

# SMILA Release Review 0.7

Project leads: I. Novakovic, S. Voigt

Planned Review Date: March, 17<sup>th</sup> 2010

Communication Channel: [[smila-dev@eclipse.org](mailto:smila-dev@eclipse.org)]

# SMILA - SeMantic Information Logistics Architecture

## “Unleashing the potential of unstructured data sources”

- The amount and diversity of information is growing exponentially, mainly in the area of unstructured data, like emails, text files, blogs, images etc. Poor data accessibility, user rights integration and the lack of semantic meta data are constraining factors for building next generation enterprise search and other document centric applications. Missing standards result in proprietary solutions with huge short and long term cost.
- **SMILA** is an extensible framework for building search solutions to access unstructured information in the enterprise. Besides providing essential infrastructure components and services, **SMILA** also delivers ready-to-use add-on components, like connectors to most relevant data sources. Using the framework as their basis will enable developers to concentrate on the creation of higher value solutions, like semantic driven applications etc.

# Features

- Enterprise Search Framework
- Crawling/Agent Framework
- Processing Framework
  - BPEL
- Search Framework
  - BPEL
- General
  - Monitoring / Management (JMX)
  - Distributable (embedded JMS)

# Non-Code Aspects

- User Documentation
  - Every Feature/Bundle is well documented (<http://wiki.eclipse.org/SMILA>)
- JavaDoc
- Many JUnit-Test cases (557 on the last count) to validate that APIs are working
- Examples:
  - [http://wiki.eclipse.org/SMILA/Documentation for 5 Minutes to Success](http://wiki.eclipse.org/SMILA/Documentation%20for%205%20Minutes%20to%20Success)
- How-tos
  - Several How-tos for Setup / Build / Creation of bundles / Integration of Components (for Developers/Integrators)

# Non-Code Aspects: Talks / Demos

Type	Event	Location	Date	URL
<b>2008</b>				
Talk	Entwicklertag 2008	Karlsruhe, TPK	07.05.2008	<a href="http://nontypo.andrena.de/Entwicklertag/2008/Vortraege/EILF.html">http://nontypo.andrena.de/Entwicklertag/2008/Vortraege/EILF.html</a>
Talk & Demo	2. Partnertreffen ORDO	KL, empolis	26.06.2008	
Talk & Demo	TEXO Development Day	Karlsruhe, SAP-Research	10.07.2008	
Talk	TEXO Plenum	GT, empolis	09.10.2008	
Poster	ISWC 2008	Karlsruhe, Kongresszentrum	28.10.2008	<a href="http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-401/iswc2008pd_submission_44.pdf">http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-401/iswc2008pd_submission_44.pdf</a>
Training	Empolis Evolution	GT, empolis	12.11.2008	
Talk	ESE 2008	Ludwigsburg, Forum am Schlosspark	20.11.2008	<a href="http://www.eclipsecon.org/summiteurope2008/sessions?id=41">http://www.eclipsecon.org/summiteurope2008/sessions?id=41</a>
Webinar	-	Web	17.12.2008	<a href="http://live.eclipse.org/node/617">http://live.eclipse.org/node/617</a>
<b>2009</b>				
Talk & Demo	EclipseCon 2009	Santa Clara, CA, USA	26.03.2009	<a href="http://www.eclipsecon.org/2009/sessions?id=398">http://www.eclipsecon.org/2009/sessions?id=398</a>
Tutorial	BIS 2009	Poznan, Polen	28.04.2009	<a href="http://bis.kie.ae.poznan.pl/12th_bis/tutorial.php?tut=smile2009">http://bis.kie.ae.poznan.pl/12th_bis/tutorial.php?tut=smile2009</a>
Talk	OpEN.SC 2009	B, Charité Campus	08.05.2009	<a href="http://opensc.charite.de/index.php">http://opensc.charite.de/index.php</a>
Talk	IKS Workshop	Salzburg, Österreich	28.05.2009	<a href="http://www.iks-project.eu/">http://www.iks-project.eu/</a>
Tutorial	International THESEUS Symposium 2009	B, BMWi	30.06.2009	<a href="http://www.theseus.joint-research.org/index.php?id=43#si1">http://www.theseus.joint-research.org/index.php?id=43#si1</a>
Tutorial	LWA 2009	Darmstadt	23.09.2009	<a href="http://lwa09.informatik.tu-darmstadt.de/bin/view/LWA/Tutorial">http://lwa09.informatik.tu-darmstadt.de/bin/view/LWA/Tutorial</a>
Workshop	Workshop with InfoChem	Bielefeld	26.10.2009	
Talk	ESE 2009	Ludwigsburg, Forum am Schlosspark	28.10.2008	<a href="http://www.eclipsecon.org/summiteurope2009/sessions?id=876">http://www.eclipsecon.org/summiteurope2009/sessions?id=876</a>
Talk	TEXO Meeting	Dresden	06.11.2009	
Talk & Demo	Eclipse DemoCamp	Berlin	23.11.2009	<a href="http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Berlin">http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Berlin</a>
Talk & Demo	Eclipse DemoCamp	Braunschweig	25.11.2009	<a href="http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Braunschweig/Hanover">http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Braunschweig/Hanover</a>
Talk & Demo	Eclipse DemoCamp	Kaiserslautern	26.11.2009	<a href="http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Kaiserslautern">http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Kaiserslautern</a>
Talk & Demo	Eclipse DemoCamp	Karlsruhe	03.12.2009	<a href="http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Karlsruhe">http://wiki.eclipse.org/Eclipse_DemoCamps_November_2009/Karlsruhe</a>
<b>2010</b>				
Workshop	Workshop Fraunhofer IAIS	St. Augustin	21.01.2010	

