

# TBASSCO Capabilities and Collaboration Plan (Year 1 and Preliminary Year 2)

Bob Neches, Ke-Thia Yao, In-Young Ko  
Information Sciences Institute  
University of Southern California

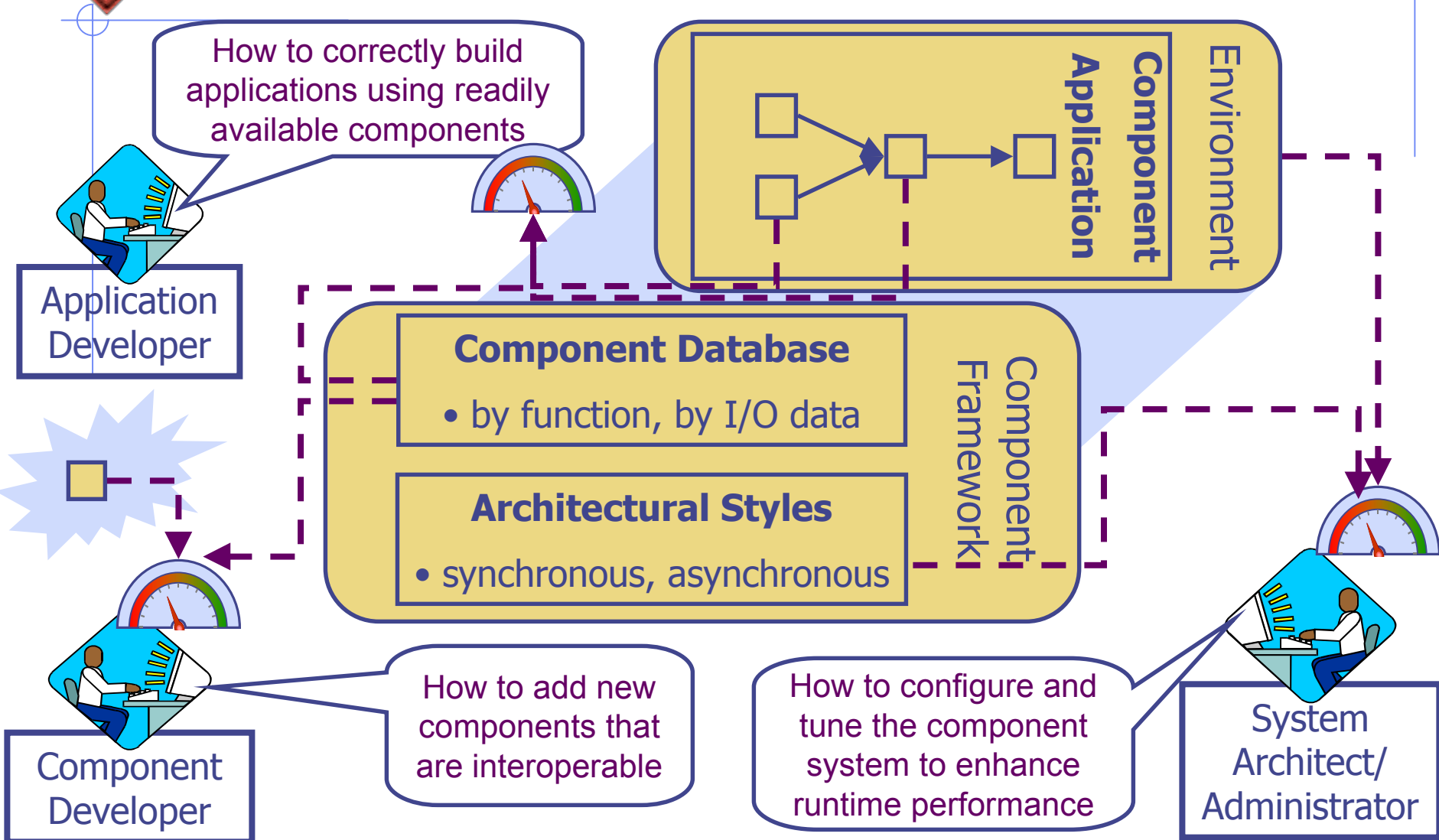


# TBASSCO Coordination Plan Summary and Roadmap

- ◆ Products/Capabilities Produced for Others
  - FY2001: Data flow scripting, component metadata schemas, scripting gauges, component insertion gauges (*see Foils 5-6, 8-11, 15*)
  - FY2002: Data flow to ACME translators, architectural selection gauges (*see Foils 7-9, 11*)
- ◆ Demos
  - FY2001: Four design and runtime scenarios centered around USC ISI's GeoWorlds System at PACOM and JFCOM (*see Foils 14-17*)
  - FY2002: Further scenarios including runtime configuration (e.g. for installation in classified environment)
- ◆ Capabilities needed from others (*see Foils 9, 11*)
  - Semantic distance metric (Georgia); runtime performance gauges (BBN, Columbia/WPI, OBJS), runtime quality gauges (Columbia/WPI)
- ◆ Requirements for other tools to use products (*see Foil 10*)
  - Components described semantically using metadata, runtime performance and quality gauges expressed in group accepted XML standard
- ◆ Evaluation
  - Five criteria to be evaluated collaboratively by the organizations at PACOM and JFCOM using GeoWorlds (*see Foil 18*)



