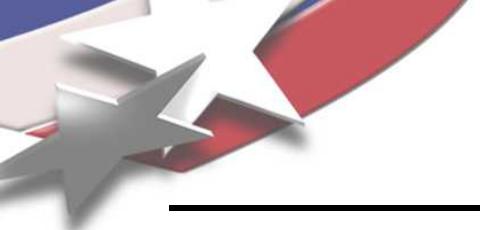
- Need the ability to represent negative (antagonistic) relationships
- It is very desirable to be able to do entity & relationship extraction
  - From existing metadata
  - From Raw Text
- Hook to GIS analysis capabilities
- Would like custom serialized output
  - Entities & attributes, entity relationships, and relationship attributes
  - Format compatible with RIPPLE
- Need Custom filters
  - Filters for cognitive model generation (domain specific, input and output)
  - Ability to handle unusual formats (non-text)
- Need to Support Multi-core platforms
- Need the ability to force events to happen at a specified time
- Delete individual agents during the course of the simulation
- Agents with greater influence (larger relationship energy?)
- Improving the default values for Seldon & Cog Foundry
  - Probability values for interacting with agents
- Organizational behavior –
  - how does an organization respond when key people are deleted?
  - How does an organization form/grow/die?
- Hooks into the model database
  - Formatting output and providing access
- Initialize social network with real data.



## **Three possible integration scenarios have been envisioned for RIPPLE and Cognitive Seldon**

---

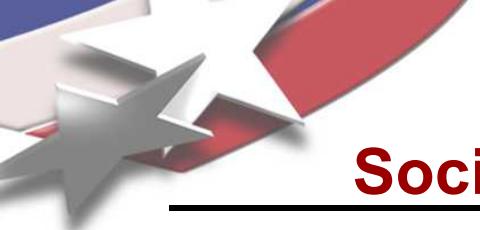
- Automatic generation of 'starter' collections of role-players and their relationships from raw unstructured text sources
- Validation of thread models constructed in RIPPLE using Cognitive Seldon simulations
- After action review of simulated alternative approaches and outcomes as compared to actual actions taken by trainees



# Use Case 1: Automatic Generation of Role-Players

---

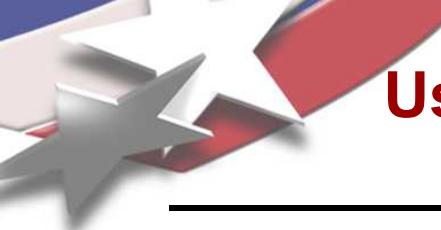
- Scenario: Planner wishes to obtain an automatically generated ‘starter’ collection of role-players and their relationships, constructed automatically from source documents and presented in the RIPPLE Intel View
  - Sandia tool ingests data from a variety of information sources, such as:
    - Media
    - USMC/US Military Intel documents from the theater
  - Sandia tool can also gather new data from web spiders
  - SNL tool mines role player names, organization names & types, and networks (mutual relationships among people and organizations) from ingested data. SNL tool also mines biographical incidents (“biography snippets”) associated with each role player and any other attributes (e.g. tribe name, sunni versus shia, geographical location, etc.)
  - SNL tool outputs models of role players, organizations and networks to RIPPLE.
    - These models are characterized by: role player names, biography snippets and other role player attributes, organization names, associations (person-person, person-organization, organization-organization), and association types (municipal, criminal, family, etc.) and any other attributes (subordinate vs superior, father vs son, etc.)
  - RIPPLE creates role players, organizations, organization structures and social networks.
  - RIPPLE populates role player biographies with the biography snippets as well as any other role player attributes determined by the SNL tool
  - Scenario planners flesh out role players, organizations, organization structures and social networks in greater detail for use in live exercises through RIPPLE Intel View.
  - SNL tool could refine the various models and periodically send updates to RIPPLE
- While not all of these capabilities currently exist, the text processing and social simulation tools form a promising basis from which to begin development
  - The result would save training experts from having to fill in the role information manually, and reduce the overall time required to conduct an exercise.



