



STEM 1.0.0 Release Review

Matthew Davis, Stefan Edlund, Daniel Ford,
and
James Kaufman

2/10/2010



STEM Links

Web site

<http://www.eclipse.org/stem>

Downloads

<http://www.eclipse.org/stem/downloads.php>

Newsgroup *(Preferred Communications Channel)*

http://www.eclipse.org/forums/index.php?t=thread&frm_id=72

Project Plan

<http://www.eclipse.org/projects/project-plan.php?projectid=technology.stem>

IP Log

http://www.eclipse.org/projects/ip_log.php?projectid=technology.stem



Introduction

- **The Spatio-Temporal Epidemiological Modeler (STEM)** is a tool designed to help scientists and public health officials to study emerging infectious diseases. STEM uses mathematical models of diseases to simulate the development or evolution of a disease in space and time (e.g., H1N1 or salmonella). These models could aid in understanding and potentially containing the spread of such diseases. STEM also comes pre-configured with a vast amount of reference or denominator data for the entire world.

Furthermore, STEM provides tools to help users compare and validate models. As an Eclipse project, the ultimate goal of STEM is to support and encourage a community of scientists that not only use STEM as a tool but also contribute back to it. STEM is designed so that models and scenarios can be easily shared, extended, and built upon.



Features

- First Release, all features are “new”
- Core Framework based on EMF
 - Defines the world as a graph
 - Graph labels and decorators describe disease models, populations, transportation, etc.
- Denominator Data for 244 ISO defined countries and dependent regions
 - Human Population
 - Region areas
 - Relationships between places
 - Adjacency
 - Roads in North America
 - Air travel for world calibrated against FAA data



Features

- **Basic Disease Models**
 - Stochastic and Deterministic SI, SIR, SEIR compartment models
- **Human Population Models**
 - Allows user specified Birth/Death Rates
- **Differential Equation Solvers**
 - Finite Difference (fast, good for demos)
 - Runge-Kutta (accurate numerical integration)
- **Multiple Runtime Interfaces**
 - Graphical Interface built on Eclipse Workbench
 - Three workspace perspectives: Simulation, Designer, Analysis
 - BIRT-powered views for monitoring and charting active simulations
 - Headless option for running automated batch mode
 - File system logging of results
- **J-Unit Test Suite for each component**



Non-Code Aspects

- Documentation
 - Continuously updated to reflect new features
 - Recently moved to Wiki
- Localization
 - Attempts to use NLS as often as possible
 - Input data as UTF-8
- Examples and Tutorials
 - Example disease models included by default. Example geographic models and simulations posted on the Web site
 - Tutorial video posted on YouTube. Translated in English, Spanish, and Hebrew
- Articles and Publications
 - Numerous peer-reviewed academic publications
 - Numerous press mentions, most recently for work with Mexico



APIs

- Significant time spent building and refining the underlying data models and APIs
 - Designed around epidemiology use case with extensibility for general use cases
 - Supports any type of compartment model
 - Easily extendable (SIR builds on SI, SEIR builds on SIR, etc).
 - All APIs are fully implemented
- Most *core* APIs are “Platform”
- Data File Format and Generation APIs are “Platform”
- Disease Modeling and Simulation APIs are “Provisional”
 - New features, including multi-disease modeling, will require modification. Will attempt backwards compatibility.
- All major APIs include an extensive test suite and implemented in the STEM UI



Architectural Issues

- Like APIs, considerable time spent in architecture phase
- Overall architecture is considered *stable*
- Ready to separate STEM into Eclipse *Features* if/when we decide to support other modeling paradigms in the future (e.g., agent based models).



Tool Usability

- STEM is designed to overcome a major problem facing public health : *Epidemiologists generally aren't computer scientists*
- STEM provides tools to compose, simulate, and analyze disease models using a drag-and-drop interface
- Running a simulation can be as easy as dragging items and clicking “run”
 - Users don't have to understand Java
 - Users don't have to collect and verify denominator data
 - Users don't necessarily need to write their own disease models
- ...but, if needed, STEM provides APIs to extend the platform
- Hide the *technical* complexity of epidemiology modeling, while preserving the flexibility of “doing it yourself”



End-of-Life

- First official release
- No feature/API deprecation or removal between milestones and release
- **Note:** Models and simulations from versions before move from OHF unsupported (namespace references)
 - Documented on Downloads Page, will remove old versions after release



Bugzilla status

- 57 open bugs (none labeled as release blockers)
- All open bugs migrated from OHF/STEM to STEM Bugzilla
- All contributions added to IPLOG
 - All contributors IBM Employees.