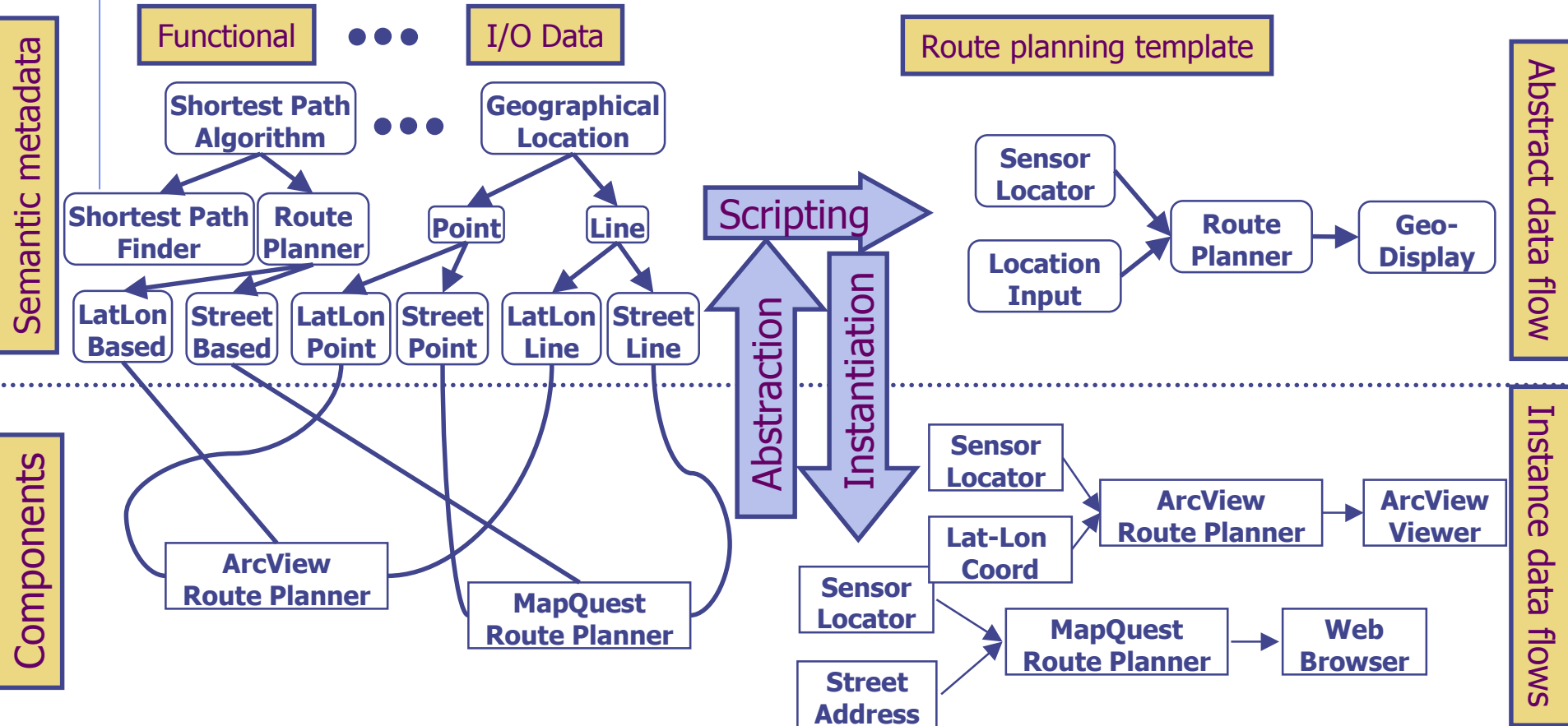
# TBASSCO: Component-based Software Development