## Use Case 2: Social Simulation Validation of RIPPLE Scenario

---

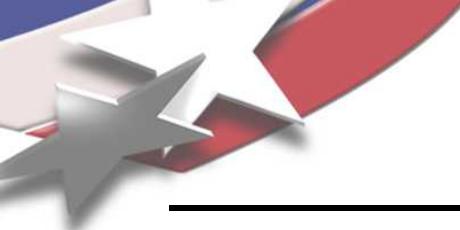
- Scenario: planner(s) wish to validate expected outcomes of social models constructed in the RIPPLE Intel View as they play out according to sequences of events in scenarios constructed in the RIPPLE' Planning View.
  - Scenario planners construct role players and relationships in RIPPLE's Intel View.
  - Scenario planners build thread (causal chain of events leading to an undesirable action that the exercise force must prevent) models in RIPPLE's Planning View.
    - Each thread is associated with one or more training objectives.
    - Threads have alternate paths defined by 'forks' and 'joins' in the flow of events that occur at the inject points.
  - RIPPLE exports the thread models (Planning View) and social model (Intel View) into Cognitive Seldon.
    - The models are defined by actors (role players, units, organizations, etc.), events, and anticipated (planned) outcomes.
  - Cognitive Seldon simulates the threads, and determines possible outcomes (e.g. attitudes of individual persons, attitude of entire town, state of specific organizations such as formation of new insurgent networks and/or disruption of existing ones, etc.) for each thread event.
    - Compares the simulated outcomes with the anticipated outcomes and provides a measure of validation for the threads.
  - SNL tool exports the measures of validation for each thread at various levels – overall, at each flow decision point (fork/join), and for each event.
  - RIPPLE maps the measures of validation for each thread to the exercise objectives and provides a measure of effectiveness for each thread (vis-à-vis exercise objectives), flow decision point and each event.
  - Planner refines the thread to improve effectiveness vis-à-vis exercise objectives
- When the setup process for Cognitive Seldon is complete, trainers and other users can run the simulation and use it to explore how the target population will respond to information released through media agents.
  - This can be used for course of action analyses that can provide trainers with feedback as they set up new exercises.
  - Cognitive Seldon simulations can also serve as an initial form of validation for exercises set up manually.



## Use Case 3: After Action Review of Alternative Approaches and Outcomes

---

- Note: we envision this use case as an expanded version of use case 2. Whereas use case 2 deals with exercise preparation, this use case deals with exercise review and lessons learned.
- Scenario: Observer-Controller wishes to be able to explore alternative outcomes to training scenario with trainees by showing simulated results of alternative courses of action
  - During training event, observer-controllers capture data about actions taken by trainees in terms of key interactions with role players.
    - Actions taken by trainees at key decision points during the scenario
  - Following a training event, run simulations on a Cognitive Seldon-modeled version of the scenario to show how outcomes change (in terms of the social response of the modeled role players) depending on what actions are made at key decision points.
  - Trainers & trainees can use tool to show how changing their approach at various key interactions could have potentially led to a different social outcomes



# Conclusions

---

- Cognitive Seldon and RIPPLE Systems of Systems Interoperability Findings:
  - Discovery Task
    - Understanding of each other's vocabulary/subject matter is essential
    - Exchange of detailed technical content and walkthroughs are necessary
    - Assessment of current functional and technical limitations of each system was helpful for assessing interoperability
    - Sustained effort was required in order to keep the team on track
  - Road Mapping Task
    - Envisioning a combined System of Systems that is bigger than the sum of its parts was challenging, but rewarding
    - Educating the client had an opportunistic effect – resulting in development of newer, more innovative use cases that extended the current utility of the individual systems
    - Operational integration would clearly pose more of a challenge than System/Technical Integration
      - Need to focus on doing this holistically, considering the socio-technical nature of a System of Systems