# APIs

- API is stable since 0.5M2
  - Processing Service API (for Integrators)
  - Data Connector API (for Integrators)
  - Monitoring and Performance Counter API (is based on JMX)
- All APIs are provisional and are expected to evolve based on community feedback
- Extensibility and easy integration of new components is the key aspects of SMILA

# Architectural Issues

- SMILA consists of several different parts
  - Data Connectors / Connectivity
  - Processing
  - Search
  - Persistence

Each of these contains several bundles, but each can be started on different hosts, and they communicate with an embedded JMS Queue (Apache ActiveMQ)
- The use of declarative OSGi Services wherever possible
  - OSGi Services are used as API (interface definition)
  - SMILA contains several implementations for these services
  - Services can be replaced by integrators/customers

# Tool Usability

- Management of SMILA components
  - SMILA contains a small JMX client (remote)
  - Extension of the OSGi console
- JConsole can also be used for fully framework management and monitoring



# Bugzilla

Severity	Status					
	NEW	ASSIGNED	RESOLVED	VERIFIED	CLOSED	Total
	blocker	.	<u>3</u>	.	.	<u>3</u>
	critical	.	<u>1</u>	.	.	<u>1</u>
	major	.	<u>2</u>	<u>1</u>	.	<u>3</u>
	normal	<u>3</u>	<u>81</u>	.	<u>5</u>	<u>89</u>
	minor	<u>1</u>	<u>3</u>	.	<u>1</u>	<u>5</u>
	trivial	<u>2</u>	<u>1</u>	.	.	<u>3</u>
	enhancement	<u>14</u>	<u>4</u>	.	.	<u>44</u>
Total	<u>17</u>	<u>7</u>	<u>117</u>	<u>1</u>	<u>6</u>	<u>148</u>

# Standards

- Attempt to base as much as possible using OSGi standards
  - OSGi declarative services for the internal communication
- Internal distributed communication uses JMS
- BPEL for the definition and execution of workflows
- XML / XML schema for configuration files (including validation)

# Schedule

- **SMILA Version 0.5 M1**

- This milestone was published on March 23rd 2009.
- Basic architecture settled and implemented
- Simple search application available
- Diverse connectors (File System, Web, DB)
- Incremental update
- Conception and implementation of search APIs
- Implementation of the security concept
- Ontology Service (introduction of semantic layer)

- **SMILA Version 0.5 M2**

- This milestone was published on May 29th 2009.
- Simple compound management
- More connectors (RSS & Atom)

- **SMILA Version 0.5 M3**

- This milestone was published on July 31st 2009.
- Bulk operations (increasing the data throughput)
- More connectors (XML, Job file)
- Alternative binary storage implementation (based on JPA)

# Roadmap

- **SMILA Version 0.7 - Release 0.7**

- This release is planned for March, 17th 2010.
- Web Service API support
- Solr-Integration
- Completed IP process

- **SMILA Version 0.8**

- This milestone is tentative planned for June 25th 2010.
- Cluster readiness (remote communication between components / REST API)
- General configuration management support

- **SMILA Version 0.9**

- This milestone is tentative planned for September, 30th 2010.
- Aperture integration
- BPEL editor extensions
- Concept for debugging of BPEL processes
- Completed IP process