# TBASSCO Technical Approach

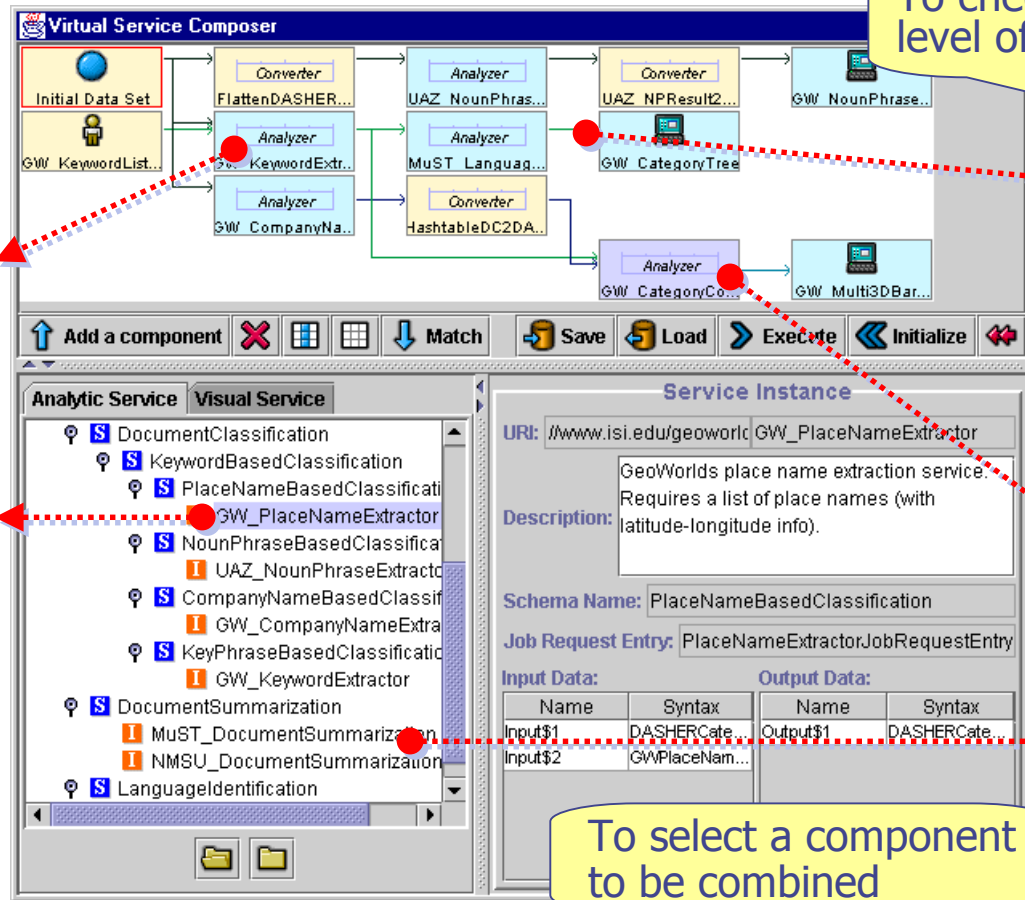
Template-based, multi-view abstraction ADL extensions that model and reason about the *semantics* of component application systems.





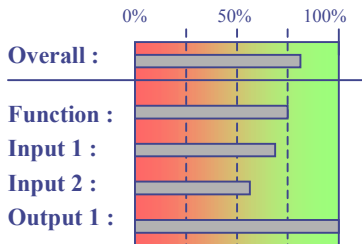
# Design-time Semantic Gauges for Application Scripting

Helps developers of new/improved systems combine component services appropriately



To check the interoperability level of a connection

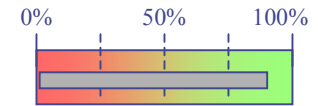
## Compatibility Gauge



Source: Keyword Extractor  
Repl. : Place Name Extractor

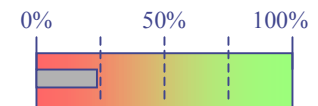
To select an alternative component

## Interoperability Gauge



From :  
Language Identifier  
To :  
Category Tree Viewer

## Interoperability Gauge



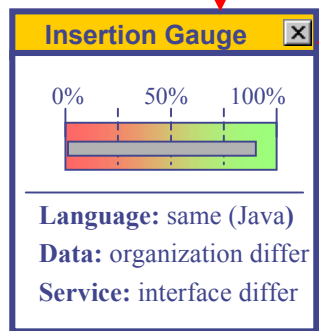
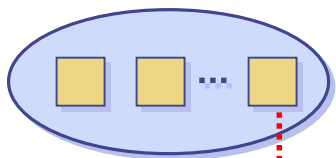
From :  
Category Comparison  
To :  
Document Summary

To select a component to be combined

# Design-time Semantic Gauges for Component Service Insertion

Helps developers of new component services evaluate whether to improve the interoperability of their offerings

