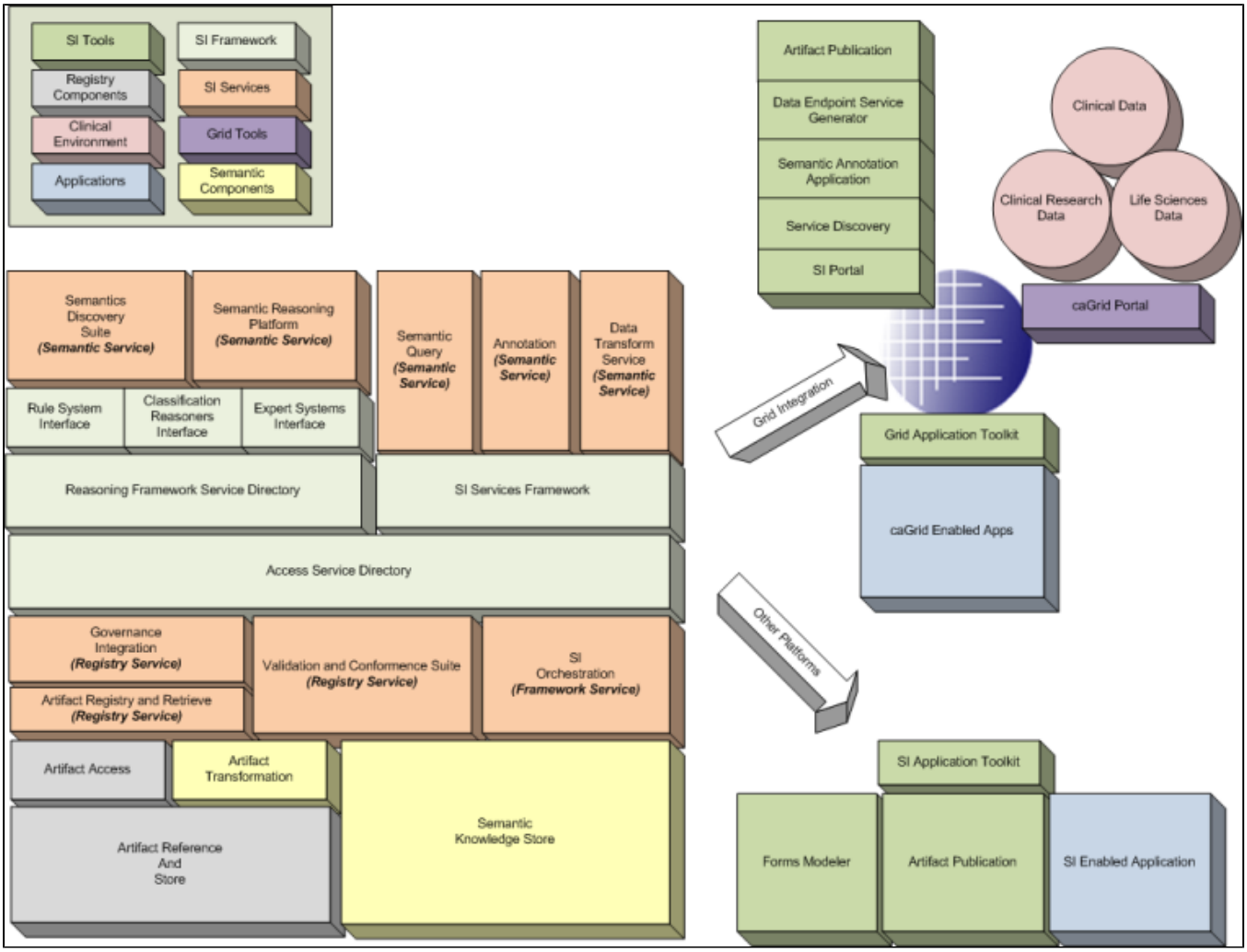
6.1 - Overview of Semantic Infrastructure 2.0 Architecture

Current Working Draft

This section provides an overview of the proposed architecture, which includes a set of core services and tools. Section 6.2 - Overview of Semantic Infrastructure 2.0 Capabilities and Services summarizes the profile of the solution with mapping to appropriate requirements and use cases. Section 6.3 - Tools for Semantic Infrastructure 2.0 provides an end user's view of the tools. Section 6.4 - Tie-in with Terminology and Platform describes integration with the platform and terminology.

Overall View of the Semantic Infrastructure 2.0 Architecture



This diagram has a series of boxes, each labeled with a name and two arrows signifying relationships. These elements are described in the table that follows.

Component Name	Description
Registry Components and SI Services: Registry	Refer to the descriptions for each component in the following rows.
Box Name: Artifact Reference and Store	This registry component is a store or registry that contains references to the various artifacts. Each artifact should have a URL that can be used to physically access the file. Each artifact reference is accompanied by a checksum or some other method to be able to verify the accessed object.
Box Name: Artifact Access	This registry component provides programmatic access to artifacts in the Artifact Reference and Store.
Box Name: Artifact Registry and Retrieve	This registry SI service provides a programmatic interface for interacting with the artifact reference registry.
Box Name: Governance Integration	This registry SI service provides state mechanisms about known artifacts that can be accessed and reviewed through governance activities.