Standards

- GML
- ISO-3166
- ISO-15836 (Dublin Core)
- FIPS



UI Usability

- STEM uses Eclipse RCP to provide a drag and drop UI for disease model composition, simulation, and analysis
- Design Perspective
 - *Simulations* composed by dragging and dropping user-created or built-in (generated) models, diseases, infectors, inoculators, etc
 - Composition UI allows easy extensibility and layering of models
 - Built-in data define geography, transportation systems, and population for 244 countries and dependent areas
 - Editors allow customizable disease params, geographies, etc
- Simulation Perspective
 - Provides runtime simulation control, including start, stop, pause, and steps
 - Geographic map for spatial view current disease spread (can be disabled)
 - BIRT-powered spatiotemporal charting to monitor disease params for user selected geographic regions
- Analysis Perspective
 - Sophisticated tooling to perform comparative analysis of simulations



Schedule

- Milestone M1 (9/15/09):
 - ◆ Implementation of the Automated Experiment.
 - ◆ Batch simulation functionality that automatically walks the parameters of a disease model to minimize the error compared to a reference data set
- Milestone M2 (10/16/09):
 - ◆ General bug fixes and improvements to key denominator data.
 - ◆ Correct discontinuity in transportation across physically adjacent regions when region areas.
 - ◆ Validation and correction (as needed) for areas of administrative regions (Globally)
 - ◆ Validation and correction (as needed) for distribution of populations across subdivisions
- Release 1.0.0 (2/17/10):
 - ◆ Feature complete for 1.0
 - ◆ Including Example Scenarios at Continent and Super-Continent levels.
- ◆ Release 1.0.1 (TBD)
 - ◆ Bug fixes for 1.0, more examples
- ◆ Release 1.1.0 (TBD)
 - ◆ Complete NLS support, non-English translations
 - ◆ Refined Nelder-Mean automated experiments and documentation
 - ◆ Stochastic Poisson Disease Model
 - ◆ Global Disease Model Projections



Communities

- Contributors and Committers
 - Active developer community with committers in the US and Israel and contributors from Germany
 - Weekly conference calls to discuss status
 - PR efforts through IBM, YouTube demonstration videos, workshops
- Users
 - STEM is being adopted by public health institutions around the world and numerous other users, including:
 - Johns Hopkins School of Public Health
 - Univ of Vermont School of Medicine
 - System donated to U.S. CDC
 - Israeli CDC
 - Mexico (GDF) Ministry of Health
- Adopters
 - Extensions and data contributions being contributed by students to support graduate-level public health research



IP Log & Issues

- http://www.eclipse.org/projects/ip_log.php?projectid=technology.stem
- No known outstanding issues
- IP Log attached



Maturity and Feedback

- Maturity: STEM is being used by several institutions including U. Vermont School of Medicine, Johns Hopkins, The Israeli CDC, and now the Mexico Ministry of Health (IBM Recently Donated Servers to them).
- User feedback:
 - ◆ The Main Feedback we get is that public health users want STEM implemented as a Web Service. We will make this a top 2010 goal for STEM 2.0
 - ◆ The other feedback is the desire for more complete documentation. We gave this a top priority for the V1.0 release plan. See for example:
 - STEM home page: <http://www.eclipse.org/stem/>
 - Eclipse wiki: <http://wiki.eclipse.org/index.php/STEM>
 - Wikipedia article: http://en.wikipedia.org/wiki/Spatiotemporal_Epidemiological_Modeler
 - 1 hour YouTube Video Tutorial
 - ✦ In English <http://www.youtube.com/watch?v=LfiibQX4IFE>
 - ✦ In Spanish <http://www.youtube.com/watch?v=3S5DbjCHsx4>
 - ✦ In Hebrew <http://www.youtube.com/watch?v=AFYXvazKkms>