## New Components



- Location to insert
- Semantic differences

**InfoManager Resource Editor**

**File**

Analytic Visual User Input Data Converters

- Analytic Service
  - KeywordExtraction
    - NounPhraseExtraction
      - NounPhraseBasedClassification
        - UAZ\_NounPhraseExtractor
      - CompanyNameExtraction
        - CompanyNameBasedClassification
          - GW\_CompanyNameExtractor
      - DocumentClassification
        - KeywordBasedClassification
          - PlaceNameBasedClassification

**Service Schema**

URI: <http://www.isi.edu/geoworlds> anyNameBasedClassification

**Description:** Services of classifying documents based on company names extracted from the document collection

**Superclass:** KeywordBasedClassification, CompanyNameExtraction

**Input Data:**

Name	Content	Structure
Input\$1	Docume...	

**Output Data:**

Name	Content	Structure
Output\$1	Compan...	FlatCate...
Output\$2	Compan...	ListOfD...

**Content Classes** Organization Structure Classes

- Data Content
  - DocumentCollection
    - CategorizedDocumentCollection
      - KeywordBasedCategories
        - NounPhraseBasedCategories
        - CompanyNameBasedCategories
        - PlaceNameBasedCategories
      - DASHERCategory
      - YahooDocumentCollection
      - SearchEngineResultClassifiedBySc
      - DocumentClusters
        - DASHERCategory

**Data Schema**

URI: <http://www.isi.edu/geoworlds> ipanyNameBasedCategories

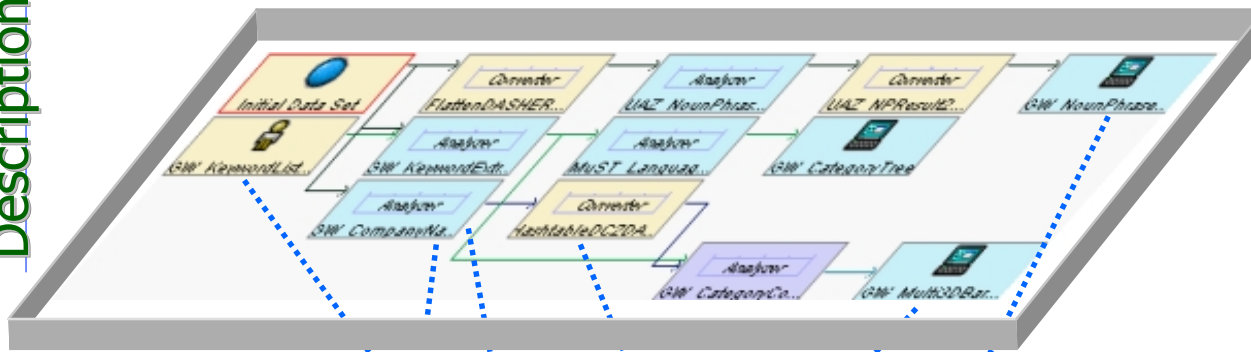
**Description:** Document collection categorized by company names

**Superclass:** KeywordBasedCategories, CompanyNameList

# Design-time Semantic Gauges for Architecture Selection

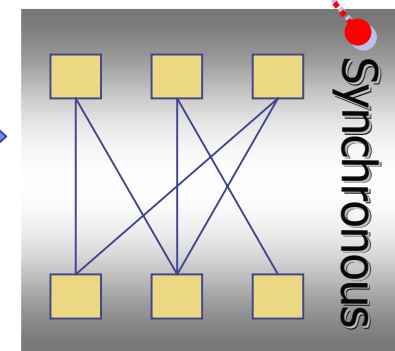
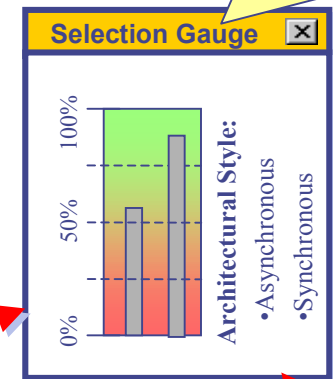
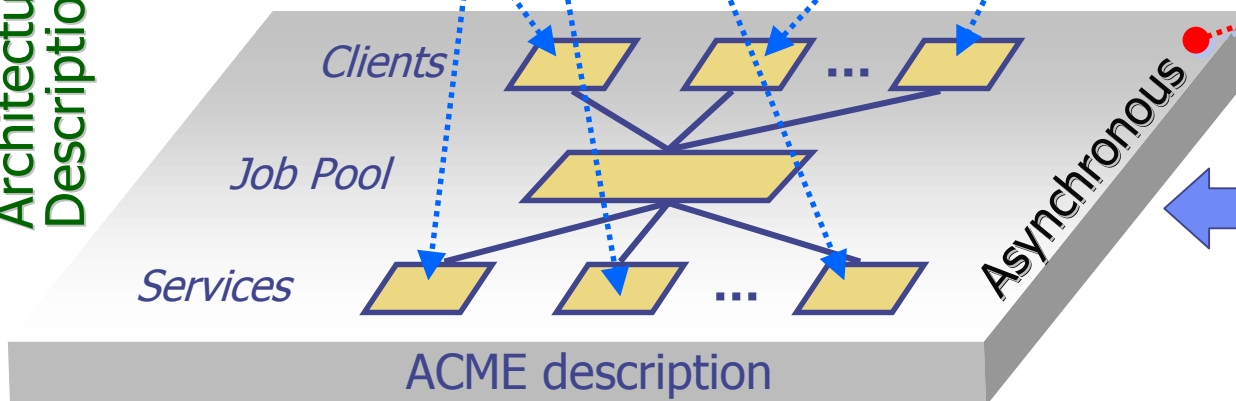
Helps system administrators best configure applications for optimal performance in their local environment

