Integrated Modeling Concepts

A clinic at the Community Surface Dynamics Modeling System (CSDMS) 2015 Annual Meeting
Boulder, Colorado
May 26-28, 2015

Jonathan Goodall
Associate Professor
Department of Civil and Environmental Engineering
University of Virginia

download slides: http://goo.gl/VRgzc9

Clinic Aims

- This clinic is intended for early career researchers interested in gaining an understanding of basic integrated modeling concepts as they relate to modeling earth science systems.
- The class will present
  - key literature in the field,
  - core concepts and terminology, and
  - different integrated modeling systems.
- Past, present, and future trends for designing integrating modeling systems will be discussed.
- Participants will also gain experience applying integrated modeling concepts using CSDMS for simplified integrated modeling examples.

Outline

- **Introduction**
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- Example modeling frameworks
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- Interface Standards
  - OpenMI
  - BMI
- Current topics
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- Tutorials
  - Getting Started with WMT (time permitting)
- Summary

Integrated Modeling Definition

"**Integrated Modeling** includes a set of interdependent science-based components (models, data, and assessment methods) that together form the basis for constructing an appropriate modeling system (EPA, 2008b, 2009)."

Source: Laniak et al., 2013

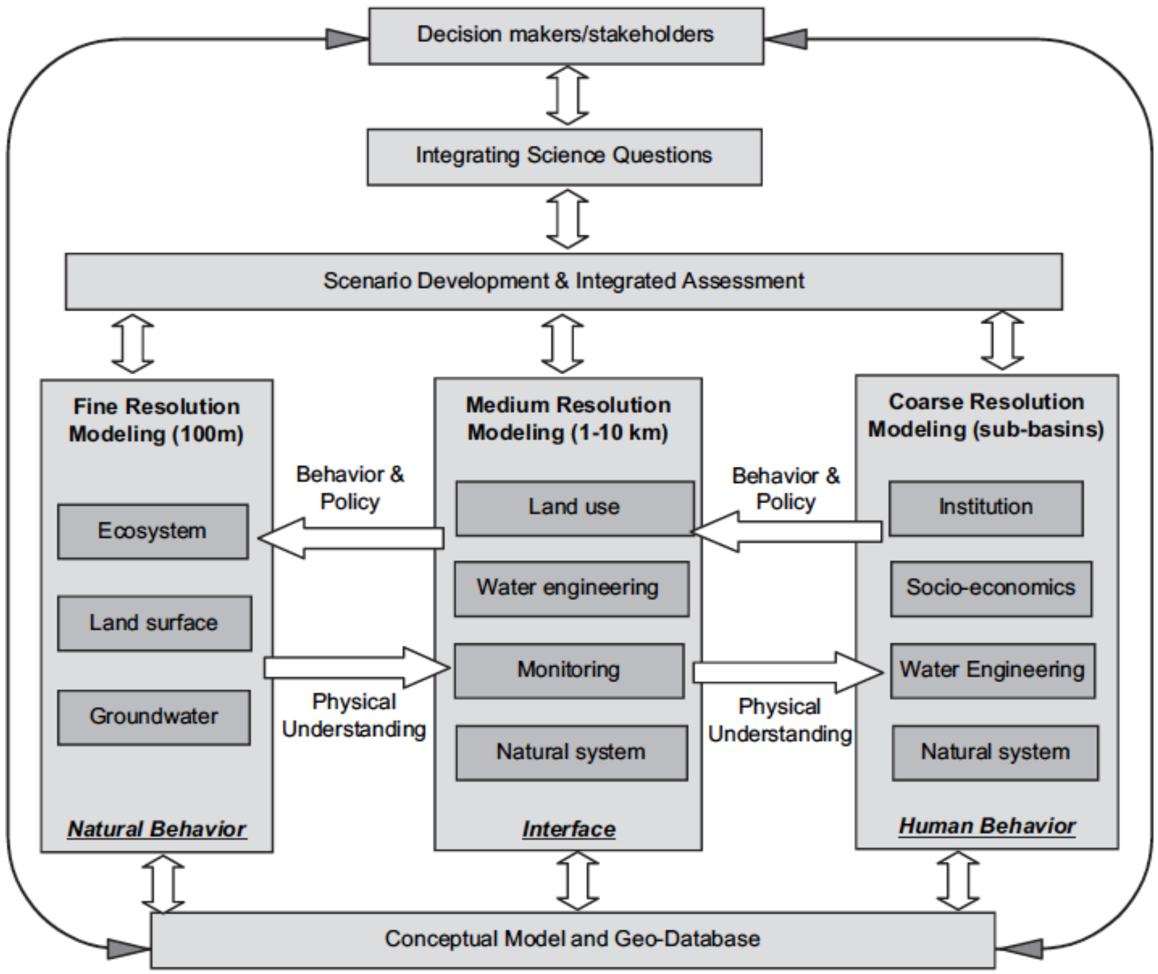
Integrated Modeling Landscape

- **IM Applications:** problem formations and solution approaches that are transparent and holistic; recognizing interdependent relationships
- **IM Science:** transdisciplinary; human-environmental systems; stresses model uncertainties
- **IM Technology:** provide a means to express, integrate, and share the science of IEM; standards and tools to facilitate the discovery, access, and integration of science components.
- A **Community** of IEM stakeholders and associated organizations

The primary focus of this talk will be on the technology aspects of IM

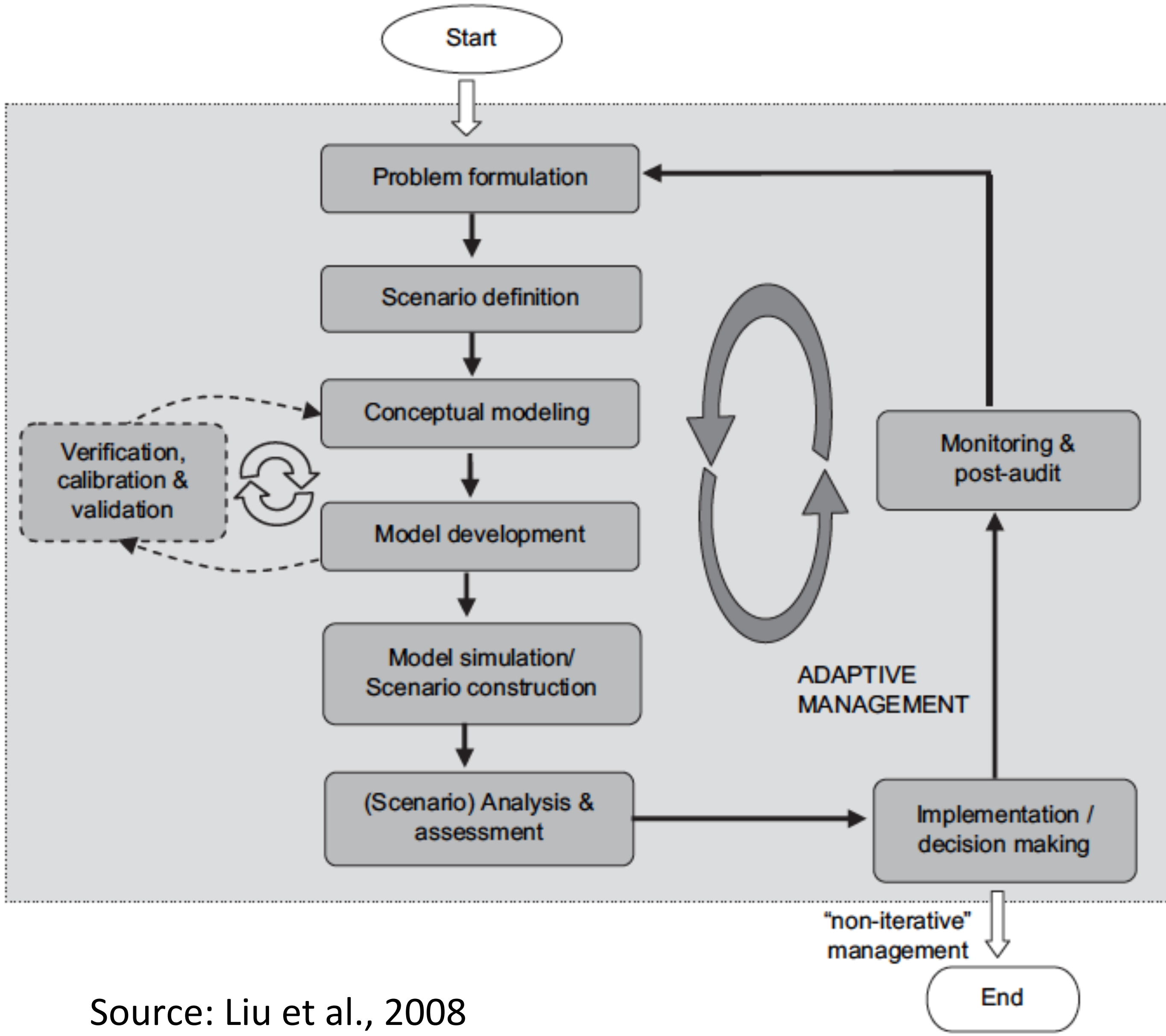
Source: Laniak et al., 2013

Example of Integrated Modeling



Source: Liu et al., 2008

The big picture workflow



Source: Liu et al., 2008

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Environmental Modelling & Software

Environmental Modelling & Software 19 (2004) 219–234

www.elsevier.com/locate/envsoft

An overview of model integration for environmental applications—components, frameworks and semantics

Robert M. Argent *

Department of Civil and Environmental Engineering, The University of Melbourne, Vic. 3010, Australia

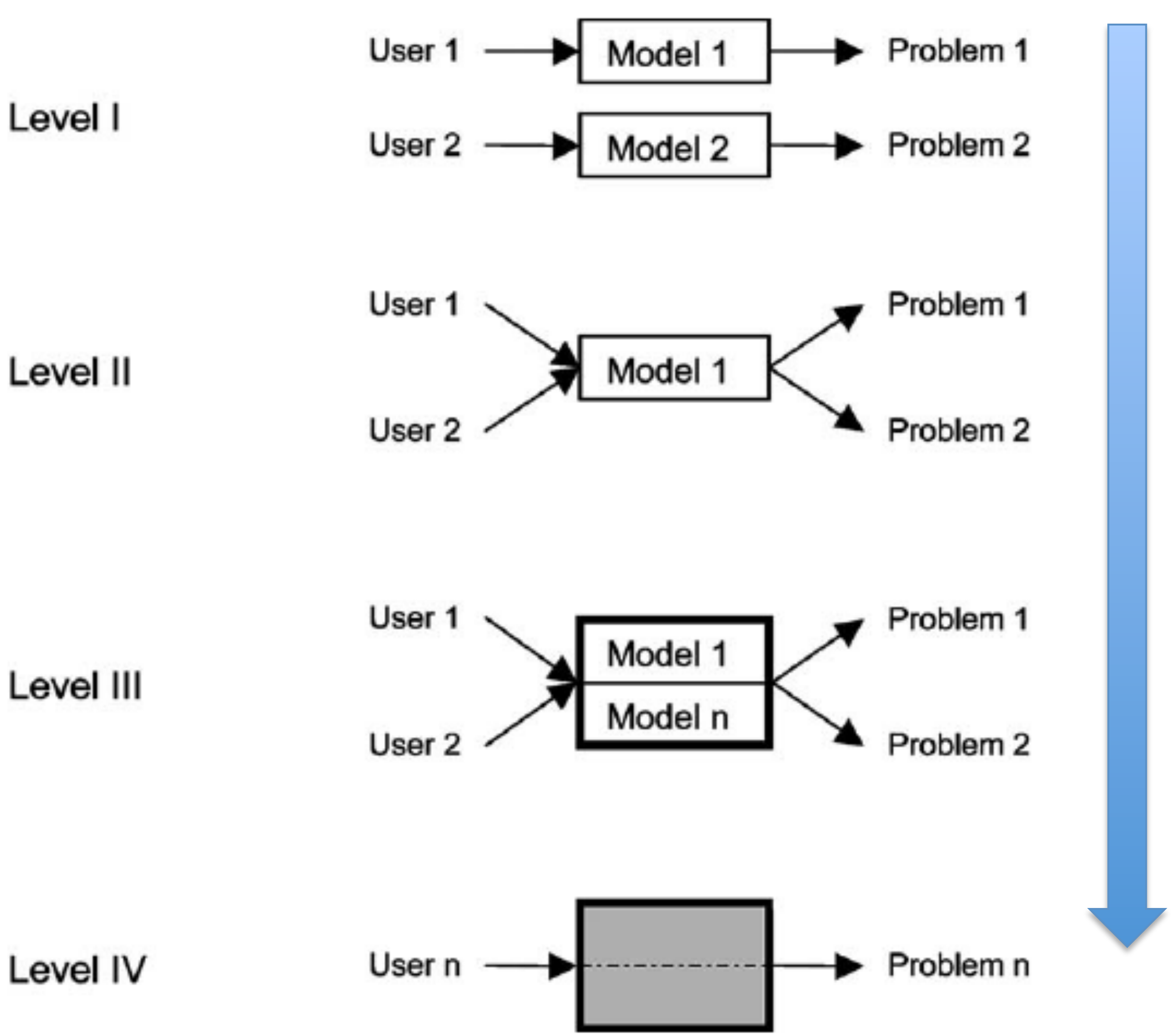
Received 4 November 2002; received in revised form 8 April 2003; accepted 14 April 2003

Abstract

In recent years, pressure has increased on environmental scientist/modellers to both undertake good science in an efficient and timely manner, under increasing resource constraints, and also to ensure that the science being performed is immediately relevant to a particular environmental management context. At the same time, environmental management is changing, with increasing requirements for multi-scale and multi-objective assessment and decision making that considers economic and social systems, as well as the ecosystem. Integration of management activities, and also of the modelling undertaken to support management, has become a high priority. To solve the problems of application and integration, knowledge encapsulation in models is being undertaken in a way that both meets the needs for good science, and also provides the conceptual and technical structures required for broader and more integrated application of that knowledge by managers. To support this modelling, tools and technologies from computer science and software engineering are being transferred to applied environmental science fields, and a range of new modelling and software development approaches are being pursued. The papers in this Special Issue provide examples of the integrated modelling concepts and applications that have been, or are being, developed. These include the use of object-oriented concepts, component-based modelling techniques and modelling frameworks, as well as the emerging use of integrated modelling platforms and metadata support for modelling semantics. This paper provides an overview of the science and management imperatives underlying recent developments, discusses the technological and conceptual developments that have taken place, and highlights some of the semantic, operational and process requirements that need to be addressed now that the technological aspects of integrated modelling are well advanced.

Keywords: Model integration; Component-based modelling; Modelling frameworks

Argent (2004) Model Integration at development and application levels I-IV



Models developed by individuals for a specific research purpose

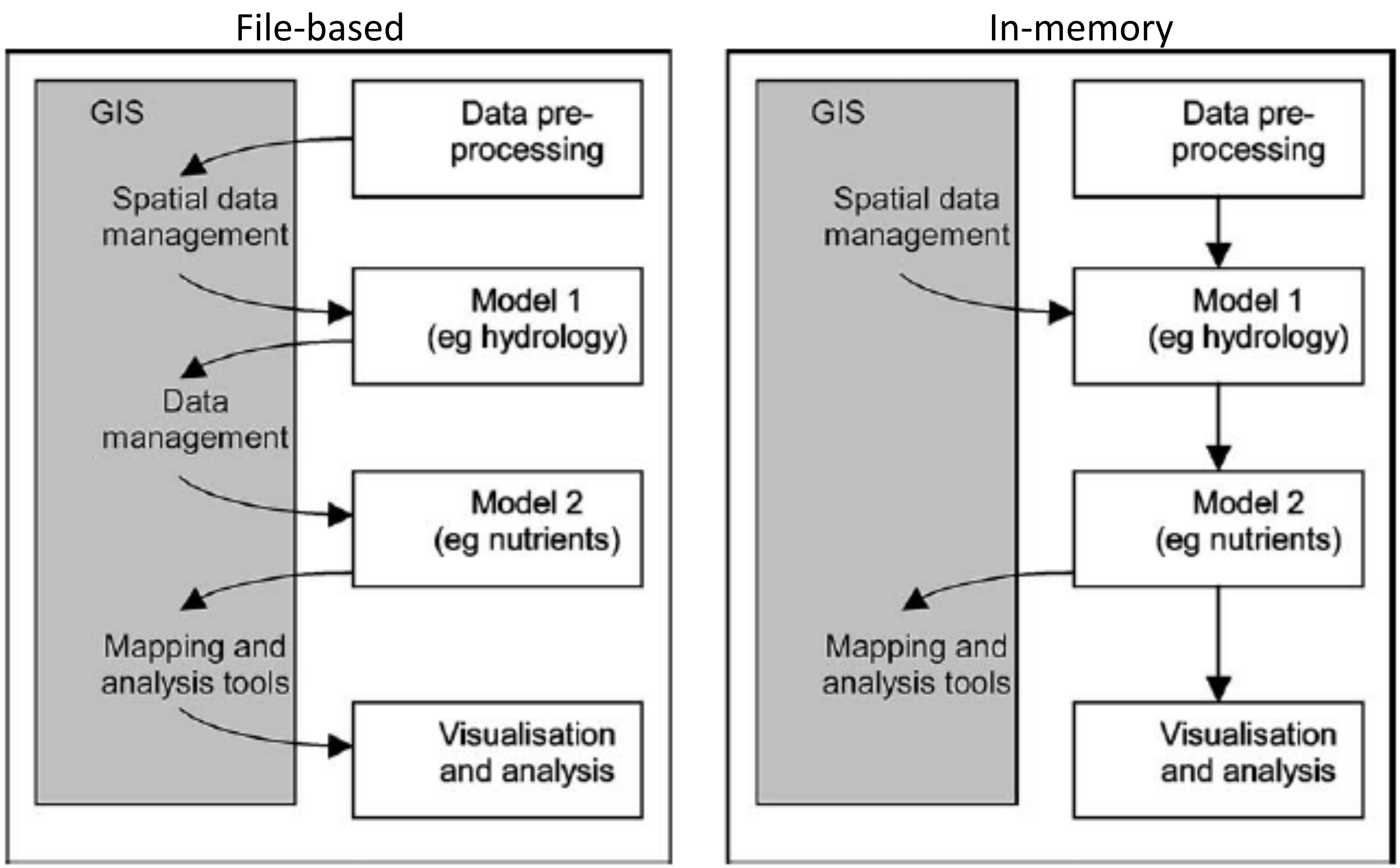
More general models that can be used for a range of research purposes

Flexible application to a range of situations; operations ready

Modelers not familiar with model workings; model operates as black box

Source: Argent, 2004

Two views of model integration



Source: Argent, 2004

Terminology and Key Concepts

- Development environment – a place where components can be built based upon existing components or component templates
- Core components – a library of components that represent various pieces of knowledge
- Support components – a library of components for common tasks (e.g., data collection, gap filling, basic analysis, output, and visualization).
- Modeling framework – the place where modelers construct and manipulate multi-component models.
- Integrated documentation system – for managing components and their metadata
- Resource discovery system – for identification and access of components
- Model execution system – supports model runs potentially in parallel from multiple users.

Source: Argent, 2004

Argent's final paragraph is still true today …

From an integration standpoint, the frameworks developed to date, and those currently under construction, offer tremendous scope for the technical achievement of the modelling practice concepts espoused earlier. We still, however, have the problems within disciplines of developing agreed component structures as well as semantic issues across disciplines for us to overcome. These are not things that can be solved by individuals. By addressing these in a shared manner, possibly through shared cross-disciplinary development and communication, we can work towards are more successful application of knowledge in solving current and future environmental management problems.

Source: Argent, 2004

What are these *semantic issues*?

- "the meaning of data, variable and parameter names that are loaded into, used by, or exchanged between, components."

- "misunderstandings will arise unless there is a clear and established meaning for various state variables and concepts, and an agreed language for communicating these."