Box Name: Validation and Compliance Suite	This registry SI service integrates with the reasoning system to validate the compliance of specific artifacts (ECCF models).
Semantic Components and SI Services: Semantic	Refer to the descriptions for each component in the following rows.
Box Name: Semantic Knowledge Store	This semantic component provides a physical representation of semantics that have either been derived through artifact analysis, or through manual annotation. This store could be represented by an RDF (Resource Description Framework) triple store.
Box Name: Artifact Transformation	This semantic component provides a function that takes as input some artifact and provides output in alternative representations. This might include a class model in UML being transformed to an OWL ontology.
Box Name: Semantics Discovery Suite	This semantic SI service takes as input artifacts or artifact transformations and extracts as many semantic representations as possible. The details of the semantics will depend on artifact type, representation, and completeness. The results are then stored in the Semantic Knowledge Store.
Box Name: Semantic Reasoning Platform	This semantic SI service provides description logic based functionality that uses inferencing to allow additional semantics (assertions) to be added about an artifact reference in the Semantic Knowledge store, supporting capabilities such as checking ontology consistency, building classification, verifying the satisfiability of concepts, and checking entailment.
Box Name: Semantic Query	This semantic SI service provides querying functionality on the semantics of an artifact reference in the Semantic Knowledge store. Different querying mechanisms may be supported, including but not limited to graph-based, rule-based, tree or directory-based, and object-based.
Box Name: Annotation	This semantic SI service provides functionality that allows additional semantics to be added about an artifact reference in the Semantic Knowledge store, and is used to augment the semantic representations which were automatically discovered.
Box Name: Data Transformation Service	This semantic SI service provides a set of transformation functions which are designed to transform data; this may include transforming data graphs into CSV, result sets into XML, or other reasonable transformations. This function may use semantics stored about artifacts to aid in the transformation function.
SI Framework Components and SI Services: Framework	Refer to the descriptions for each component in the following rows.
Box Name: Access Service Directory	This framework component provides the set of services that are available within a Semantic Infrastructure implementation which are designed to manage artifacts. This will allow for the coordination of stores and services across the grid.
Box Name: Reasoning Framework Service Directory	This framework component provides the set of services that are available within a Semantic Infrastructure implementation that provide reasoning functionality to analyze artifacts and instance representations of associated data.
Box Name: Rule System Interface	This framework component provides integrations of one or more rule systems to support to the Semantic Infrastructure in expressing business rules and behaviors.
Box Name: Classification Reasoners Interface	This framework component provides integrations for one or more classification tools. These tools are systems that process semantic and dependent information to determine relationships and associations of classes and individuals which may be expressed in an artifact, its annotated information, or instance representations of associated data.
Box Name: Expert System Interface	This framework component provides integration to one or more expert systems. These systems utilize a set of known facts and domain expert definitions to determine additional semantics and functional definitions within the artifact semantic information and instance representations of associated data.
Box Name: SI Services Framework	This framework component provides interface support to semantic and reasoning services.
Box Name: Orchestration	This framework SI service manages the internal flow of operations that can be performed. This includes automating the transformation and semantic discovery and the utilization of various rule systems or classification systems.
Integrations and applications	Refer to the descriptions for each component in the following rows.
Arrow: Grid Integration	The grid integration represents the interaction of Semantic Infrastructure services with the caGrid
Box Name: Grid Application Toolkit	This Semantic Infrastructure tool provides libraries and functions that ease the creation of new caGrid enabled applications. This tool kit will provide a method to integrate caGrid 1.0 applications to ease applications into the caGrid 2.0 environment.
Box Name: caGrid Enabled Applications	caGrid enabled applications include any application written to the caGrid specification.
Box Name: caGrid Portal	This caGrid application is a tool for accessing aspects of of the caGrid in a partner site.
Box Name: Clinical Data	This represents clinical information that may be exposed to the grid. Using the portal, an authorized user may expose data or services onto the grid; this might include outcome markers, treatment plans or other relevant information
Box Name: Clinical Research Data	This represents clinical research data that might be exposed to the grid. Using the portal, an authorized user may expose data or services onto the grid; this might include trial cohort qualifications, raw data, or publishable results.
Box Name: Life Sciences Data	This represents life sciences data that might be exposed to the grid. Using the portal, an authorized user may expose data or services on the grid; this might included gene array studies, algorithms, methodologies and data sets.
Box Name: SI Portal	This application provides a user interface for implementations of the Semantic Infrastructure framework components. The user would use this tool to access the functionality of the Semantic Infrastructure components exposed on the grid. Probably a part of the caGrid Portal
Box Name: Service Discovery	This tool and portal component provide a user with the ability to enter key words and tags or semantic queries to help determine the locations of artifacts and communication endpoints.
Box Name: Semantic Annotation Application	This tool and portal component provides a user with the ability to annotate artifacts and communication endpoints to help the user perform queries.

Box Name: Data Endpoint Service Generator	This tool allows a user to quickly create a data endpoint and make it available on the caGrid, merging the data source with a SPARQL Endpoint and structuring for access.
Box Name: Artifact Publication	This tool allows a user to take an artifact and provide a reference to the registry components of the Semantic Infrastructure framework, and provide basic annotations.
Arrow: Other Platforms Integration	This integration represents the interaction of Semantic Infrastructure services with applications and platforms that might need to utilize function of the Semantic Infrastructure.
Box Name: SI Application Toolkit	This Semantic Infrastructure tool provides libraries and functions that ease the creation of new Semantic Infrastructure Framework enabled applications.
Box Name: Forms (and Object) Modeler	This Semantic Infrastructure tool is used to create forms models, message models and other core object models from defined structures. This tool works with information in the Semantic Infrastructure to access meta-models and model definitions to construct representations of objects which can be used for data collection and information exchange.
Box Name: Artifact Publication	This Semantic Infrastructure is the non-portal version of the artifact publication found in the Semantic Infrastructure portal. This component is different, because it will provide greater access to various components, enhanced governance support and manipulation of Knowledge Store objects requiring enhanced behaviors.
Box Name: SI Enabled Applications	This represents any number of applications that might need access to Semantic Infrastructure functionality and would utilize the Semantic Infrastructure application toolkit. This may include NCI applications such as caTissue.