Data Flow  
Description



To select an architecture based on performance gauge feedback

Architectural  
Description





# TBASSCO Capabilities and Impact

FY2001

- ◆ Data flow scripting tools:  
Enable developers to specify how components interoperate
  - Aid combining alternative components and support gauge development
- ◆ **Scripting:** Functional and data gauges
  - Guide in the composition of semantically correct customized applications
  - Help to avoid garbage-in-garbage-out errors
- ◆ **Insertion:** Component insertion gauges
  - Indicate how new components are incompatible
  - Reduce developer time and effort required to add new components

FY2002

- ◆ Translators to transform data flow descriptions to ACME
  - Interface TBASSCO's contributions to other DASADA projects.
- ◆ **Selection:** Architectural level selection and deployment gauges
  - Help developers improve performance by adapting proposed system architectures to computing environment and loads.



# Producer/Consumer Relationships

TBASSCO produces

- ◆ Semantic service and data flow description capability
  - BBN, Veridian and Object Services use them to describe semantic interoperability of their services
- ◆ Service event protocol specification based on semantic service description
  - Columbia/WPI verifies services are *conforming* during runtime

TBASSCO consumes

- ◆ Semantic distance metric to measure interoperability
  - Georgia provides metrics, i.e., clustering and factor analysis
- ◆ Runtime performance to tune architecture
  - Columbia/WPI, and BBN provide performance gauges
- ◆ Runtime service quality to select alternative services
  - Columbia/WPI provides quality gauges, i.e., size of search result



# Requirements for Use

- ◆ Semantic level metadata must be supplied for components analyzed by scripting gauges
- ◆ Group agreement needed on runtime event definitions in XML Schema
- ◆ Group agreement needed on performance metric representations

# Functionality to be Illustrated: Gauges and Probes in the Software Lifecycle



## Design Time Aids

### ISI

- Gauges to select interoperable components
- Gauges to determine difficulty of adding new components
- Gauges to adapt architecture to computing environment

Georgia State

- Semantic distance metrics

## Requirements / Capability Descriptions

### ISI

- Semantic function and data descriptions

BBN

- Architecture requirement documents

## Existing Software (From Library of Available Components)

### • ISI – GeoWorlds

- Object Services – SDC
- BBN – Abstract search engine
- Veridian – GIS map layers

ACME ADL  
XML/FleXML

## Prospective Software (Same Library: Alternative Extensions, Compositions)

## Run Time Aids

### BBN

- Network (bandwidth, latency) gauges
- Uptime gauges

### Columbia / WPI

- Protocol gauges (partial matching on event posets)
- Run-time gauge plugin and modify

### Object Services

- Application profiling gauges and topology gauges on configuration, component usage
  - Component binding
  - Dead libraries
  - Versioning
  - Activity

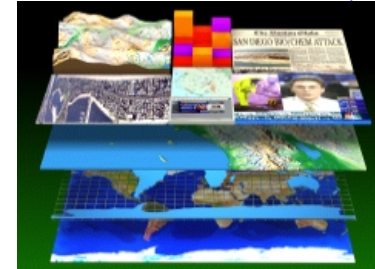


# Test-bed Application: GeoWorlds

- ◆ Large component-based system in use at PACOM & JFCOM
  - **PACOM will serve as outside evaluator**
  - **JFCOM Experimental Battle Lab is potential second evaluator**
- ◆ GeoWorlds architected from beginning as framework for adding components
- ◆ Geographic Information Systems plus Web processing
- ◆ Ops and intelligence uses, e.g.,
  - Mapping terrorist bombings
  - Locating recurring natural disasters
  - Investigating drug trafficking and piracy in various locales