Source: Argent, 2004

Argent's prediction for progress in semantic issues

- First, we need to acknowledge that these semantic difficulties exist and that the difficulties will be expanded as the scope of integrated modeling expands

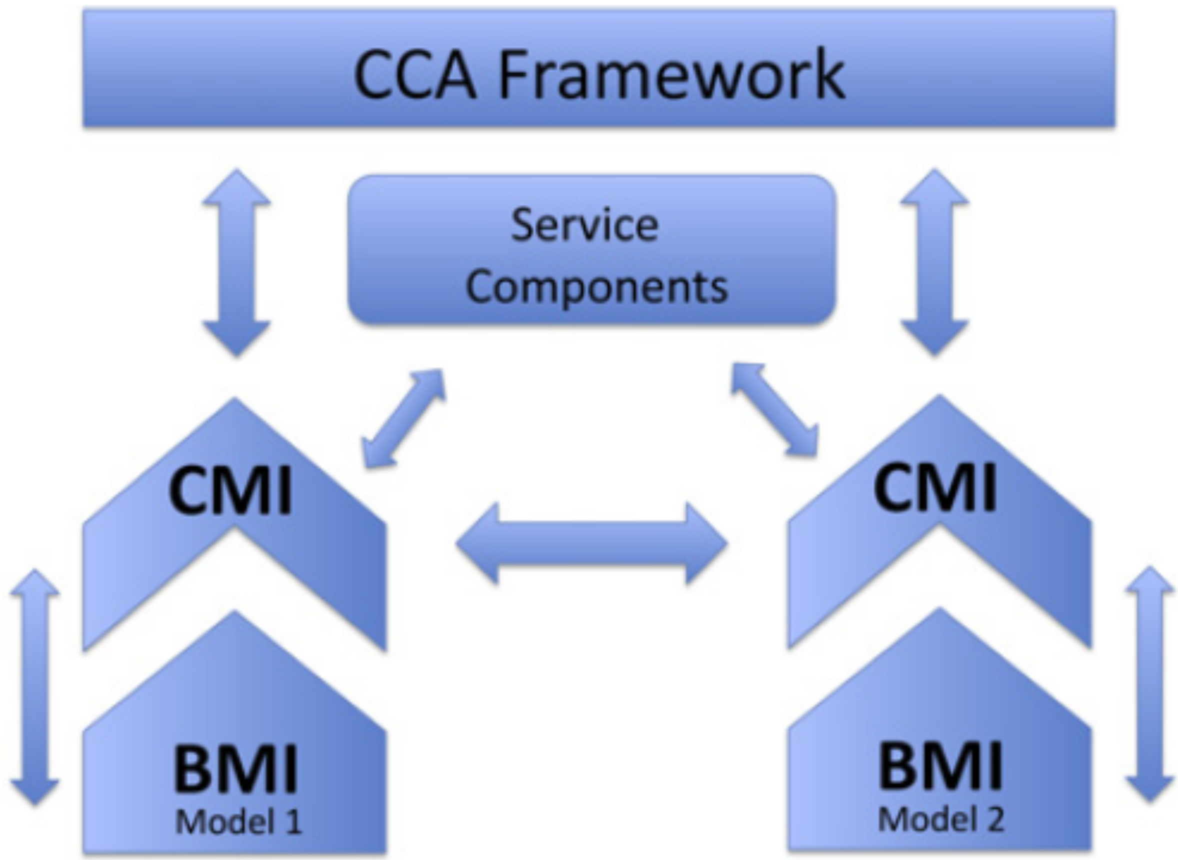
- Then, "as we gain experience in working with and explaining our terminology to members of other disciplines, we will develop a more semantically rich language"

Argent, 2004

Outline

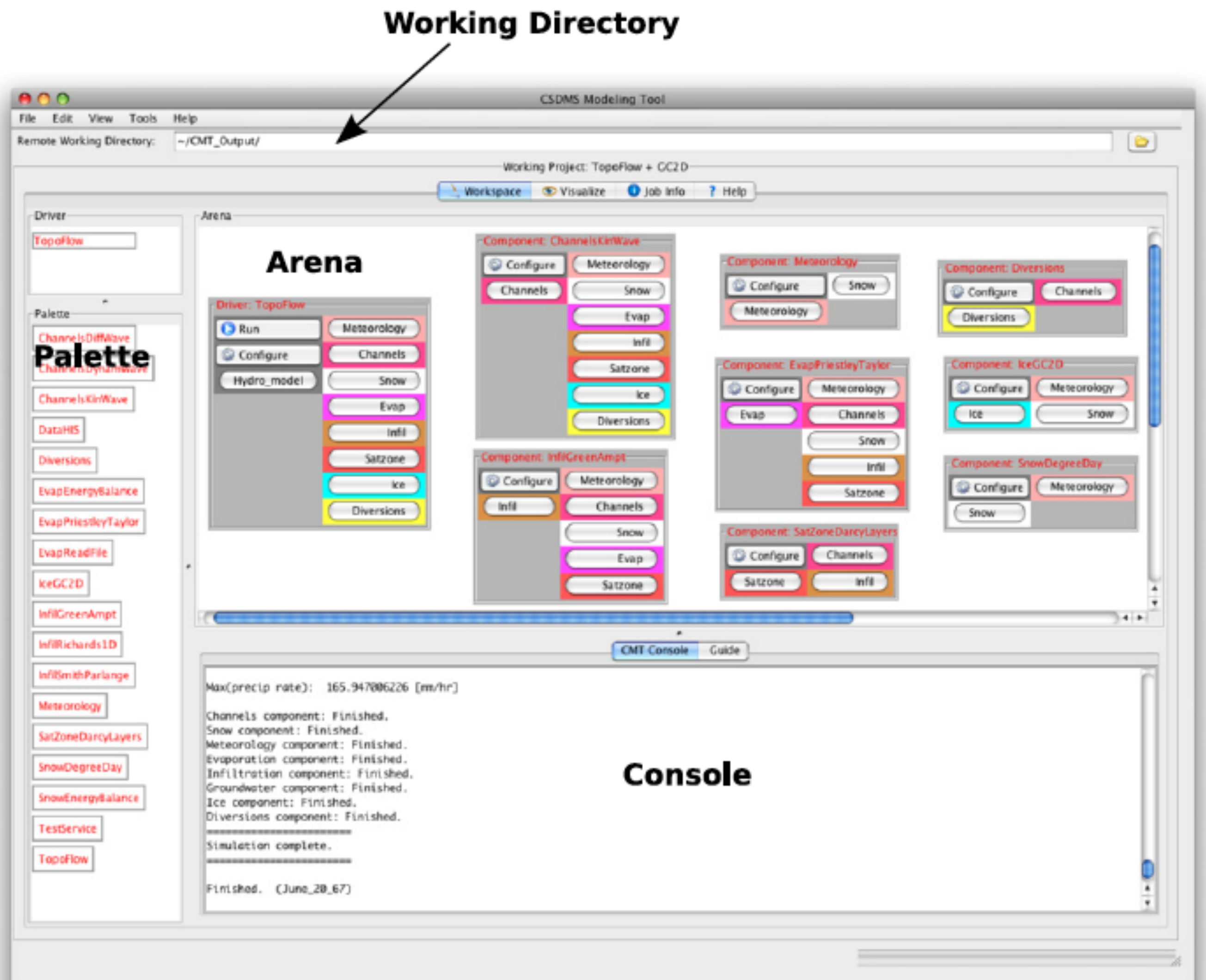
- Introduction
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- **Example modeling frameworks**
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- Interface Standards
  - OpenMI
  - BMI
- Current topics
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- Tutorials (time permitting)
  - Getting Started with WMT
- Summary

CSDMS Architecture



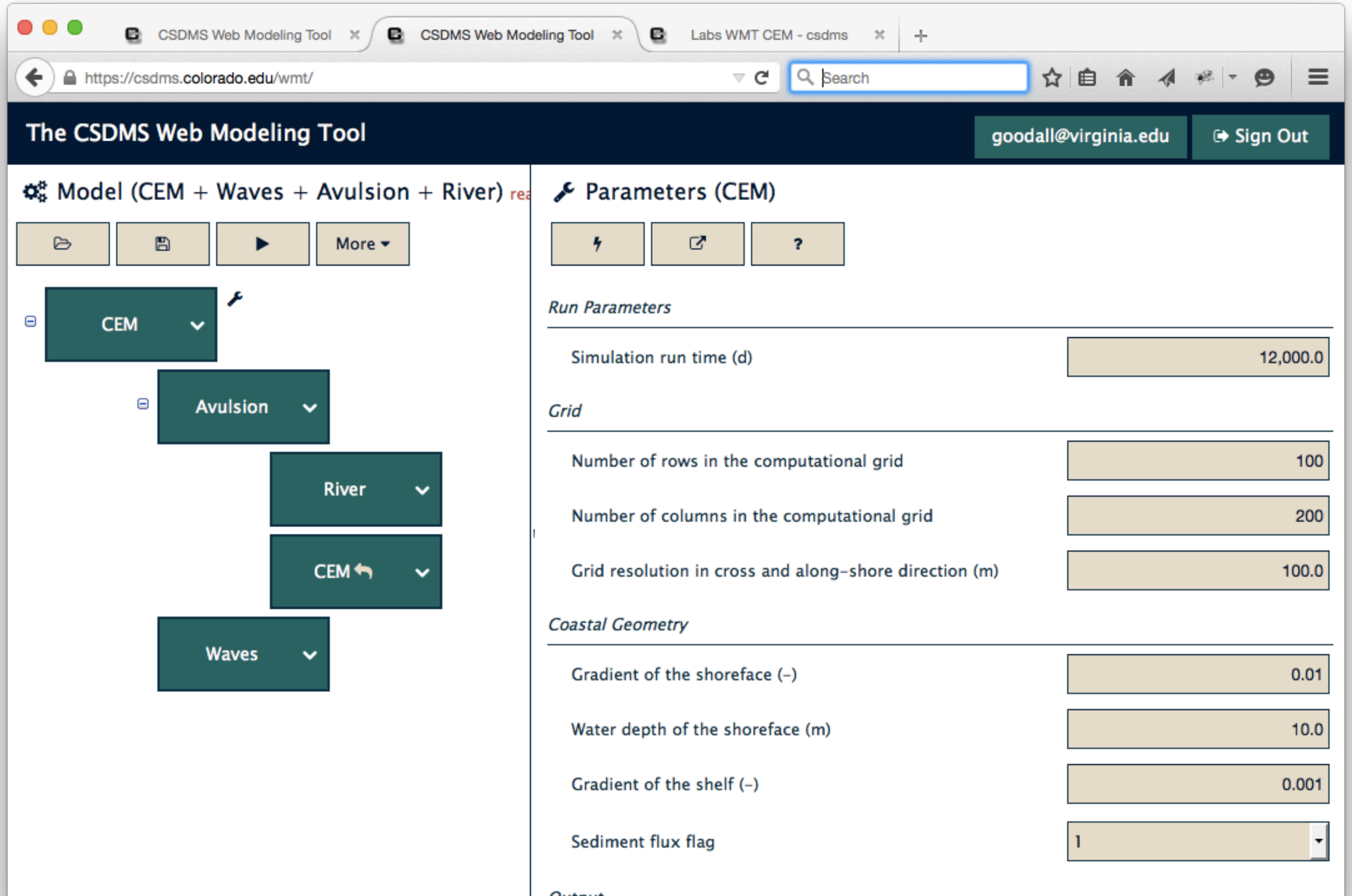
Source: Peckham et al., 2013

Original CSDSM Modeling Tool (replaced by by Web Modeling Tool: WMT)

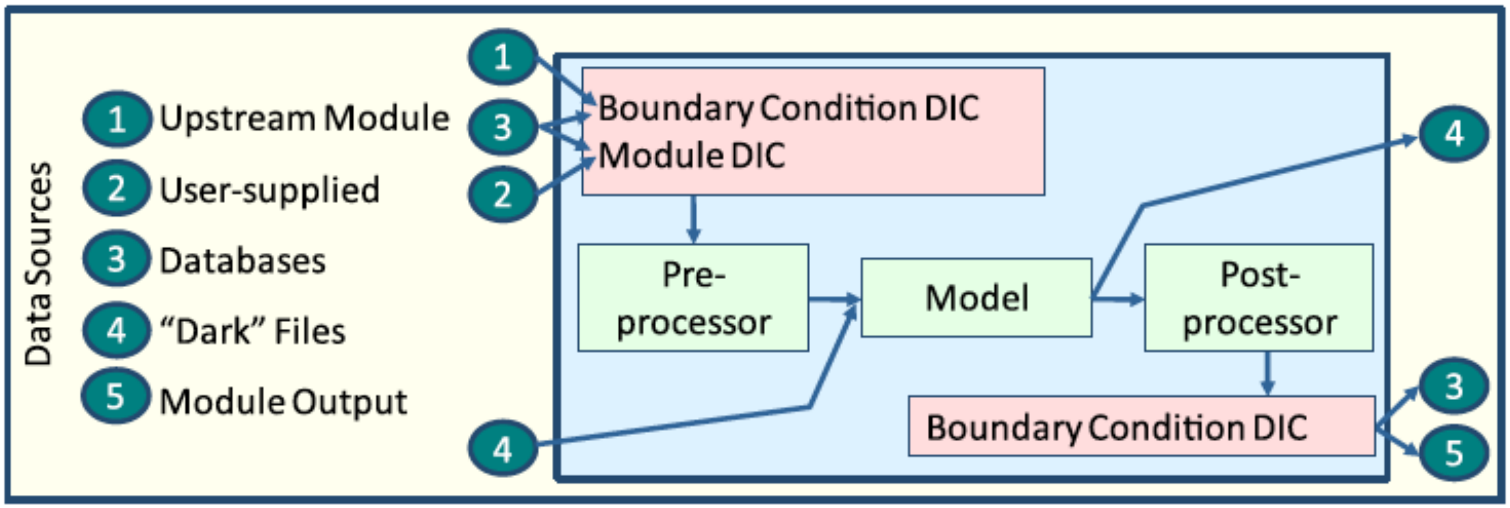


Source: Peckham et al., 2013

CSDMS Web Modeling Tool (WMT)

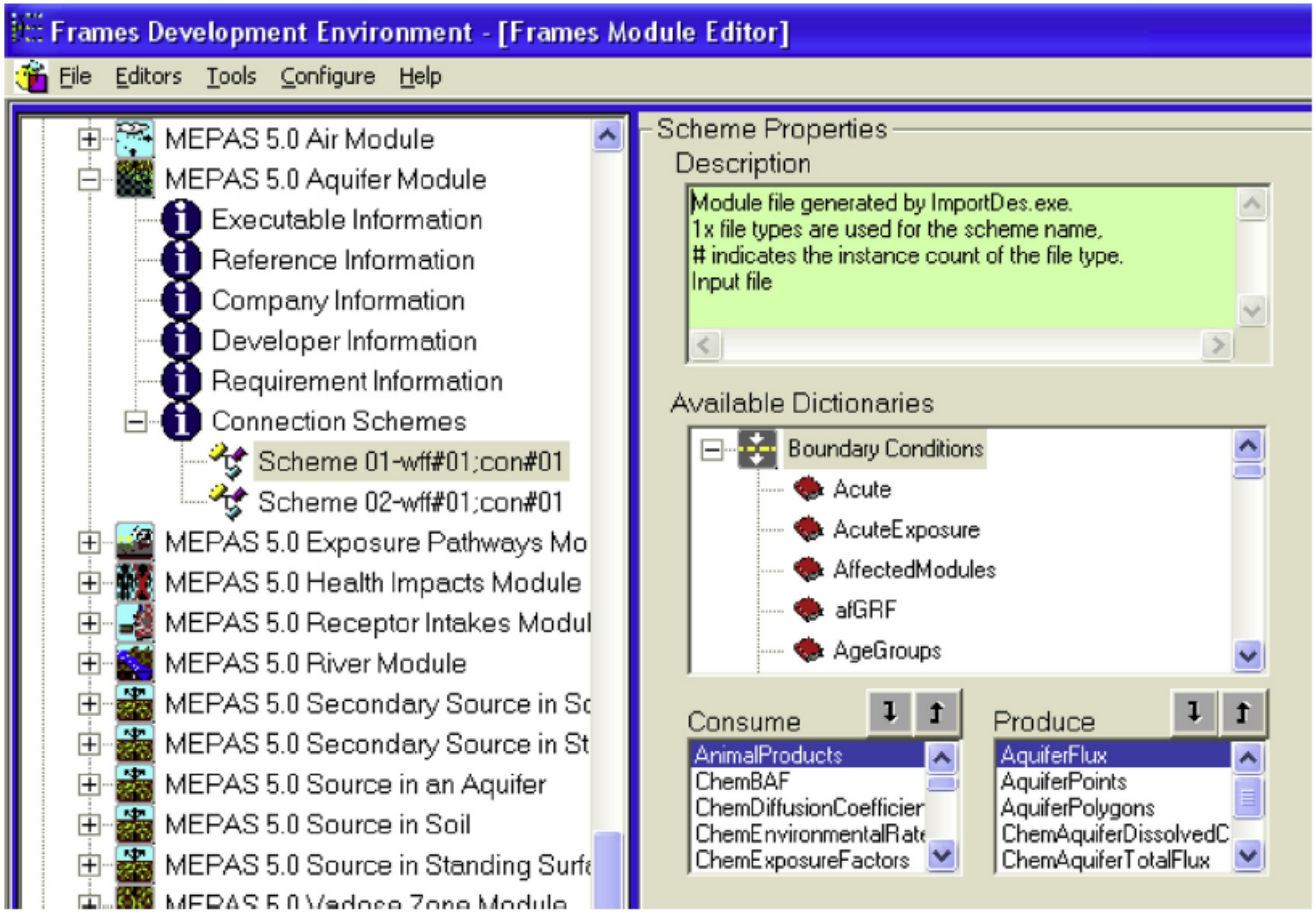


EPA FRAMES



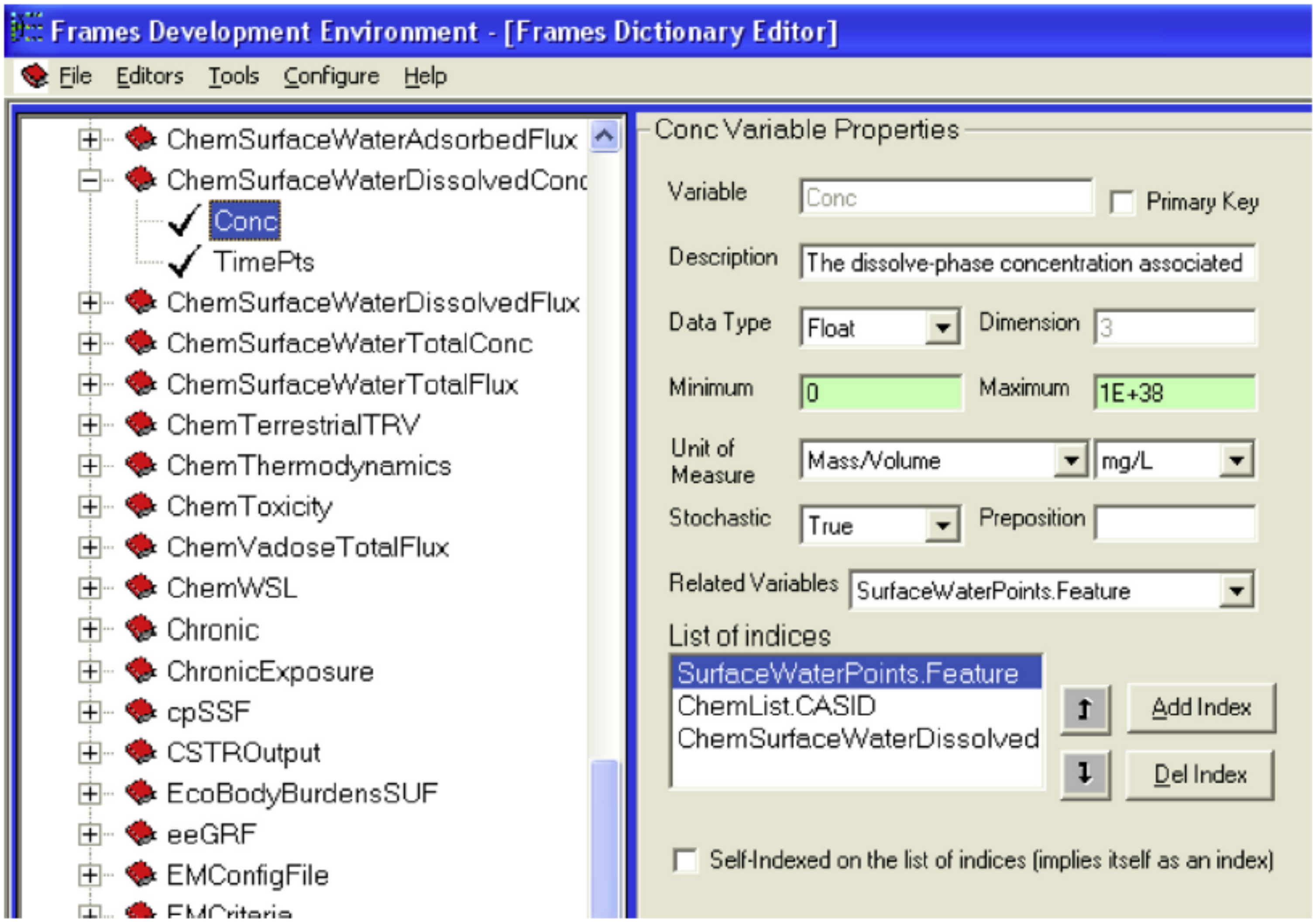
Whelan et al., 2014

EPA FRAMES Module Editor



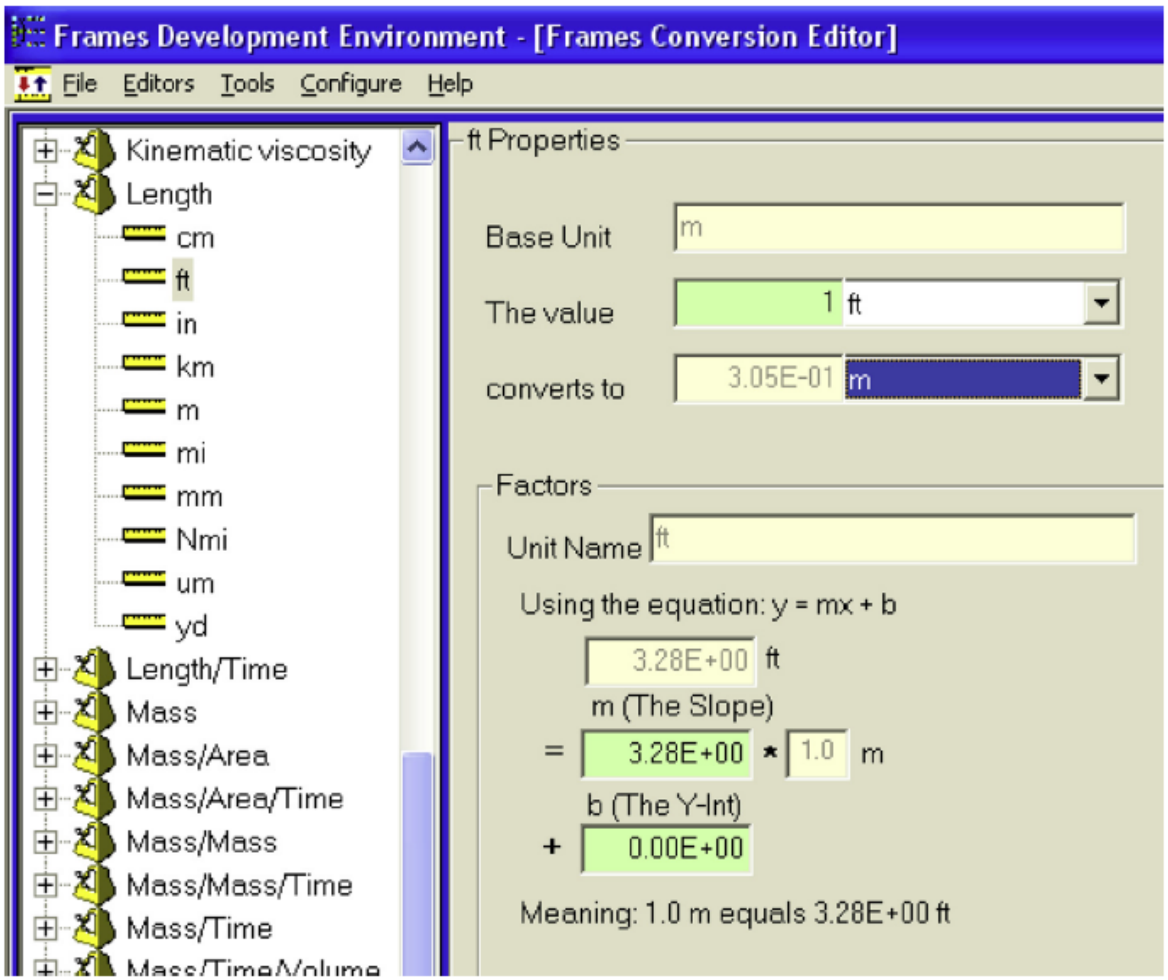
Whelan et al., 2014

Frames Dictionary Editor



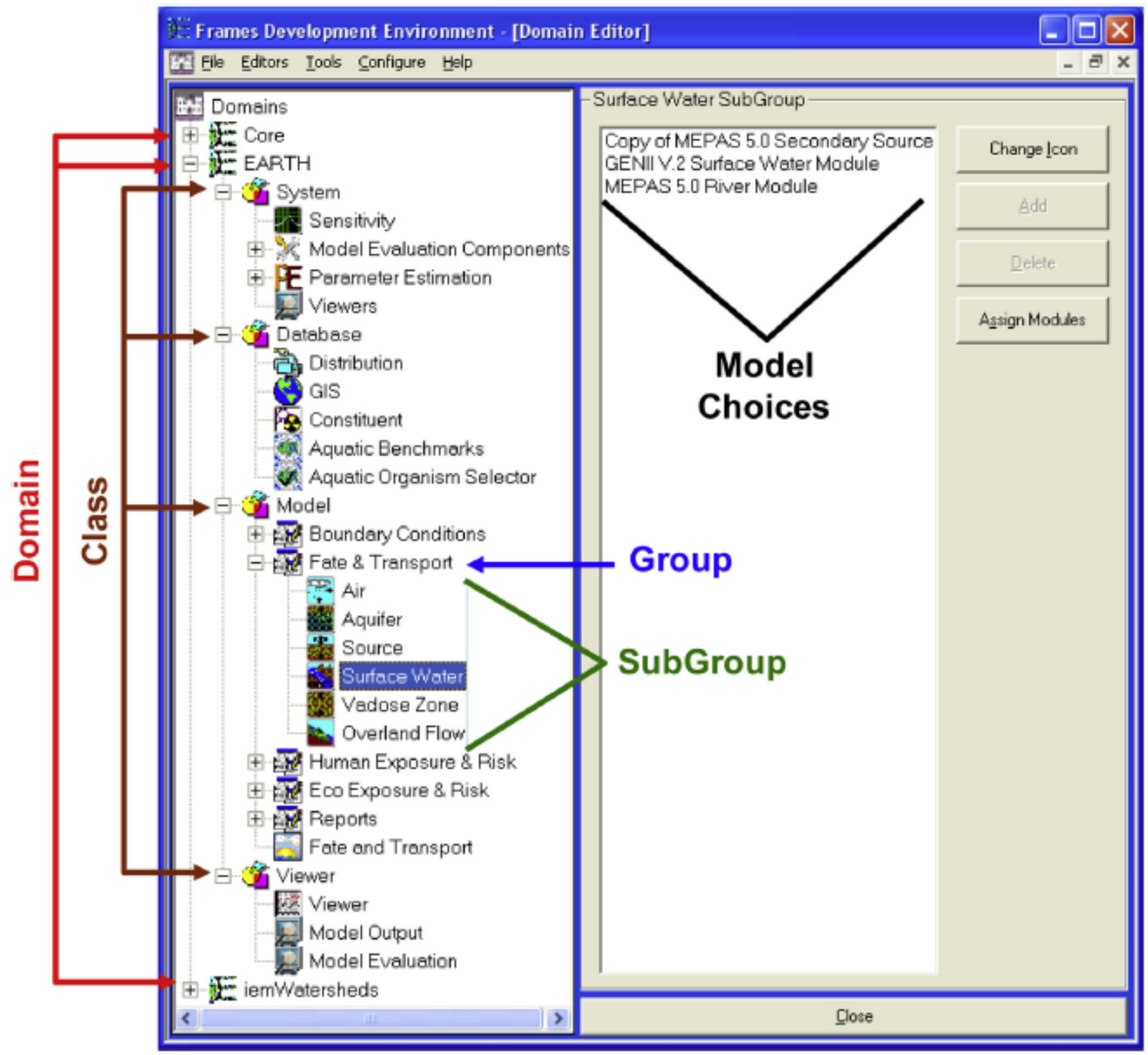
Whelan et al., 2014

Frames Unit Conversation Editor



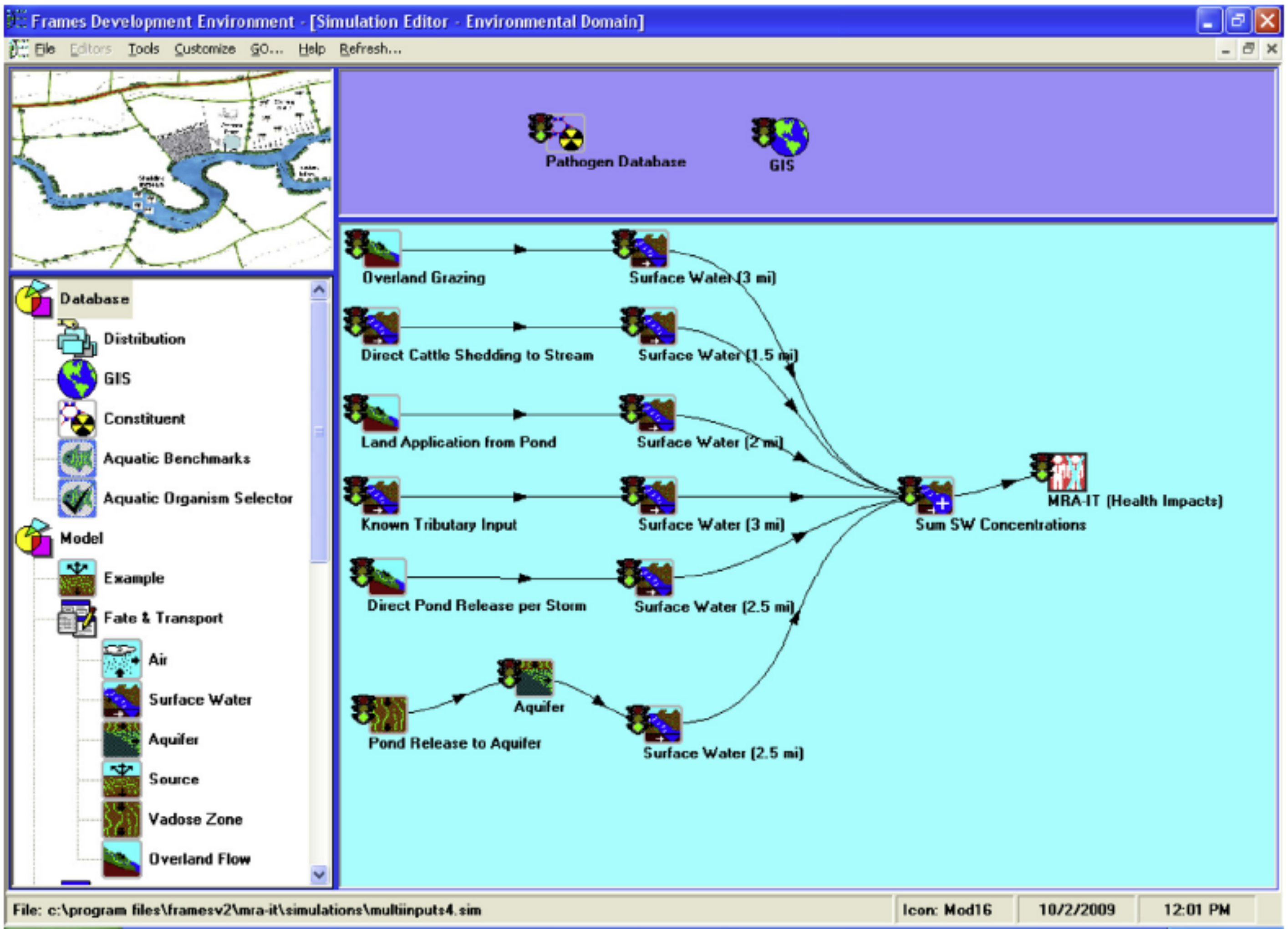
Whelan et al., 2014

FRAMES Domain Editor



Whelan et al., 2014

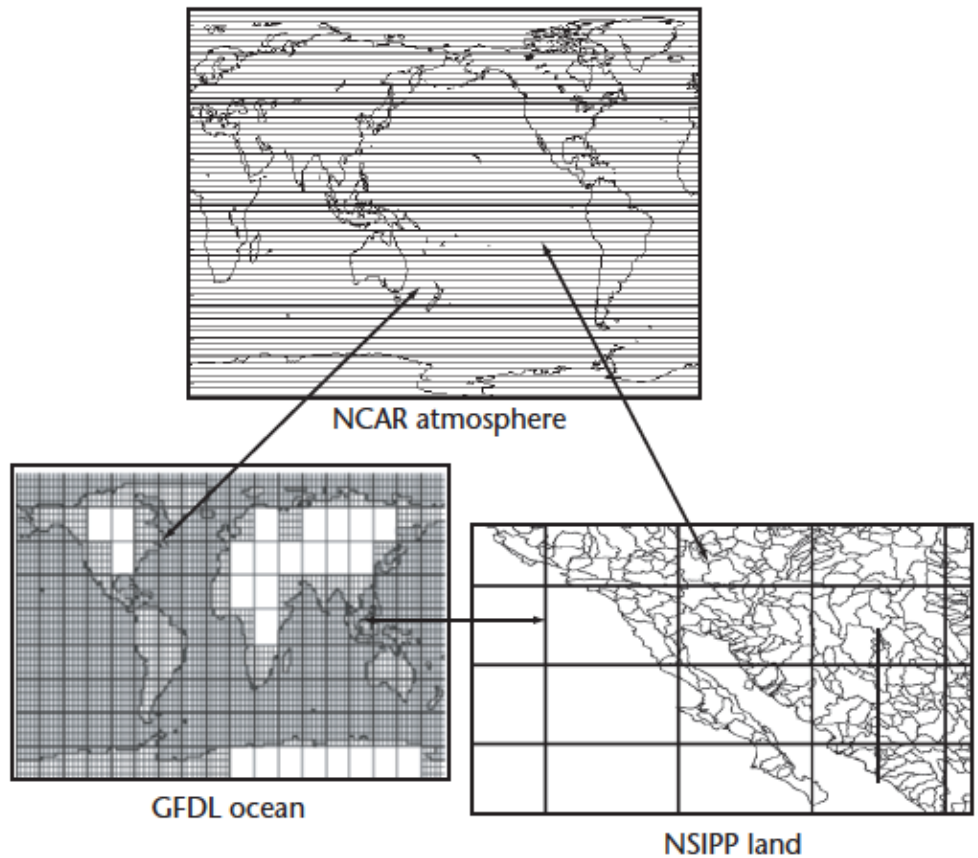
FRAMES Simulation Editor



Whelan et al., 2014

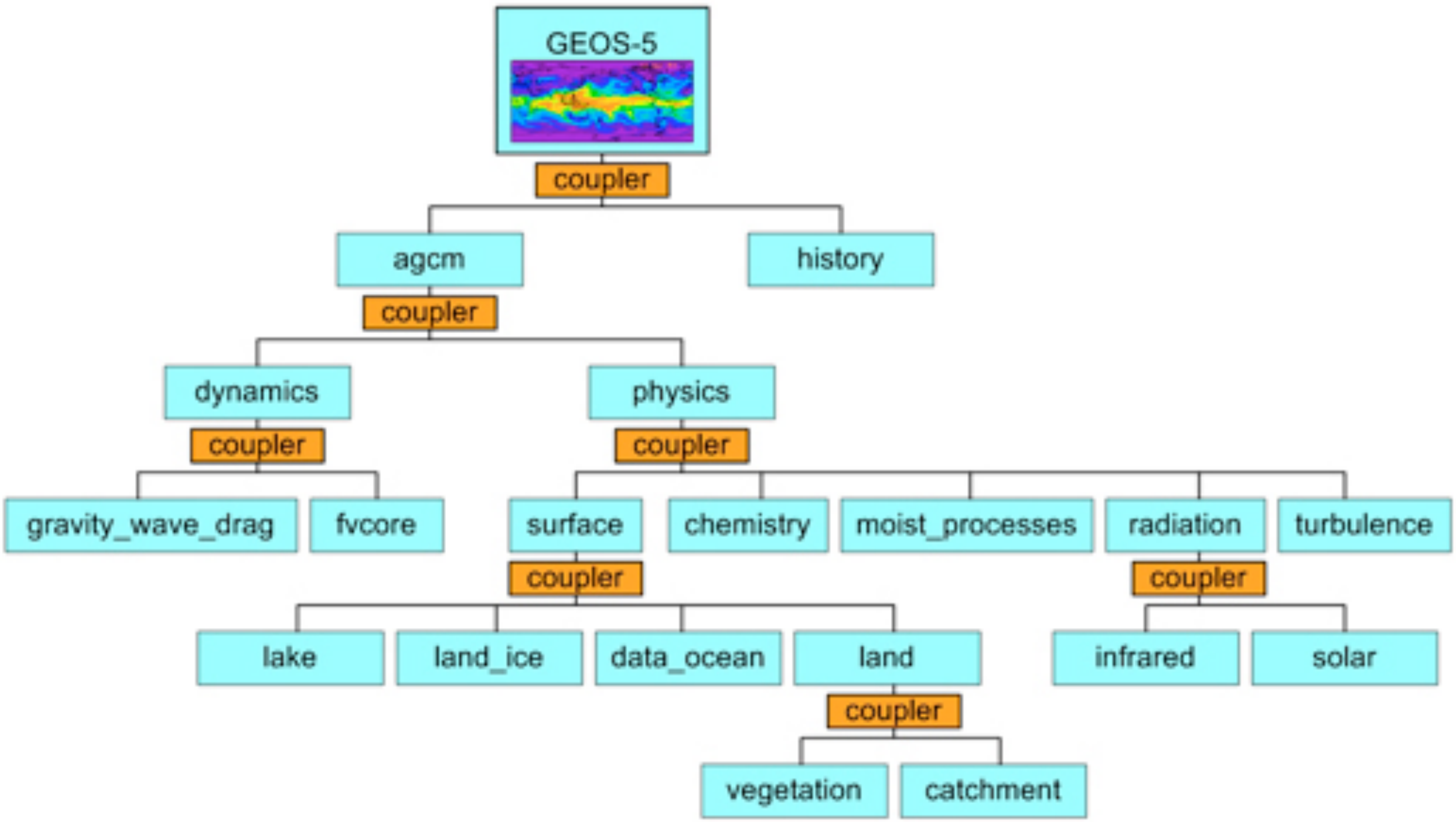
ESMF

- Earth System Modeling Framework
- Global models
- HPC emphasis
- Coupling models with gridded data structures