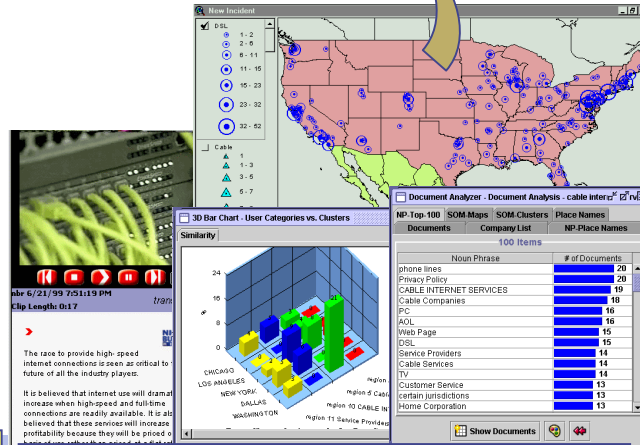
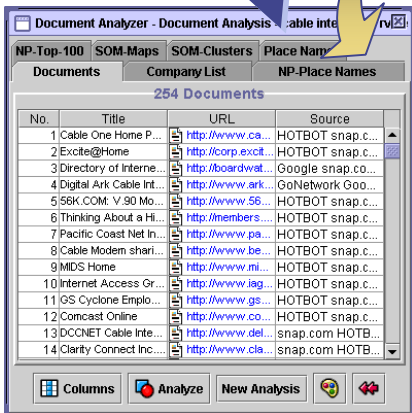
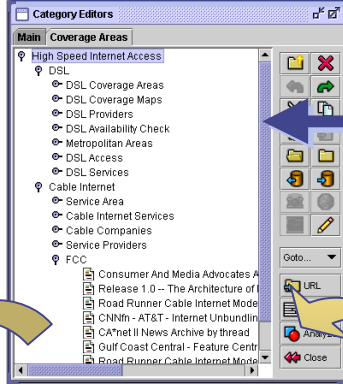
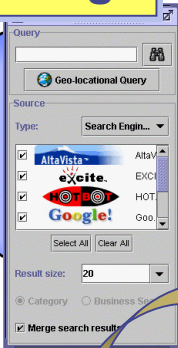
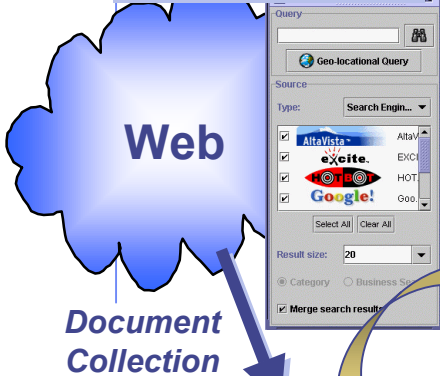
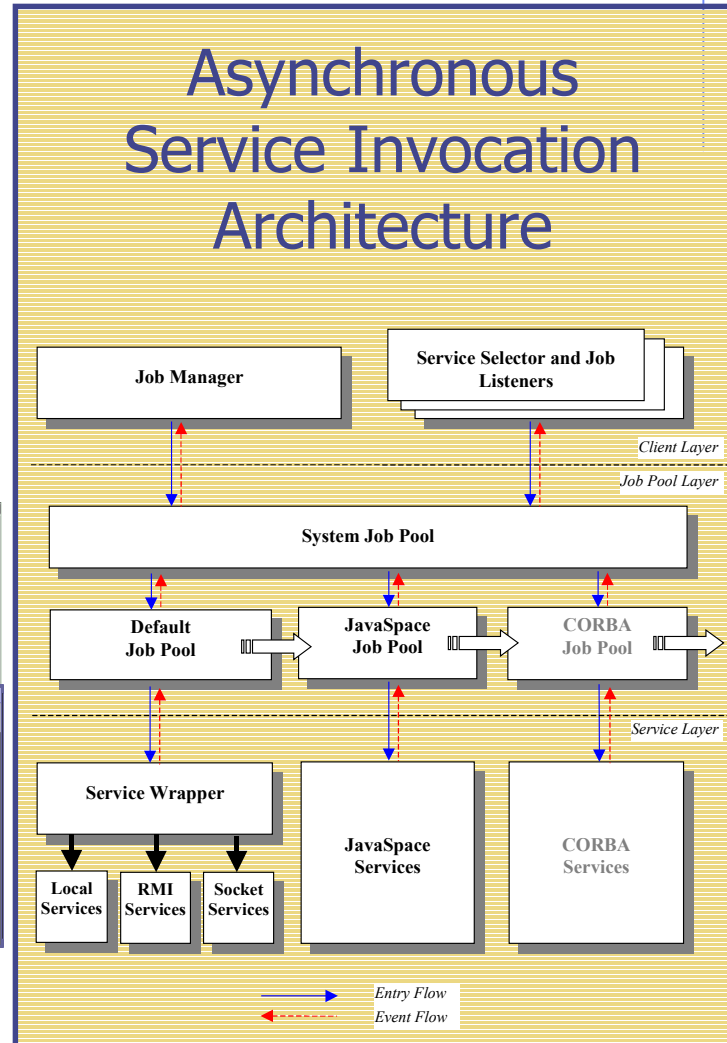
# GeoWorlds: Test-bed Application