Source: Hill et al., 2004

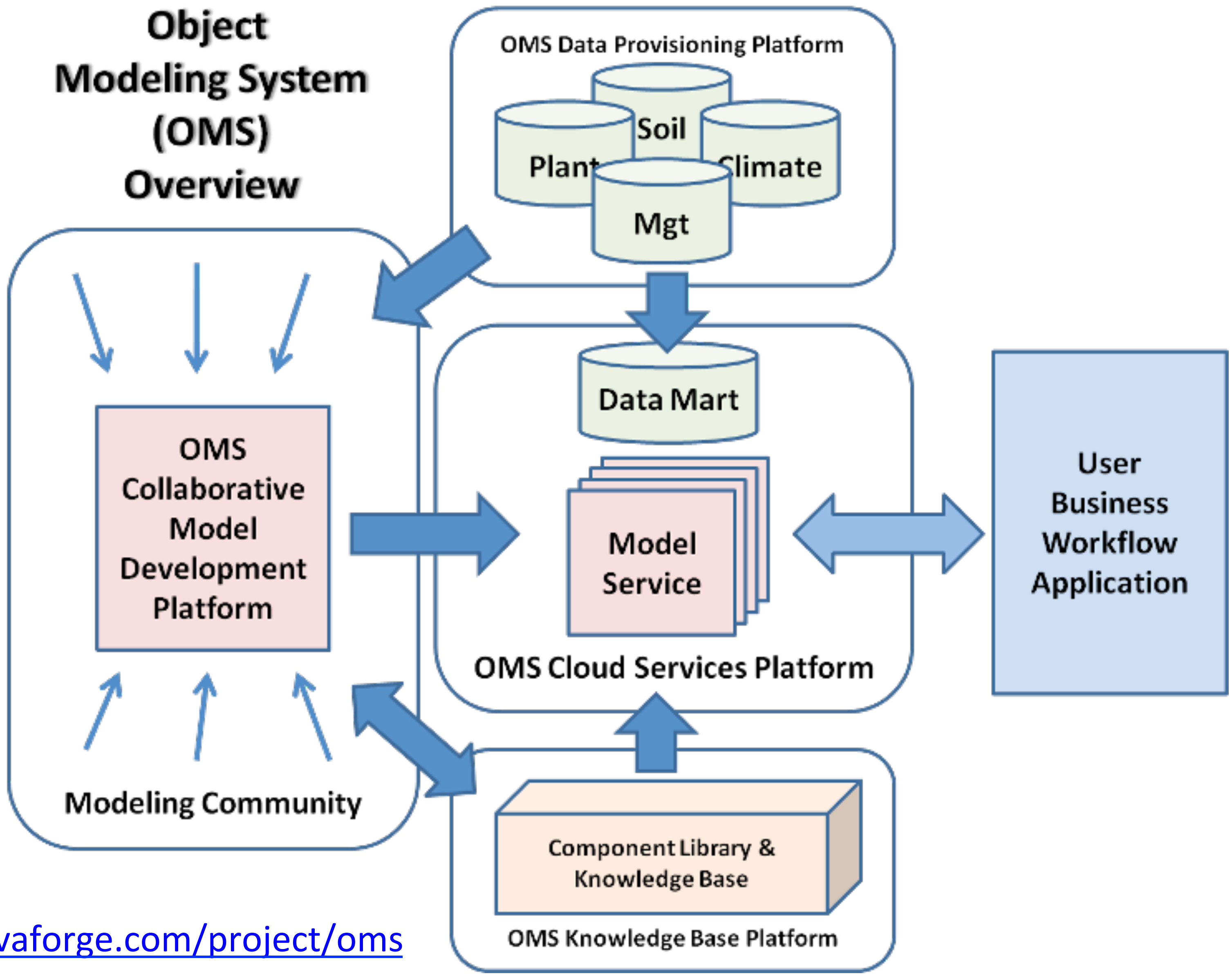
Components and Couplers in ESMF



Some of the ESMF Benefits

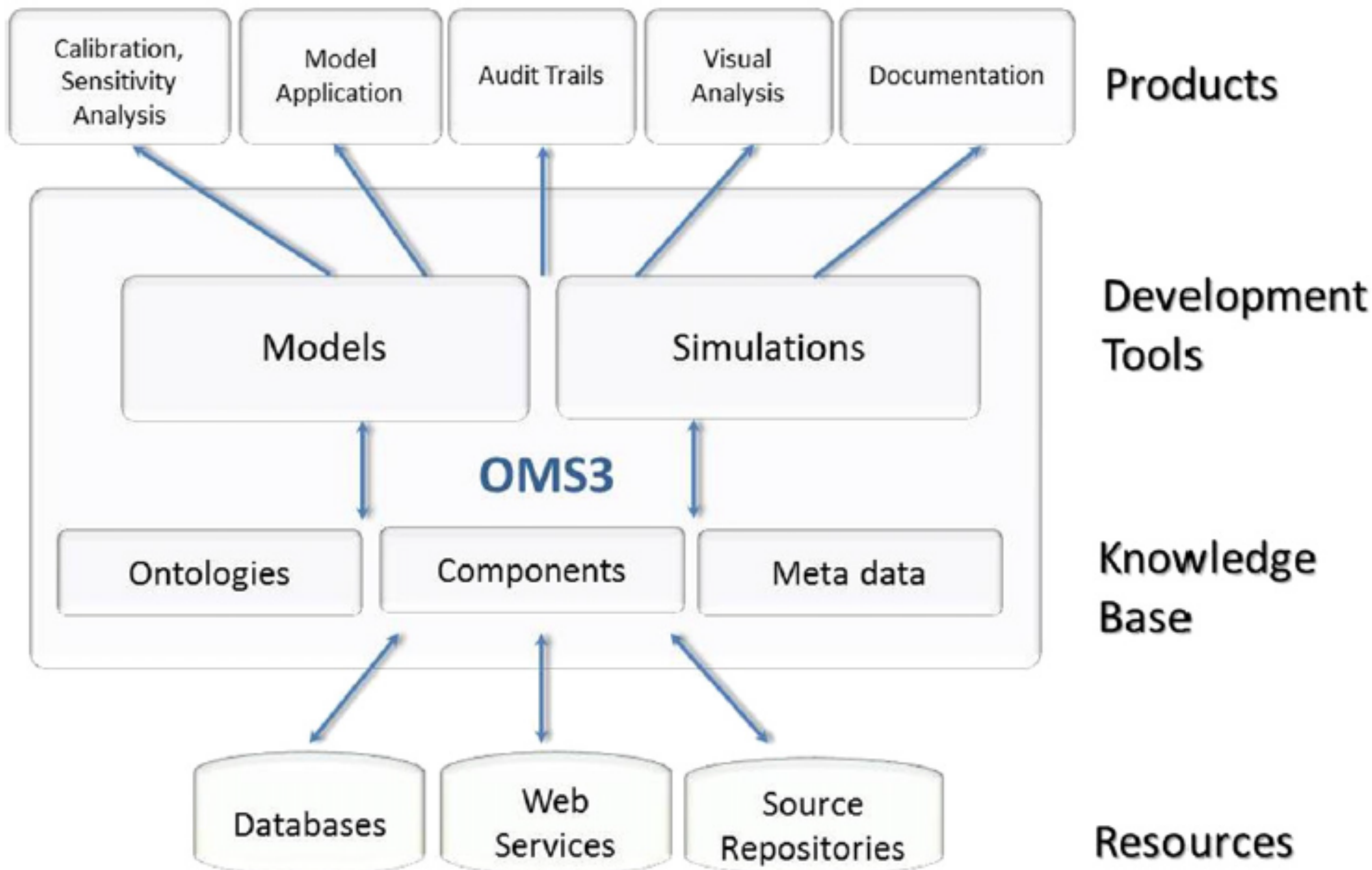
- Well tested and used in operations
- Optimized to reduce overhead due to coupling in high performance computing (HPC) environments
- Parallel regridder with Python bindings that can be used apart from ESMF (e.g., it is used by CSDMS)

Object Modeling System 3 (OMS3)



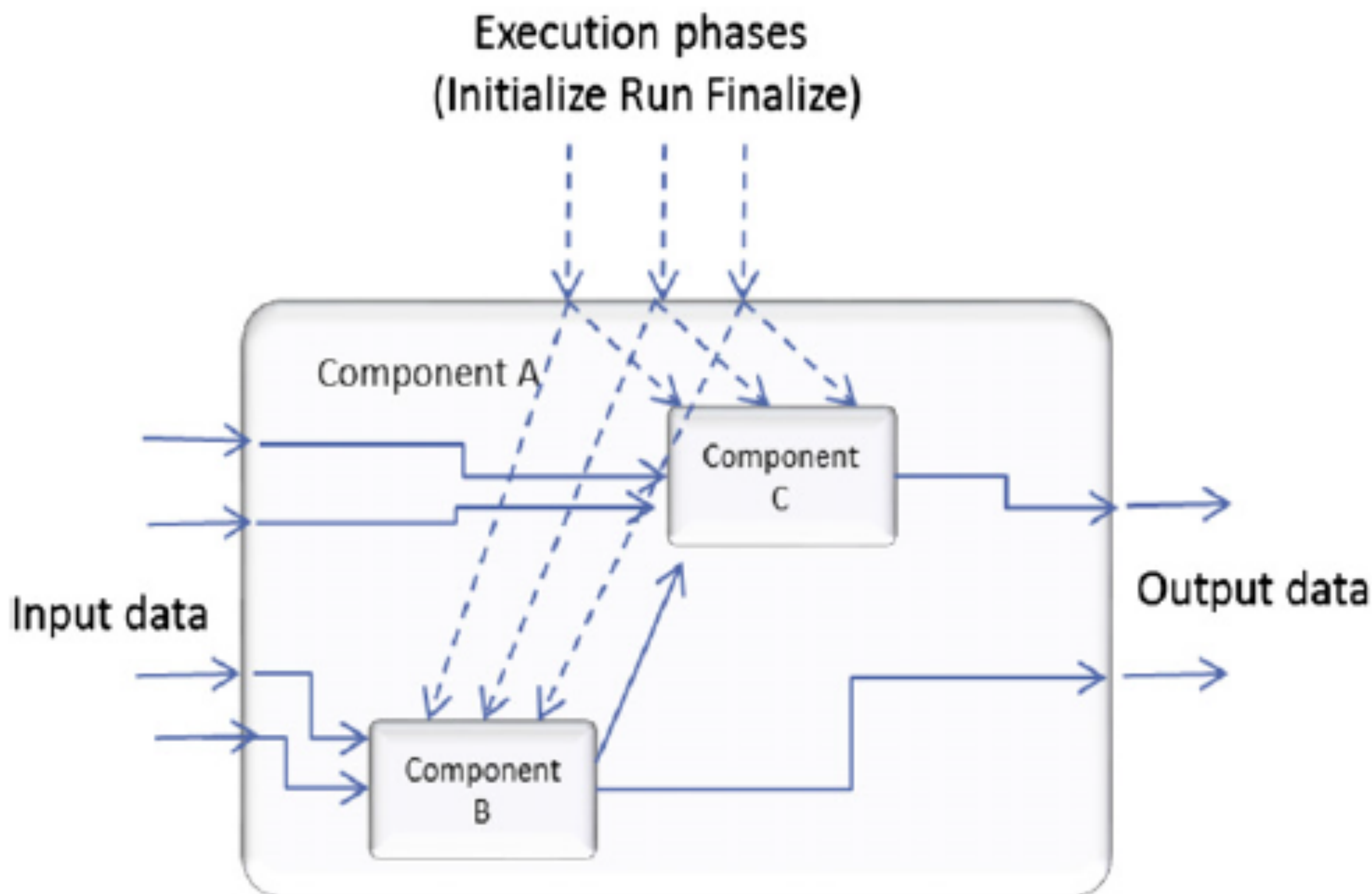
http://www.javaforge.com/project/oms

Object Modeling System 3 (OMS3): Principle Framework Architecture



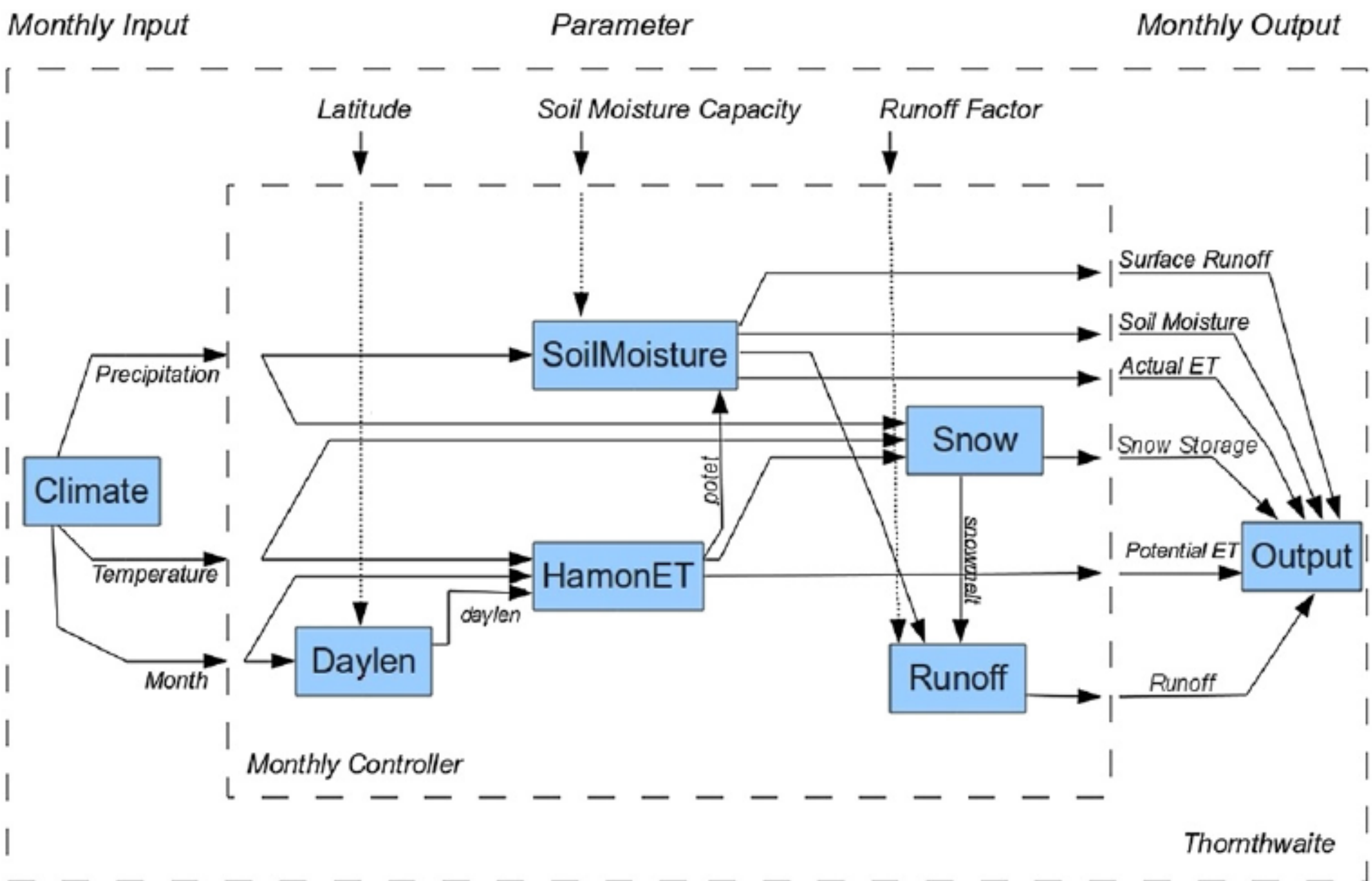
Source: David et al., 2013

OMS3 Component Architecture



Source: David et al., 2013

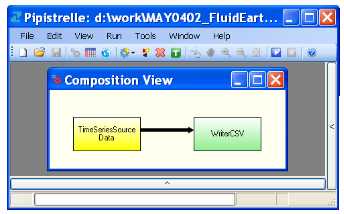
Example hydrologic model implemented as components in OMS3



Source: David et al., 2013

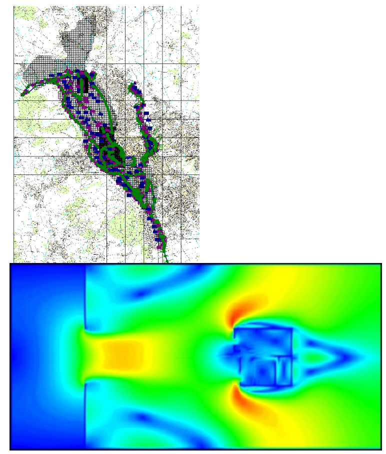
FluidEarth

FluidEarth is a functional and technical platform for using OpenMI.



Tools

Fluid Earth SDK
Pipistrelle GUI
(Reference Implementations for OpenMI 2.0)



Community

Model providers and users

eInfrastructure

http://fluidearth.net
http://catalogue.fluidearth.net
http://sourceforge.net/projects/fluidearth/

Models

A library of models available for compositions

Source: HR Wallingford, fluidearth.net

The FluidEarth Toolkit: Pipistrelle & SDK

The *FluidEarth Software Development Kit* is an OpenMI API:

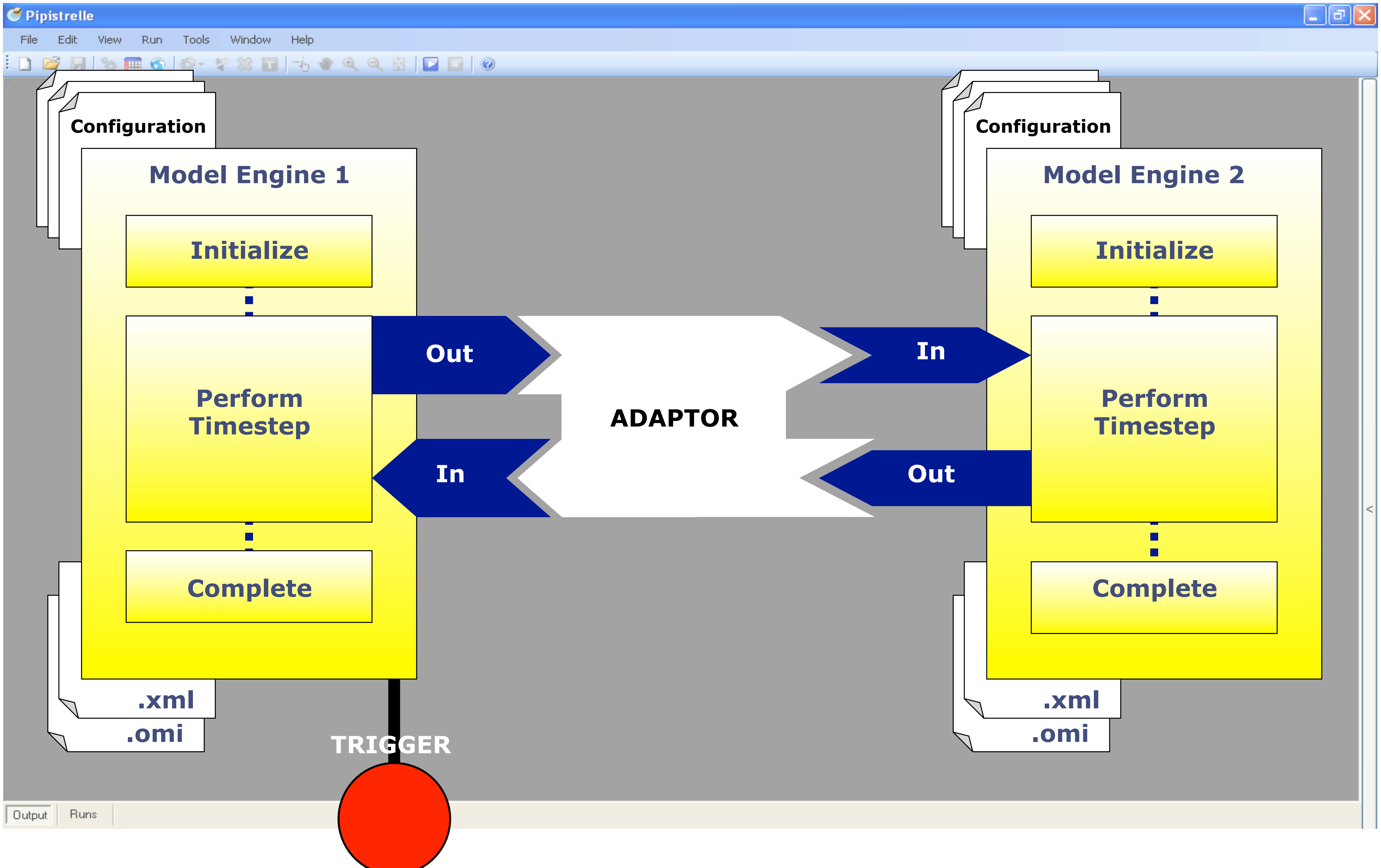
- Creation of OpenMI 2.0 in C#, VB and FORTRAN;
- Windows or Linux (under Mono);
- Available on SourceForge in "Fluid Earth" project.

Pipistrelle is a tool for running OpenMI compositions:

- OpenMI 2.0;
- Windows (with GUI) or Linux (under Mono, with console only);
- Adaptor functionality;
- Component Builder Plug-in (initial version);
- Compatibility with and conversion of FE OpenMI 1.4 components into FE OpenMI 2.0 components to follow (if Data Operations not used);
- Spatial View Plug-in (to follow);
- Available on Source Forge in "Fluid Earth" project.

Source: HR Wallingford, fluidearth.net

2-way Connection

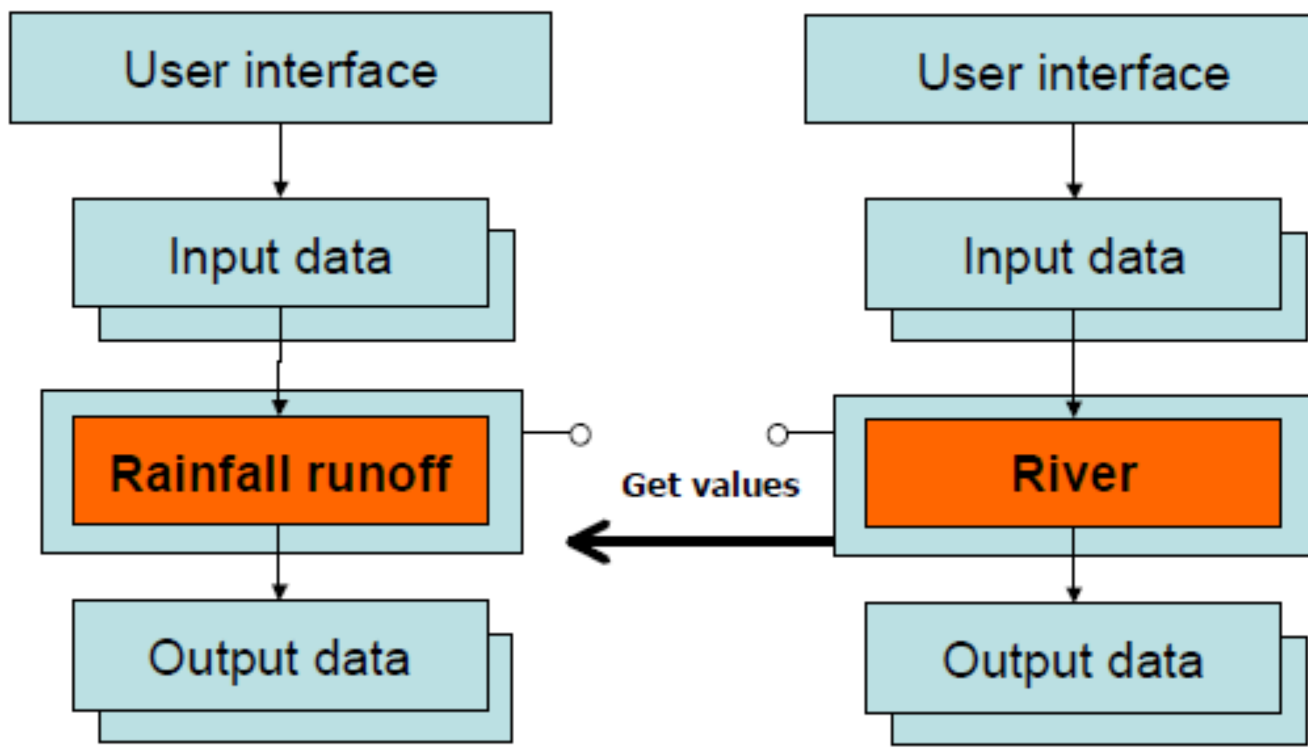


Source: HR Wallingford, fluidearth.net

Outline

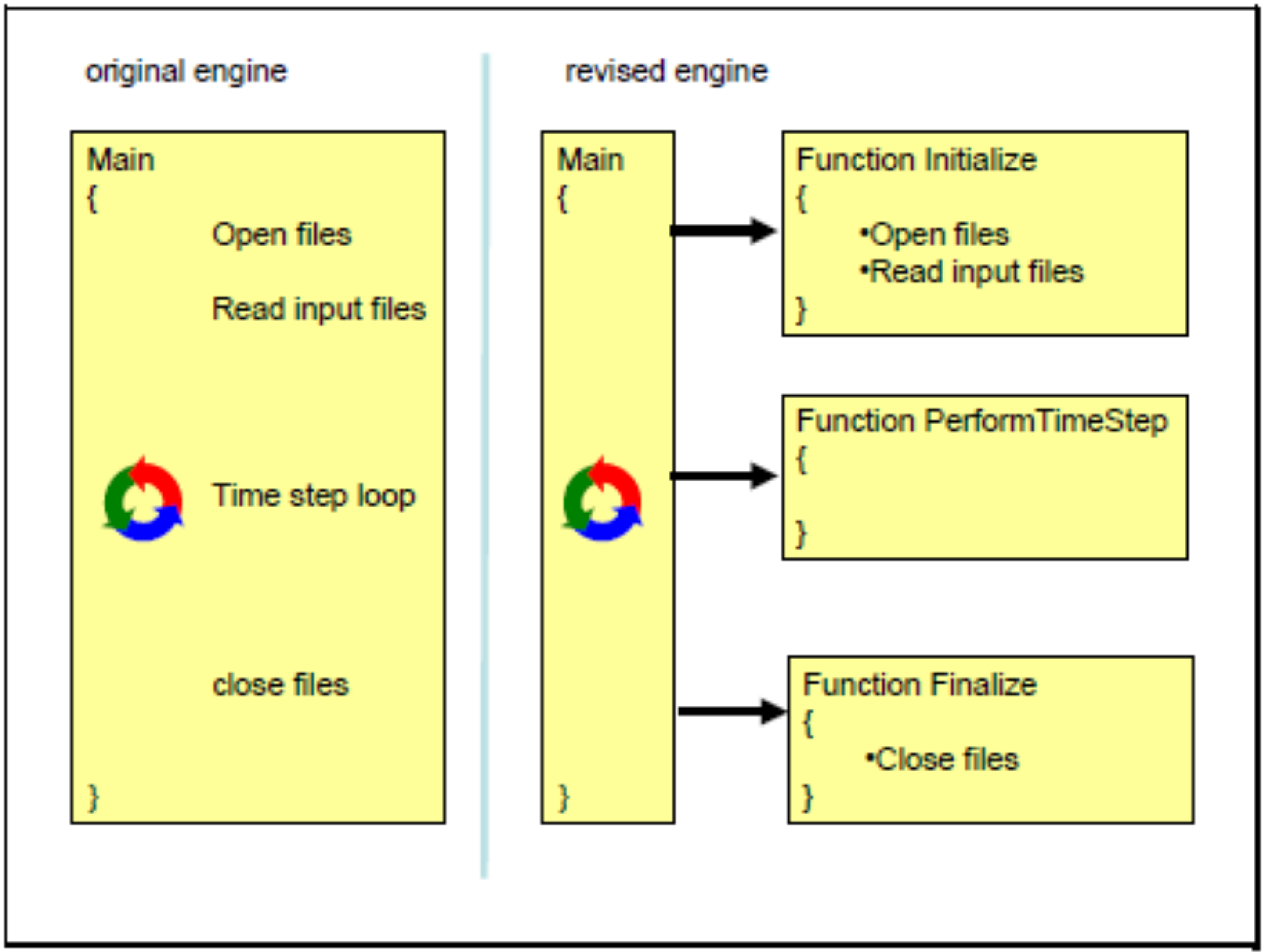
- Introduction
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- Example modeling frameworks
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- **Interface Standards**
  - OpenMI
  - BMI
- Current topics
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- Tutorials (time permitting)
  - Getting Started with WMT
- Summary

Run-time Coupling between Model Engines



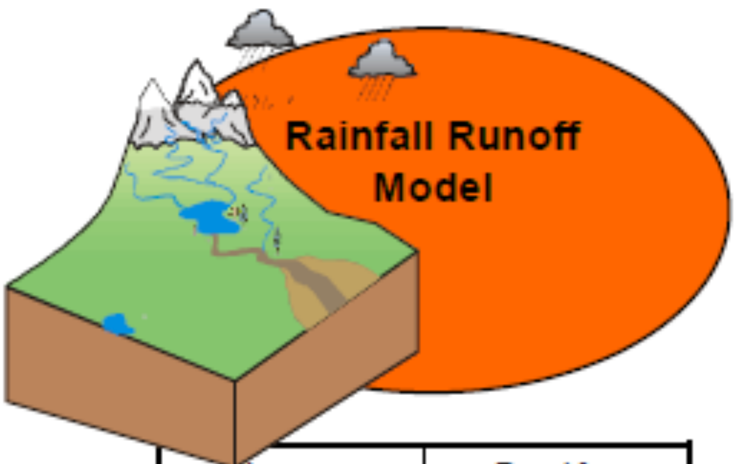
Source: Moore and Tindall, 2005

Refactoring Model to IRF Paradigm

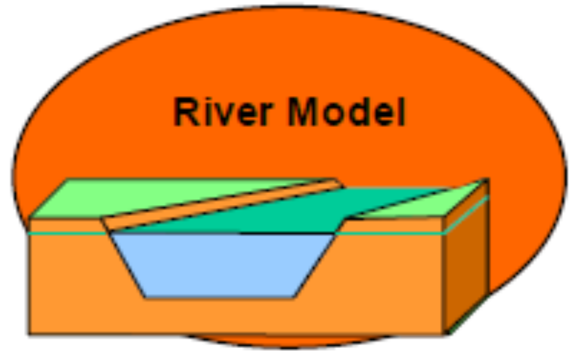


Source: Moore and Tindall, 2005

Manual Semantic Mediation



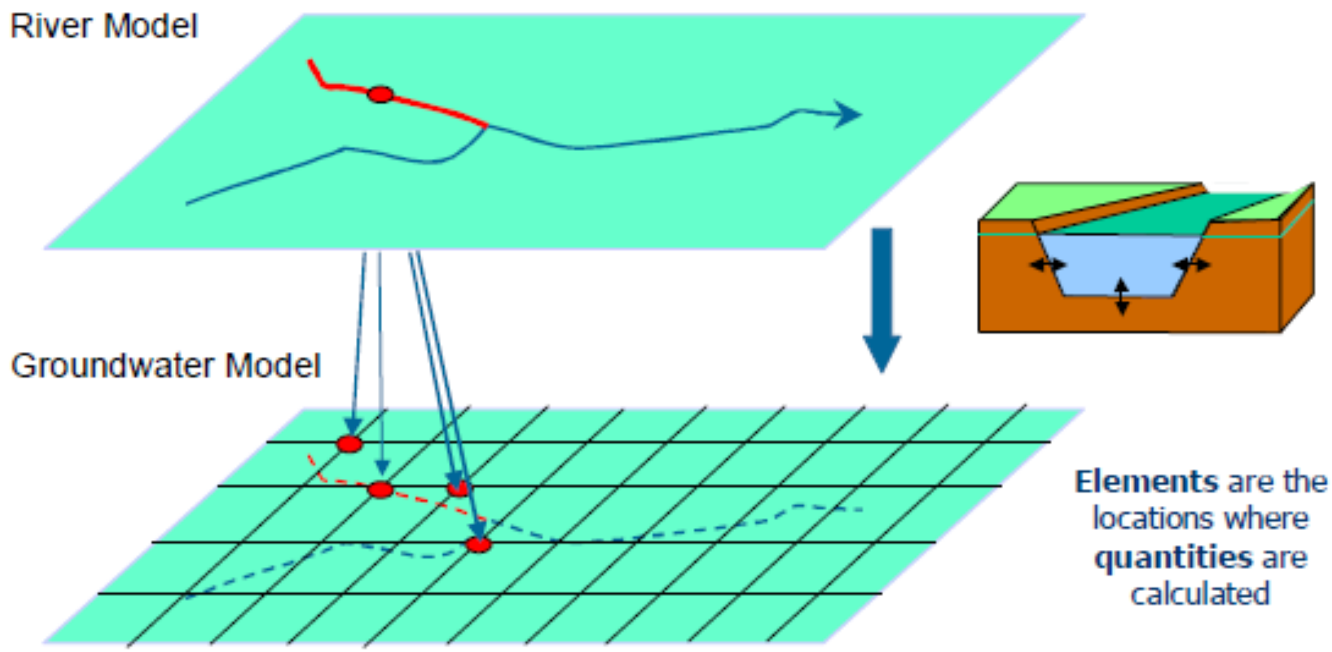
Accepts	Provides
Rainfall (mm)	Runoff (m^3/s)
Temperature (Deg C)	
Evaporation (mm)	



Accepts	Provides
Upstream Inflow (m^3/s)	Outflow (m^3/s)
Lateral inflow (m^3/s)	
Abstractions (m^3/s)	
Discharges (m^3/s)	