## Information Gathering

## Information Organization

## Asynchronous Service Invocation Architecture



## Document Analysis

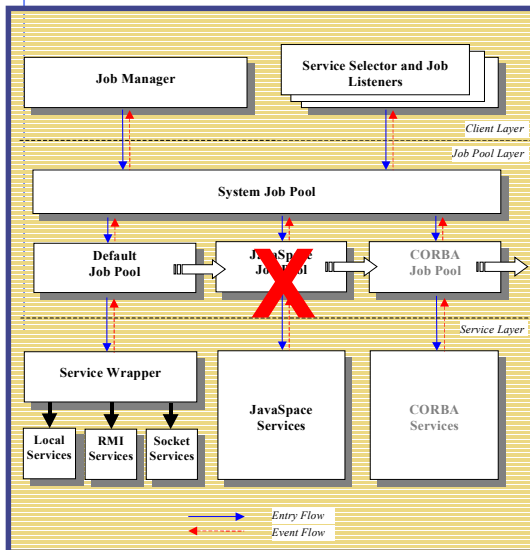
## Information Visualization





# Lifecycle Scenarios

## Installation Time



## Probes & Gauges

### Demo Scenario 1:

#### Installation not fully operational

- ◆ Assume
  - System deployed with specific probes/gauges
- ◆ Action
  - System administrator runs self-diagnostic tests
  - Probes extract information
  - Special GeoWorlds analyzer (provided by GeoWorlds) detects anomalies

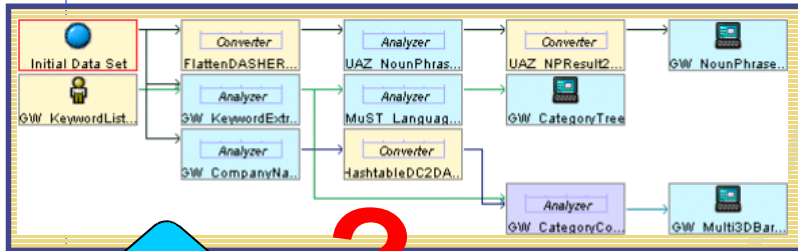
GeoWorlds Application  
Testbed from ISI

User Observes	Causes	Gauges Do	How
Install script fails	Expected component not found	Config gauge identifies missing component	Compare installed config w/ Acme spec
Installation completes, but GeoWorlds doesn't work	Version mismatch	Config gauge identifies use of different version	Compare installed config w/ a good installation
	Namespace error		
	Method invocation fails		



# Lifecycle Scenarios

## Information Management Scripting Time



Probes & Gauges

### Demo Scenario 2:

#### GeoWorlds script not working

- ◆ Assume
  - User Interface has probe palette to monitor component interactions
  - User Interface gauges to interpret probe information
- ◆ Action
  - User selects to probe various connections
  - Identifies a sequence where dataflow flawed
  - User selects alternate component from User Interface
  - Re-execute script correctly



Intelligence Analyst

GeoWorlds Application Testbed from ISI

User Observes	Causes	Gauges Do	How
GeoWorld script cannot be completed	I/O data semantic mismatch between components	Semantic gauges identify the mismatch & suggest intermediate components to resolve it	Perform reasoning on I/O semantics and find components that make a semantic connection
	Syntactic (interface) non-compliance between components	Syntactic gauges determine the cause of non-compliance and suggest adapters	Access to library of converters and wrappers
	Dataflow violation (e.g., pipe output, page input)	Dataflow gauge detects mismatch & suggests a dataflow adaptor to allow	Access to library of available converters
	GeoWorlds can't find appropriate data source	Semantic gauge subscribes data channels that meet requirement	Compare the data requirement and channel descriptions