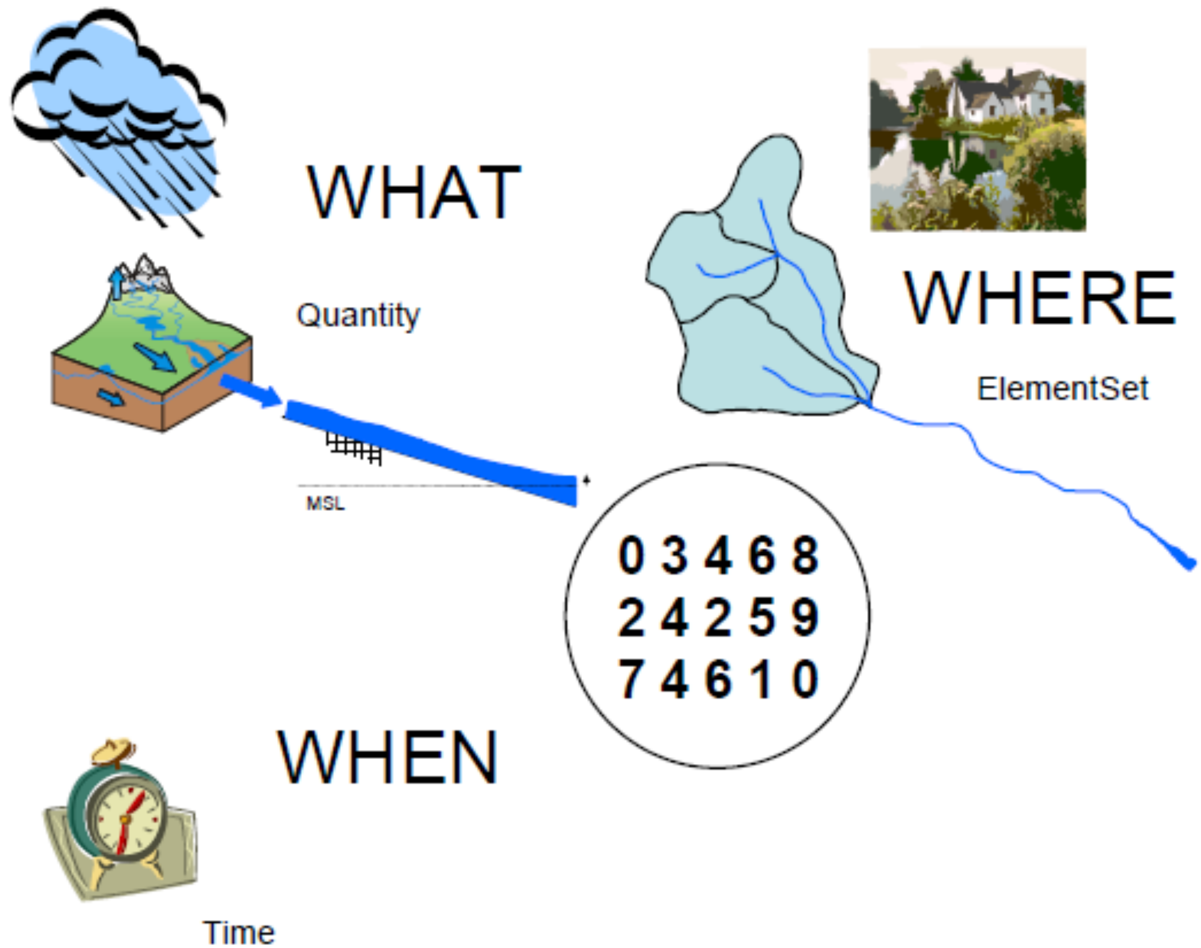
Source: Moore and Tindall, 2005

Spatial Referencing of Model Data Exchanges



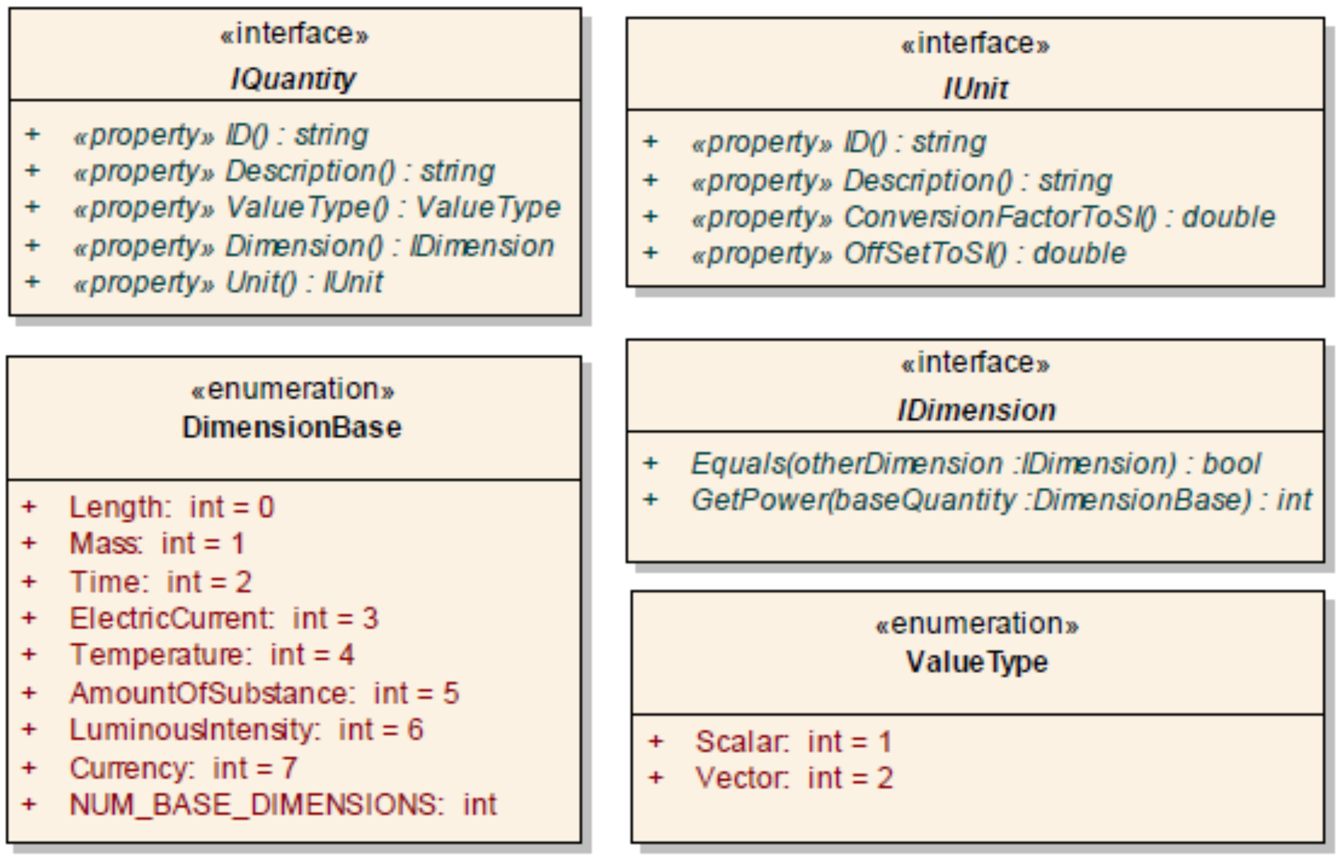
Source: Moore and Tindall, 2005

Data model for exchanges between models



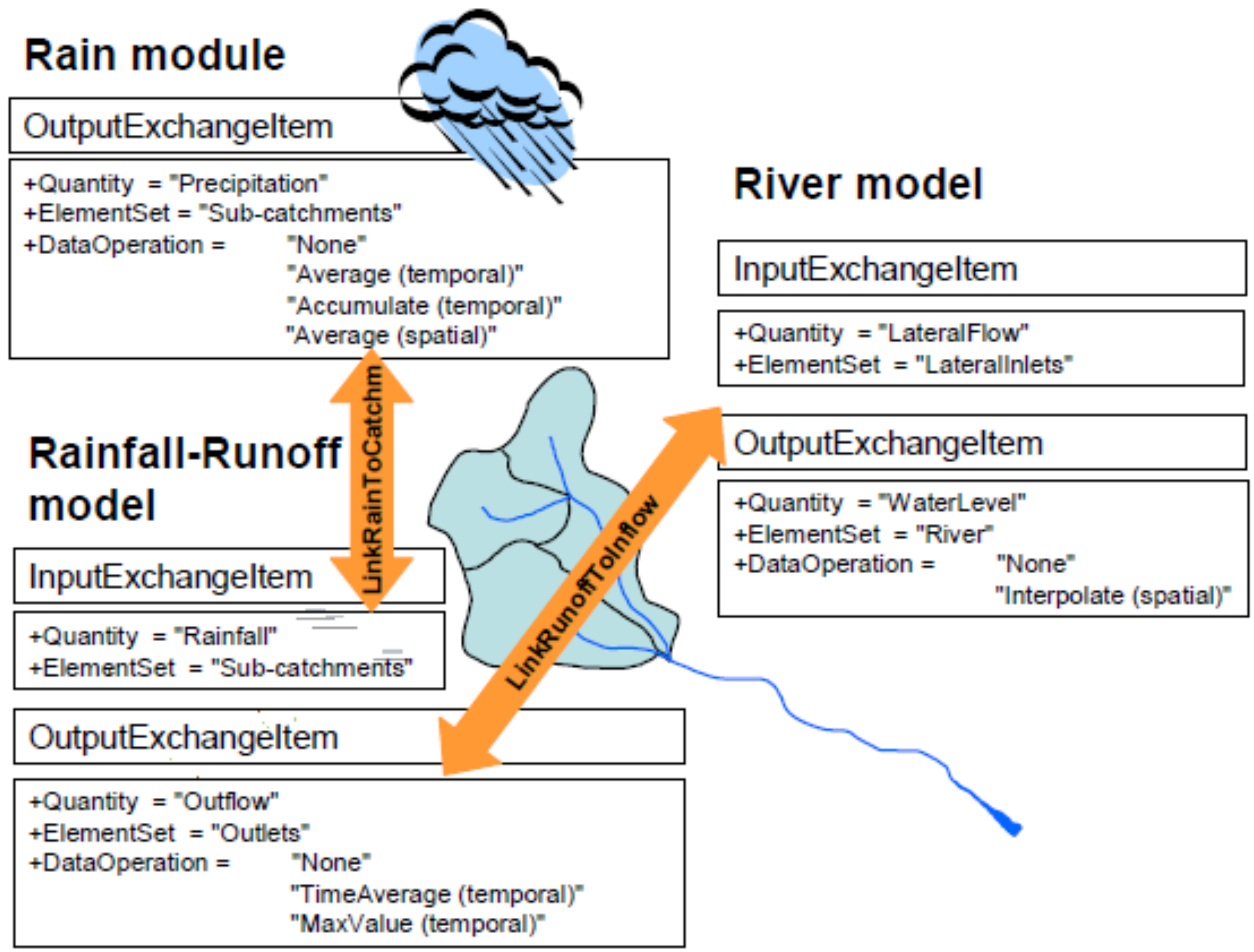
Source: Moore and Tindall, 2005

UML View of Quantities in OpenMI



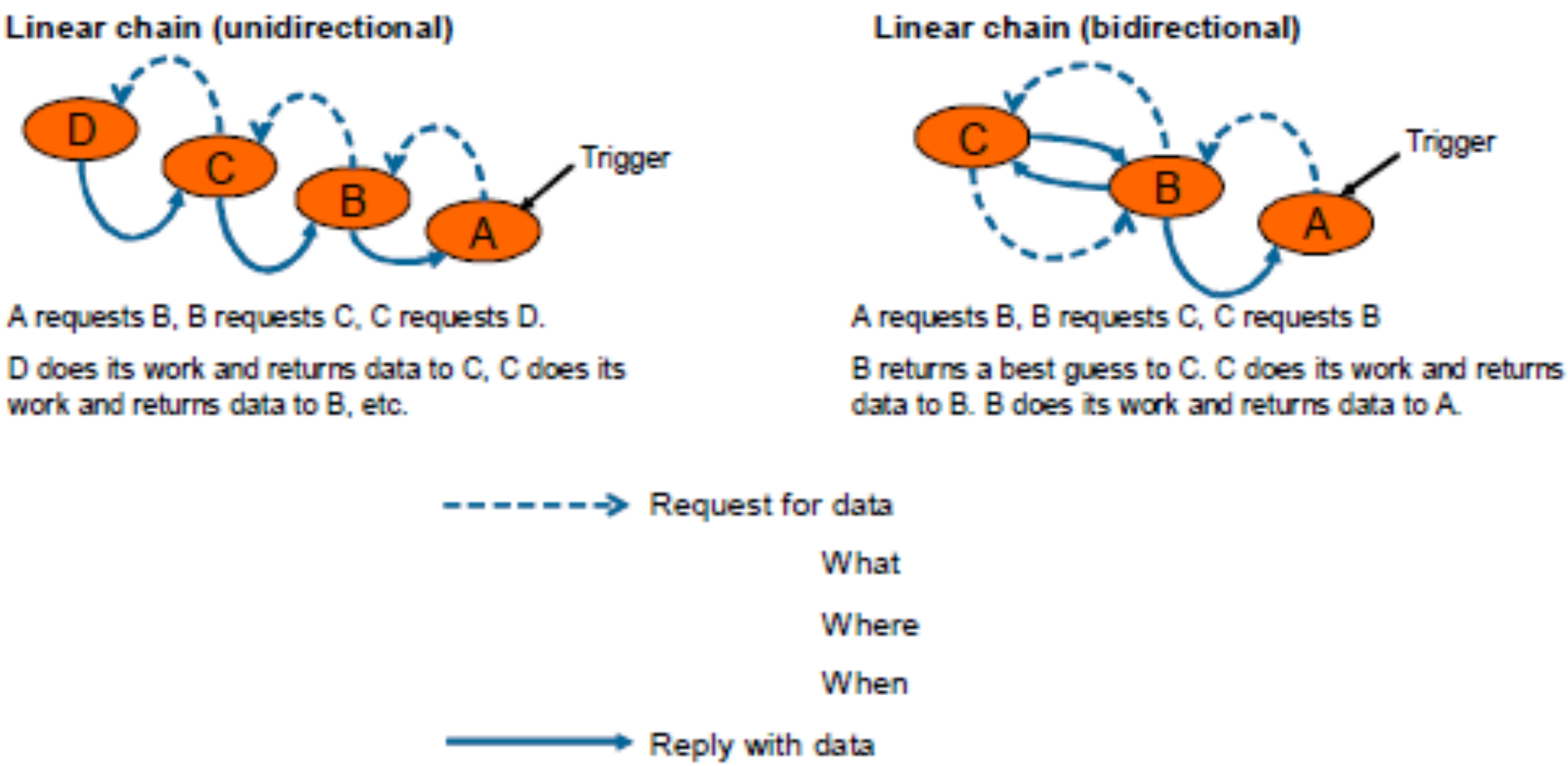
Source: Moore and Tindall, 2005

Links are used to define data exchanges between model components within a specific model configuration



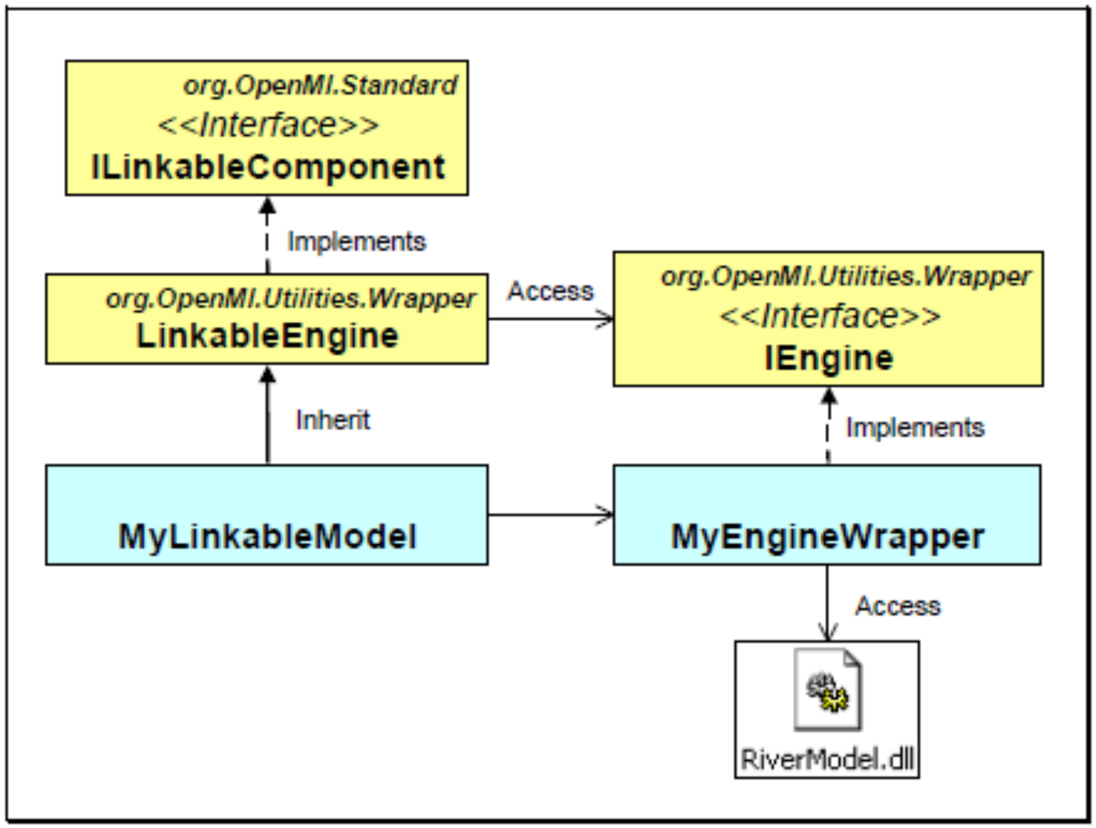
Source: Moore and Tindall, 2005

“request and reply” mechanism for data exchanges



Source: Moore and Tindall, 2005

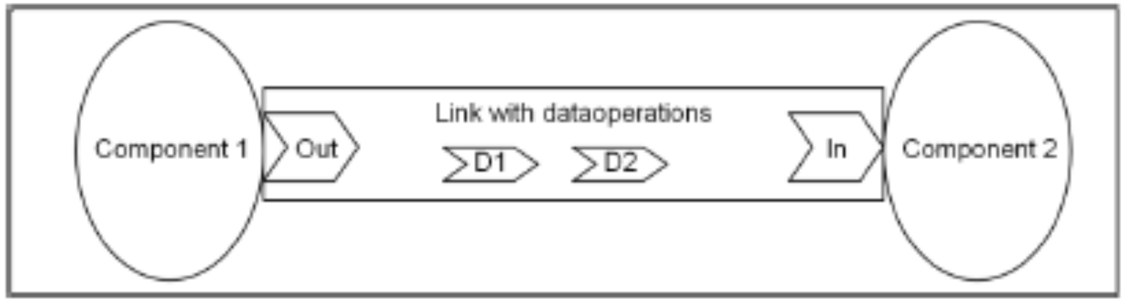
Components implement the ILinkableComponent Interface



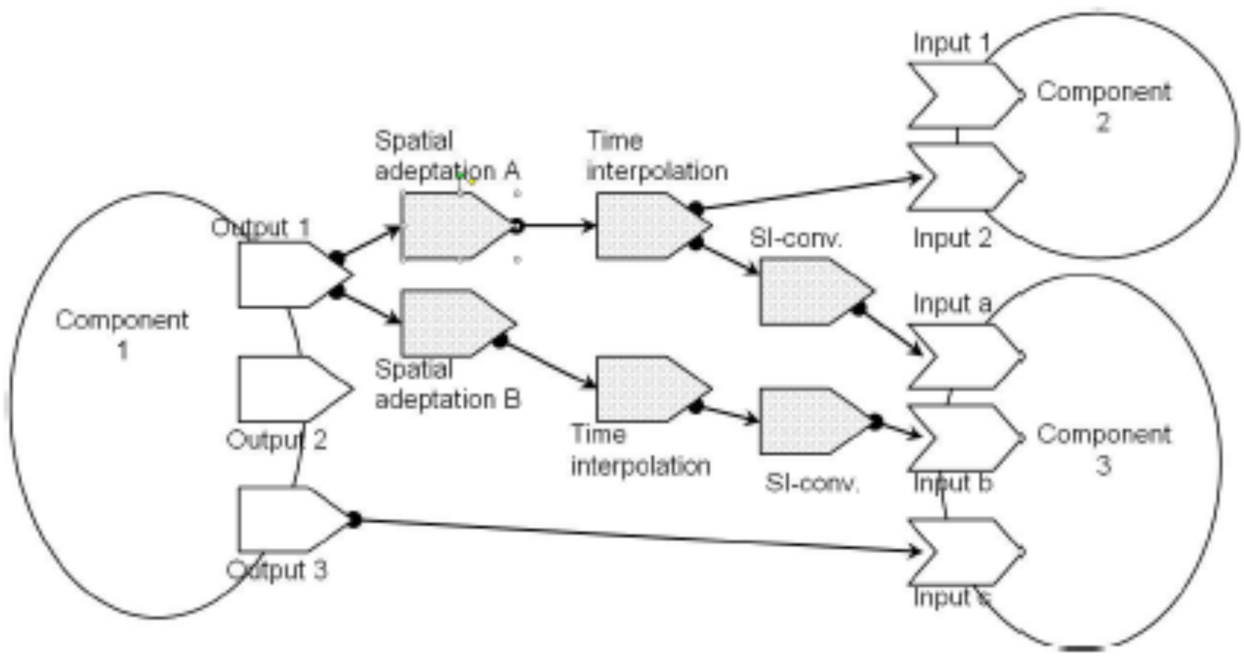
Source: Moore and Tindall, 2005

OpenMI 2.0 changes how models are linked

OpenMI version 1.4 uses data operations on links as discussed before



OpenMI version 2.0 uses the concept of data adaptors



Source: document #11-014r3 available here: http://www.opengeospatial.org/standards/openmi

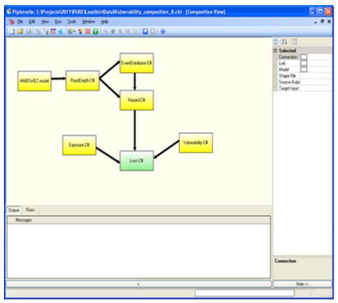
Basic Modeling Interface (BMI)

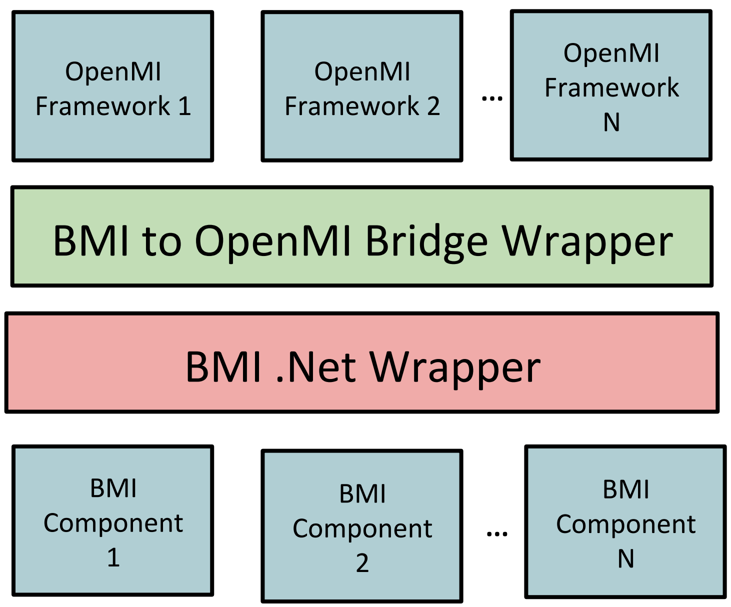
- BMI is defined using the Scientific Interface Description Lanagauge (SIDL) here: https://github.com/csdms/bmi/blob/master/bmi.sidl
- BMI is meant to be a light-weight (basic) interface for exposing models as components
- Frameworks would wrap BMI components using framework-specific model interface standards

<interface>
bmi

+ initialize()
+ update()
+ finalize()
+ get_start_time()
+ get_current_time()
+ get_end_time()
+ get_component_name()
+ get_input_var_names()
+ get_output_var_names()
+ get_value()
+ set_value()
+ get_var_units()
+ get_var_type()
+ get_var_grid()
+ get_grid_spacing()
+ get_grid_origin()
+ get_grid_shape()

Proposed BMI to OpenMI Architecture



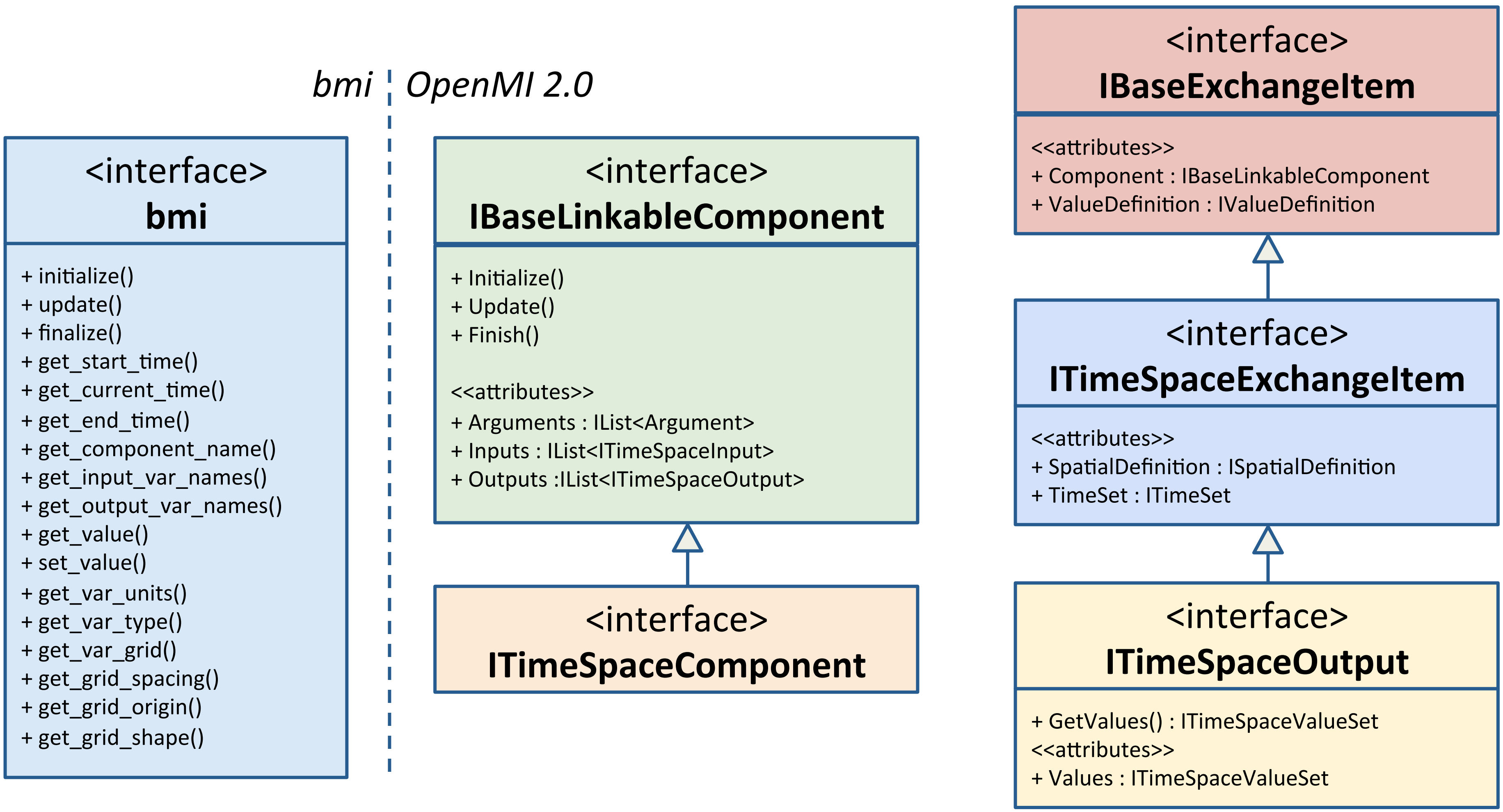


BMI components can be used in OpenMI v2.0 compliant frameworks such as FluidEarth

Two layers of wrappers for (1) C# .NET interoperability and (2) BMI to OpenMI bridging

BMI-compliant model components written in Python, C, or Fortran

Key Interfaces in BMI and OpenMI

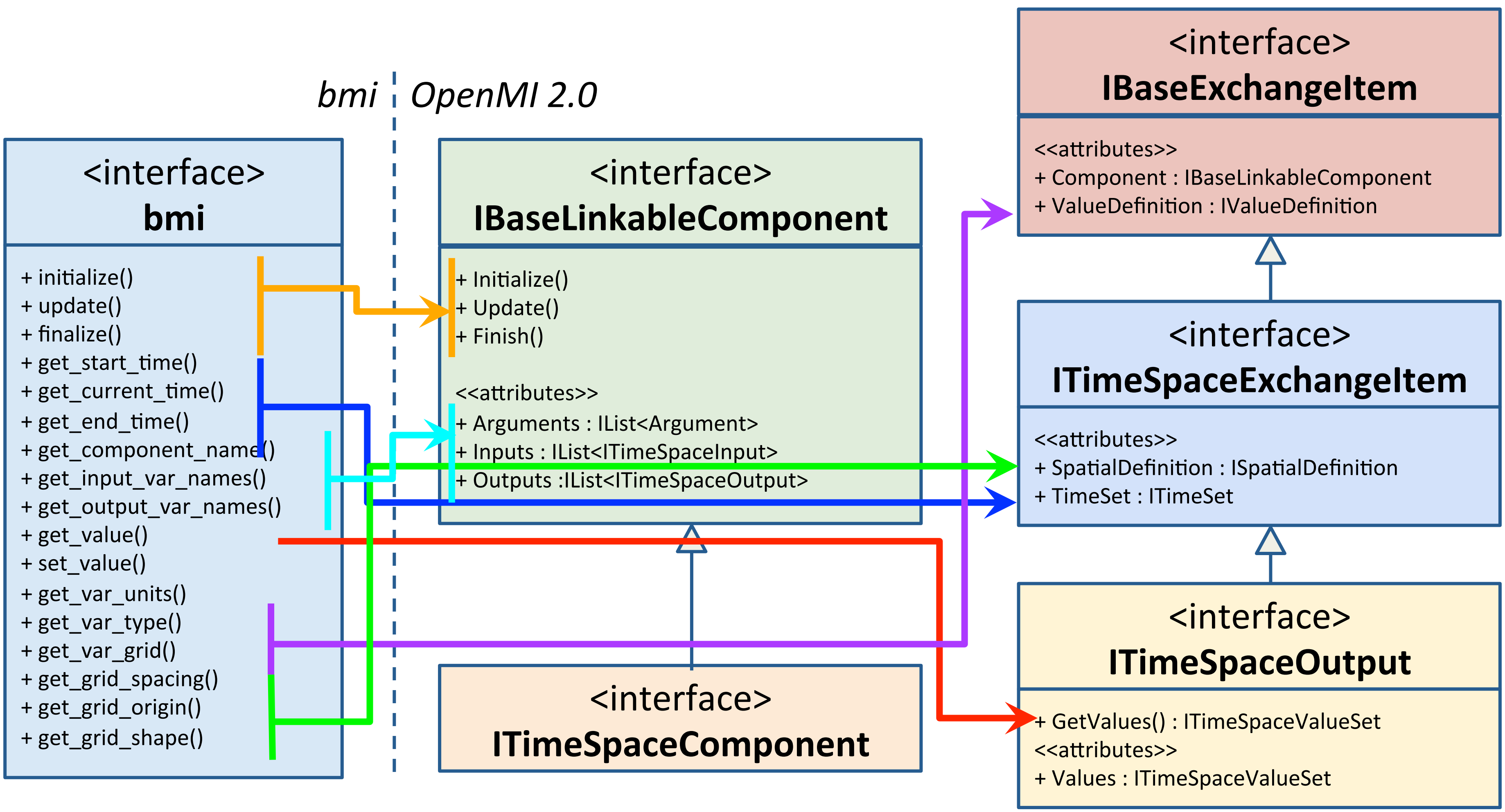


Mapping between BMI and OpenMI 2.0

This will be presented for 5 categories using the following color coding convention:

1. Initialize, Run, Finalize (IRF) methods
2. General metadata properties
3. Getters and setters
4. Time information
5. Spatial representations

Overview of Mapping From BMI to OpenMI

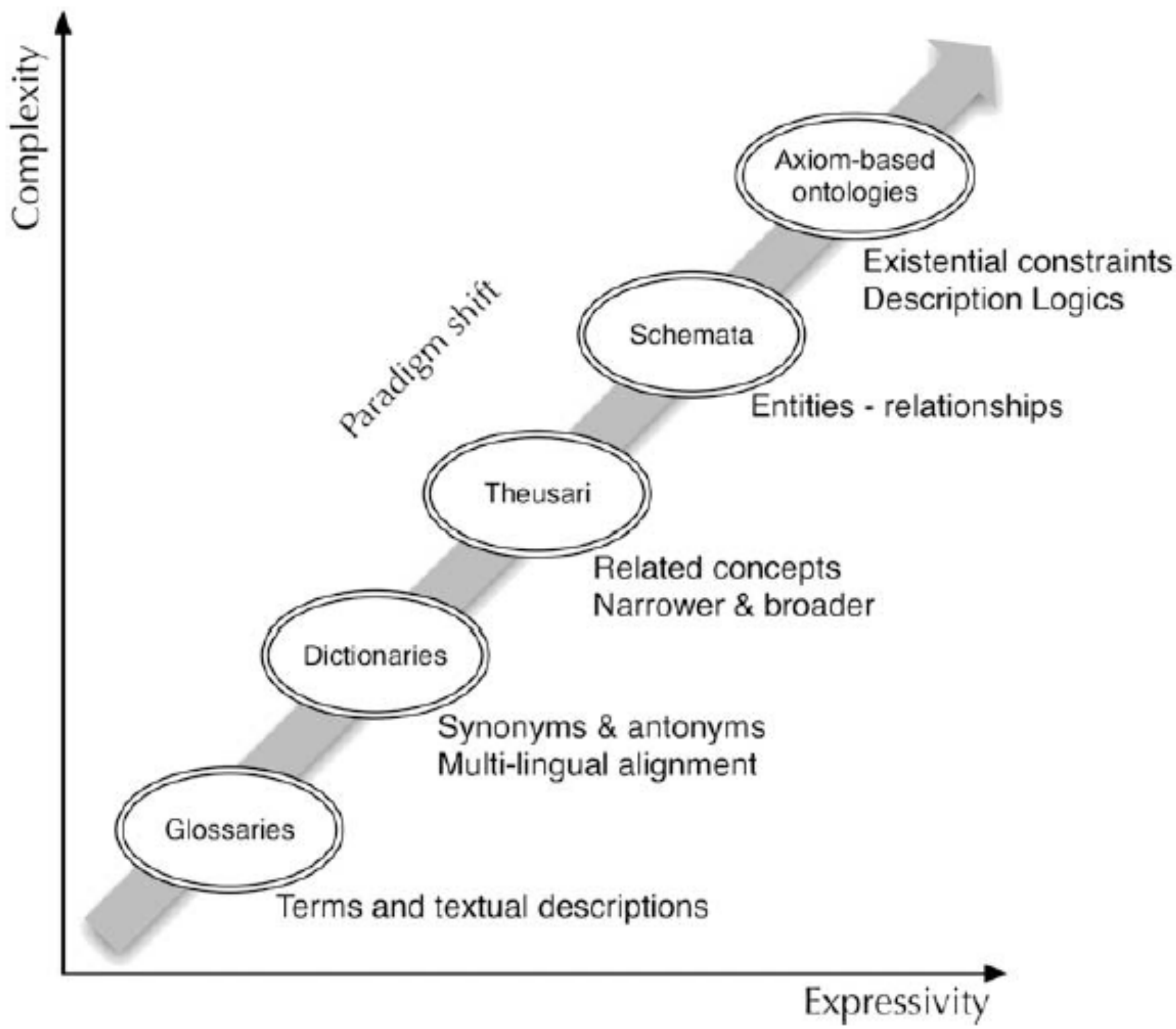


Due to space restrictions, only selected information shown.

Outline

- Introduction
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- Example modeling frameworks
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- Interface Standards
  - OpenMI
  - BMI
- **Current research topics**
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- Tutorials
  - Getting Started with WMT
- Summary

Semantics: From Glossaries to Ontologies



Villa et al., 2009

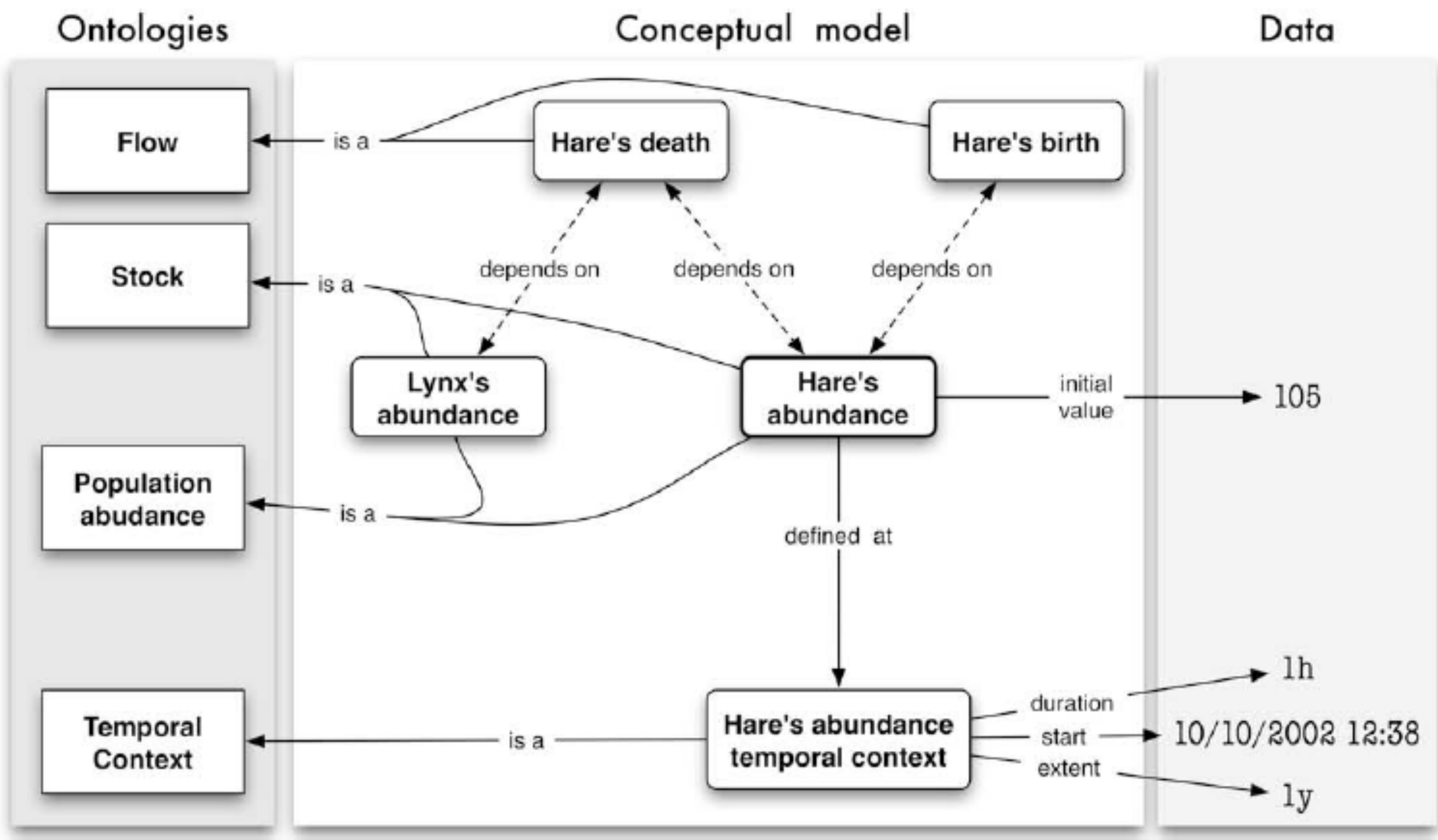
Example of Semantic Annotation for a Model Input

Text box 1. Semantic annotation of a required model input.

```
I :: =
    is-a: Temperature,
    vertically-distributed-in: PlanetarySurface,
    has_unit: Fahrenheit,
    max-value: (is-a: Temperature, has-value: 30.0, has-
    unit: Celsius)
    min-value: (is-a: Temperature, has-value: 19.0, has-
    unit: Celsius)
    distributed-in: (is-a: TimeSpan, step: 1, has-unit:
    Month).
```

Villa et al., 2009

Sematic Relationships within Models



Villa et al., 2009

Semantic Matching for Model Variables

standard variable name = object name + quantity name.

Hydro Model A

Output variables:

- streamflow
- rainrate

Hydro Model B

Input variables:

- discharge
- precip_rate

CSDMS Standard Names

- channel_exit_water_x-section__volume_flow_rate
- atmosphere_water__rainfall_volume_flux

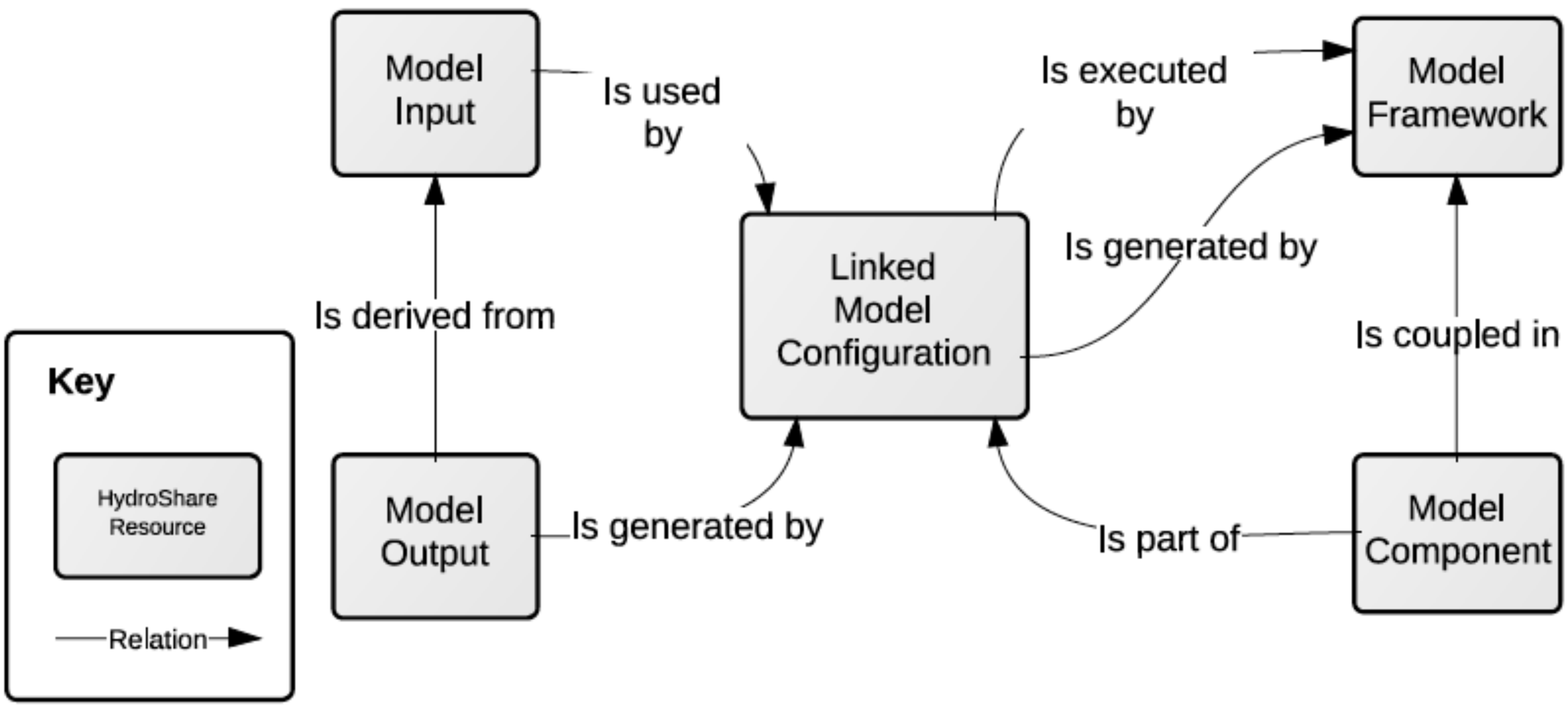
Goal: Remove ambiguity so that the framework can automatically match outputs to inputs.