ISI Scripting Gauges

ISI Scripting Gauges

ISI Scripting Gauges

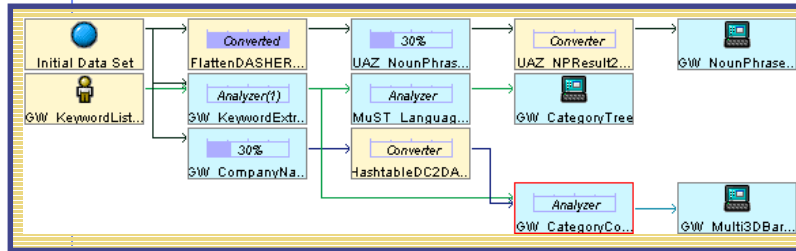
ISI Scripting Gauges



# Lifecycle Scenarios

GeoWorlds Application  
Testbed from ISI

## Script Execution Time



Probes  
&  
Gauges

User Observes	Causes	Gauges Do	How
GeoWorlds hangs during script execution	External data source has moved, disappeared, or become unresponsive	Connectivity gauge identifies the broken connection & finds new URL or drops the source	Monitoring request & response pairs and comparing timing with previous interactions
	External data source has changed its (XML) interface	Change monitoring gauge determines an XML encoding has changed	Comparing the XML Schema used in a sequence of accesses
	External service failure causes GeoWorlds script failure	Connectivity gauge identifies broken connection & suggests alternate service	Monitoring request & response pairs and comparing timing with previous interactions + knowledge of service alternates with similar interfaces
Script returns suspicious results	Spam site introduces flood of dubious responses	QoS gauge identifies aberrant site & helps build a filter to eliminate spurious results	QoS gauge identifies growth in result size
			Dataflow gauge identifies the path through the query that caused the increase

## Demo Scenario 3:

### GeoWorlds hangs or returns unexpected result

#### ◆ Assume

- Probes already in place storing relevant info
- Data source connectivity gauges
- Query difference detection gauges are available

#### ◆ Action

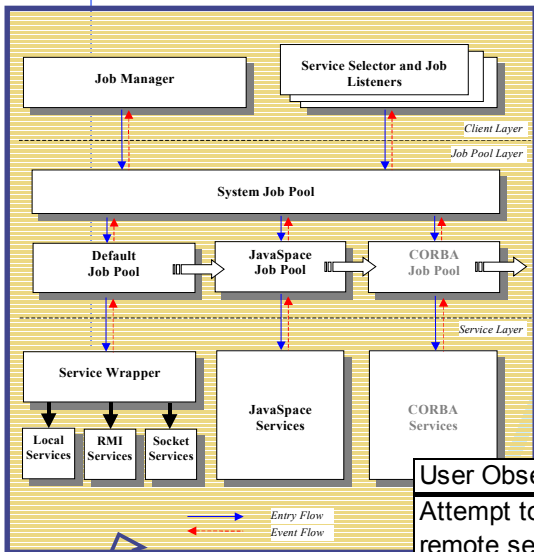
- Gauge alerts user to abnormal situation
- The user (sys admin) tracks down the wayward site and realizes that the URL has changed; reconfigures and continues
- Observe growth in queries returned by GeoWorlds
- Detect aberrant *spam* site that introduced errant info
- Create new filter to eliminate *spam* sites and introduce into GW
- Re-execute query to produce more accurate results



# Lifecycle Scenarios

GeoWorlds Application  
Testbed from ISI

## Reconfiguration Time



### Scenario 4:

#### User service causes failure in GeoWorlds

- ◆ Assume
  - Probes already in place at service layer
  - Gauges exist to verify conformance to expected service behaviors
- ◆ Action
  - User notified through gauge that their service is non-conformant
  - The gauge delivers specific information for how to be compliant
  - User modifies the service as directed by the gauge
  - System functions normally

**Brilliant new service**

User Observes	Causes	Gauges Do	How
Attempt to install new remote service fails	Should not happen - GeoWorlds will accept any service		
Users don't know how difficult it is to add a new service	Mismatch with GeoWorlds data and service specifications	Conformity gauges measure how well do I/O data and functionality conform to GeoWorlds data and function specifications	Check conformity to document models and service APIs, and data and functional ontologies. Check if any wrapper can be applied
Installation Completes, but GeoWorlds doesn't work with new service	GIS Data mismatch with request	Semantic matching of components in reconfiguring the service	Venice allows the user to reconfigure the service remotely using an ACME description for the architecture





# Evaluation Criteria

## Evaluations to be performed collaboratively with GeoWorlds users at PACOM (and, tentatively, JFCOM):

- ◆ How *efficiently* GeoWorlds can be installed in different environments and its services deployed.
- ◆ How *easily* complex information management tasks can be scripted with assured semantic and syntactic interoperability.
- ◆ How *reliably* the scripts can be executed while maintaining desired quality level.
- ◆ How *quickly* the scripts can be evolved based on resource availability and requirement changes.
- ◆ How *efficiently* can new services be added to GeoWorlds while maintaining compatibility.