Peckham, 2014

HydroShare

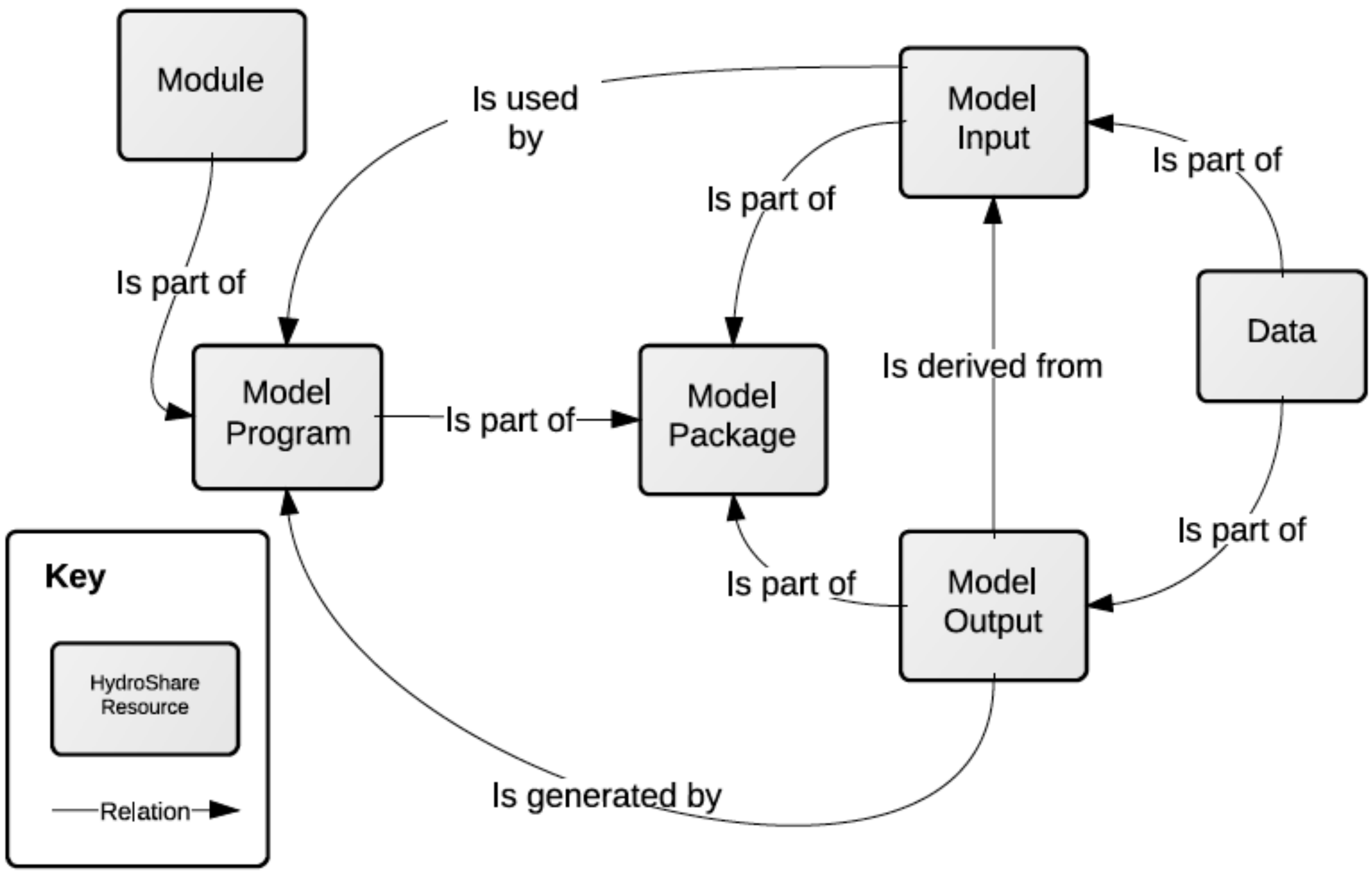
- NSF supported collaborative project building from the CUAHSI Hydrologic Information System (HIS) project
- Resource-based sharing of hydrologic data, models, analysis tools, etc.
- Share resources with metadata with collaborators or public
- Discover resources use hydrology-specific metadata
- Social objects (comments, rating, etc.)
- http://beta.hydroshare.org is live now and http://www.hydroshare.org is coming soon (the plan is for this summer)

Conceptualization of Key Model Concepts and Relationships: Component-based Modeling



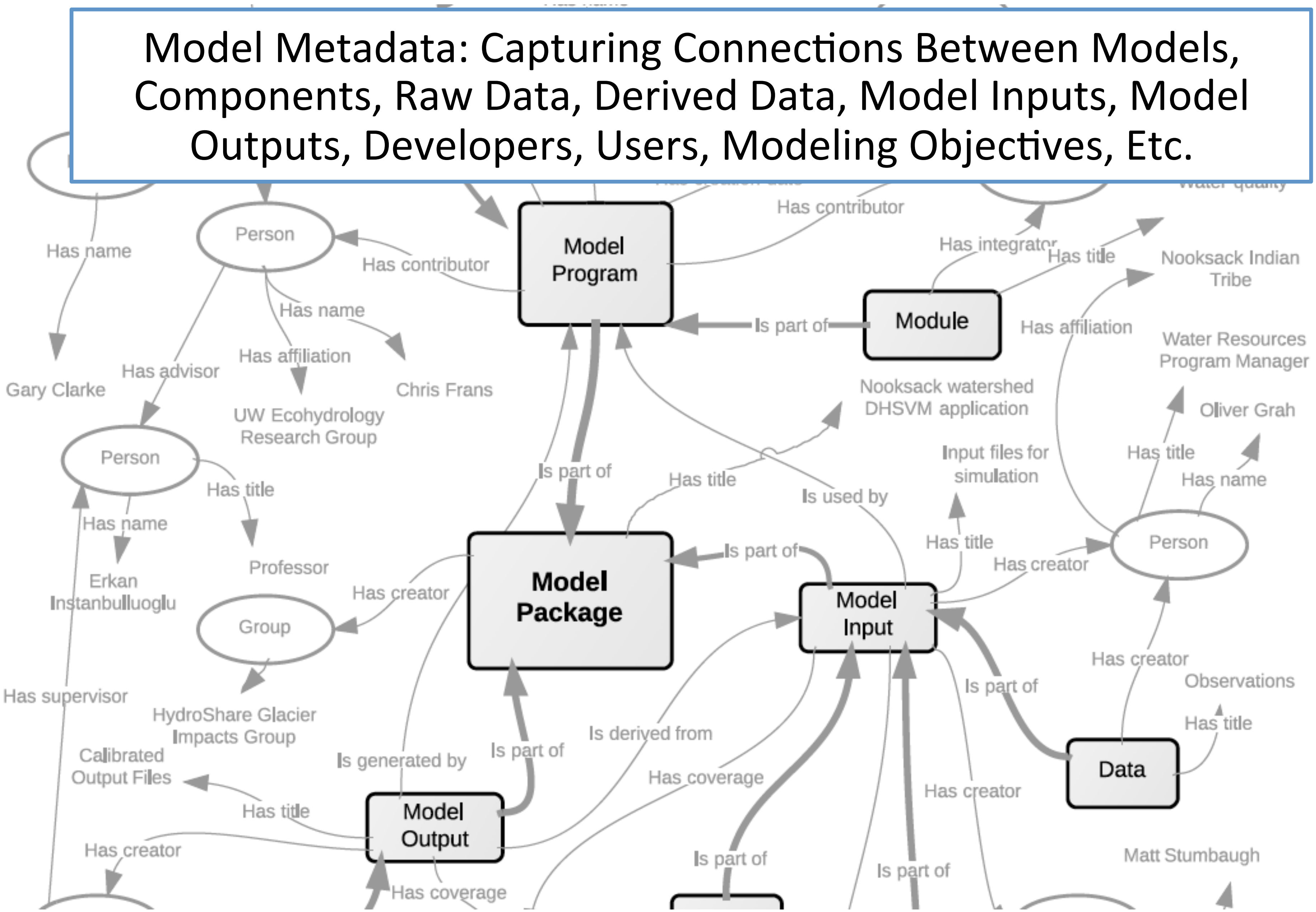
Morsy et al., 2014

Conceptualization of Key Model Concepts and Relationships: Stand-alone Models



Morsy et al., 2014

Model Metadata: Capturing Connections Between Models, Components, Raw Data, Derived Data, Model Inputs, Model Outputs, Developers, Users, Modeling Objectives, Etc.

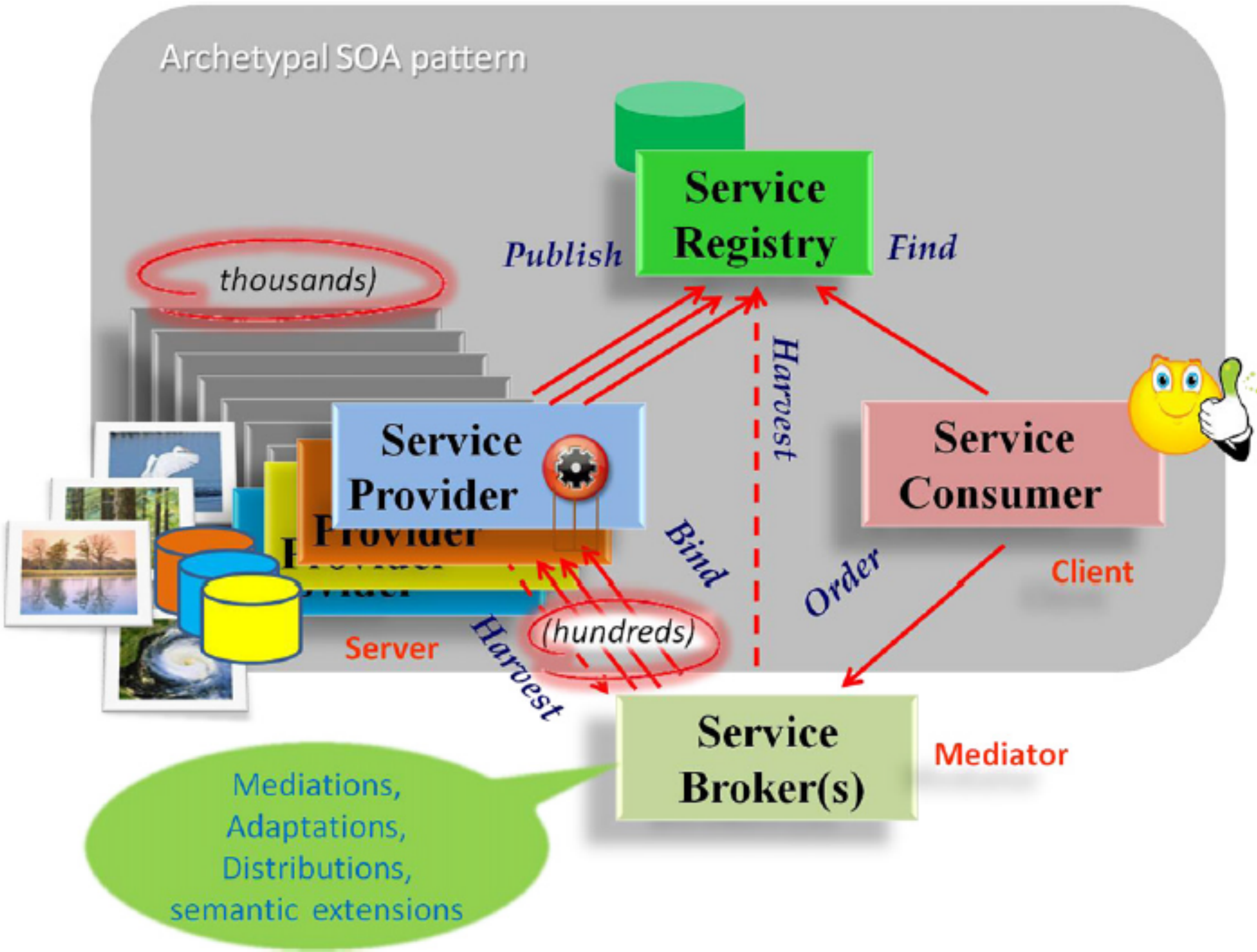


Morsy et al., 2014

Model Web Vision

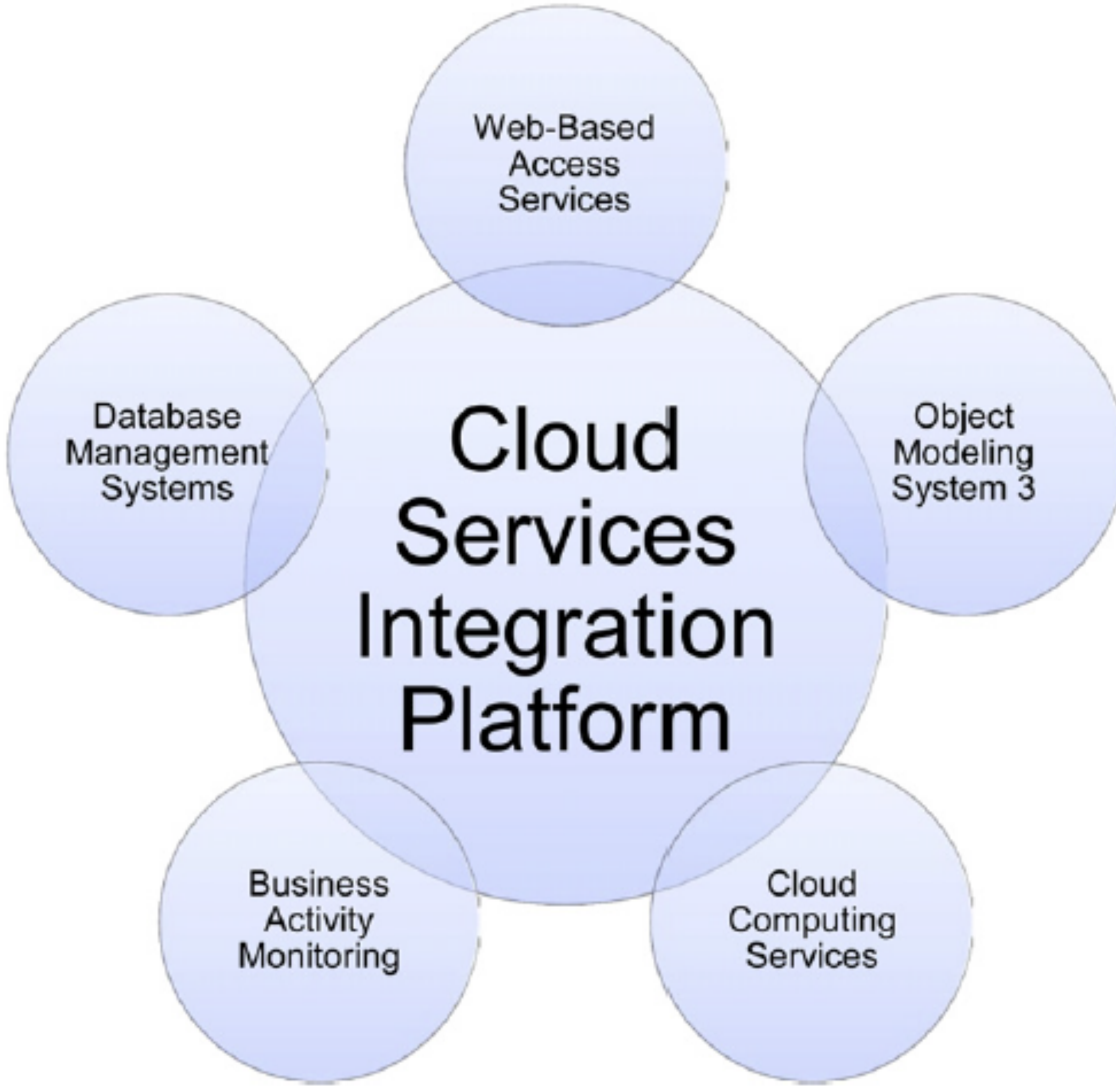
A dynamic web of models, integrated with databases and websites, to form a **consultative infrastructure** *where researchers, managers, policy makers, and the general public can go to gain insight into "what if" questions.*

Nativi et al., 2013



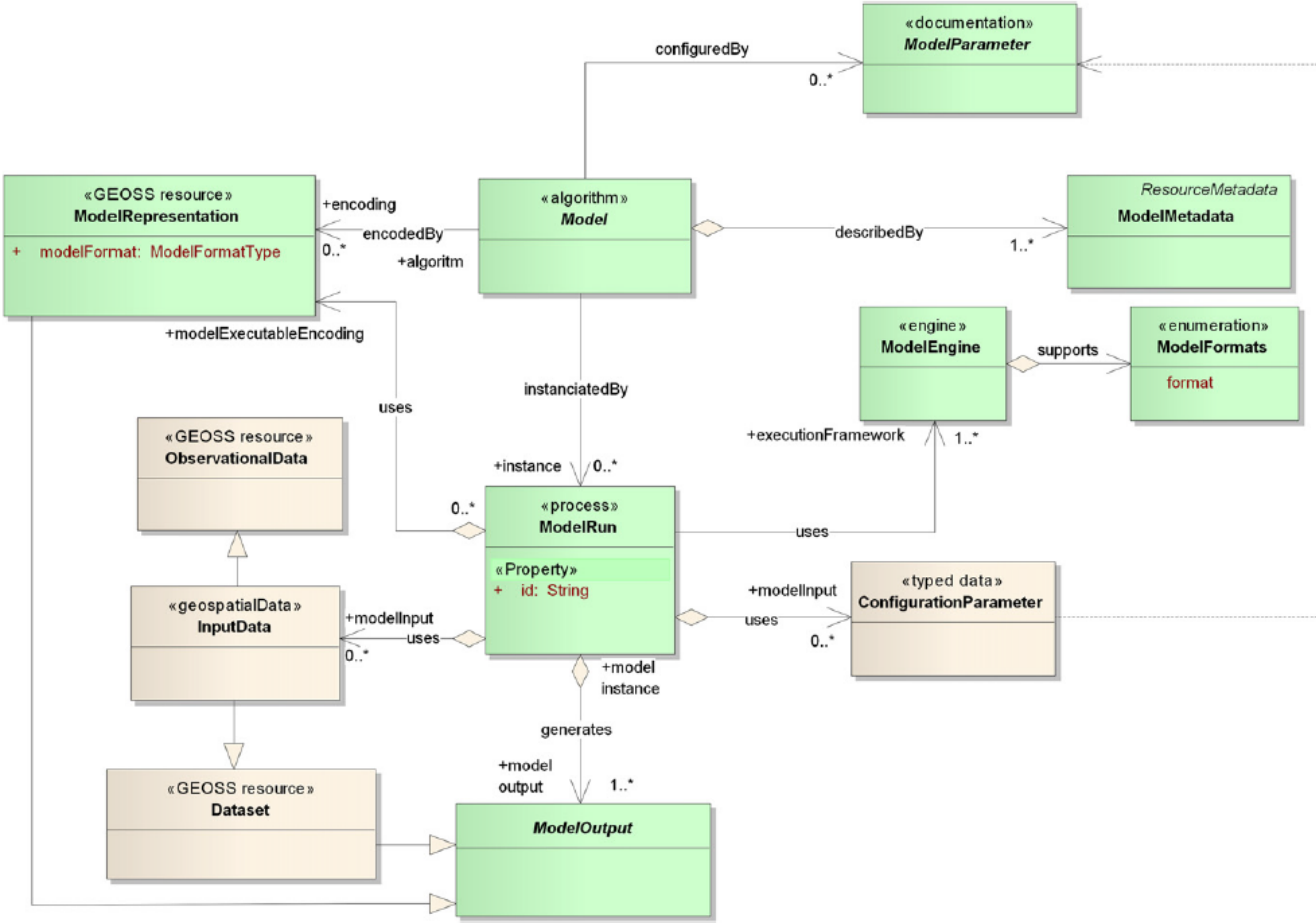
Nativi et al., 2013

OMS 3 within the Cloud Services Innovation Platform architecture



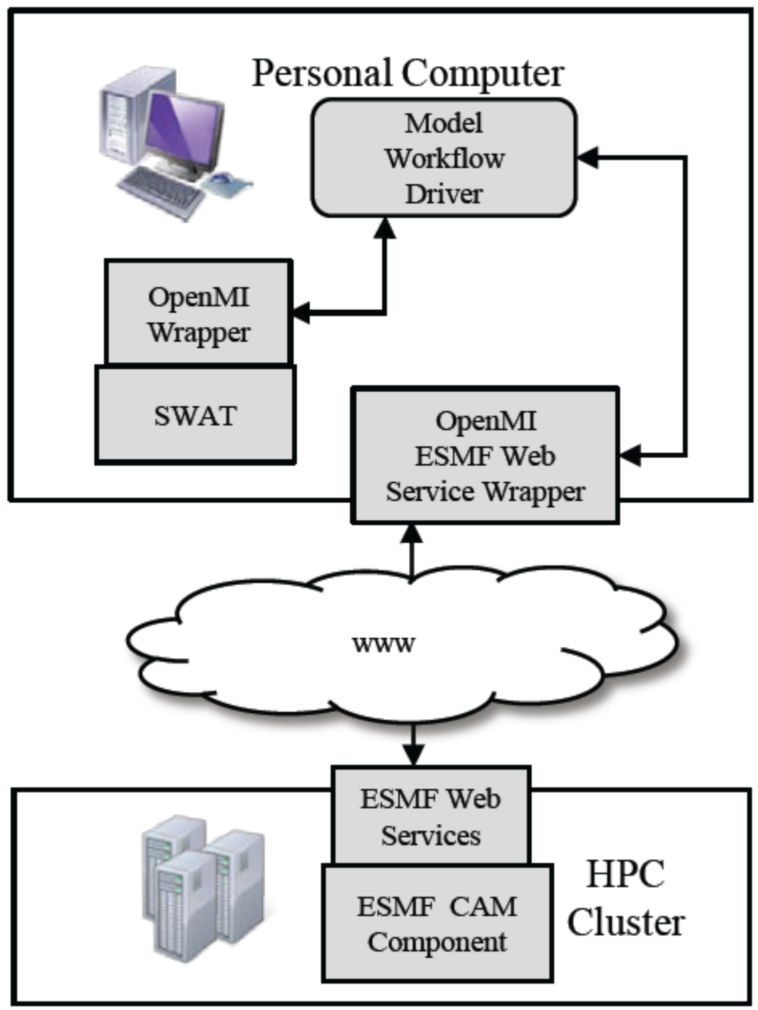
David et al., 2013

The Model Resource Data Model proposed by Nativi et al. (2013)



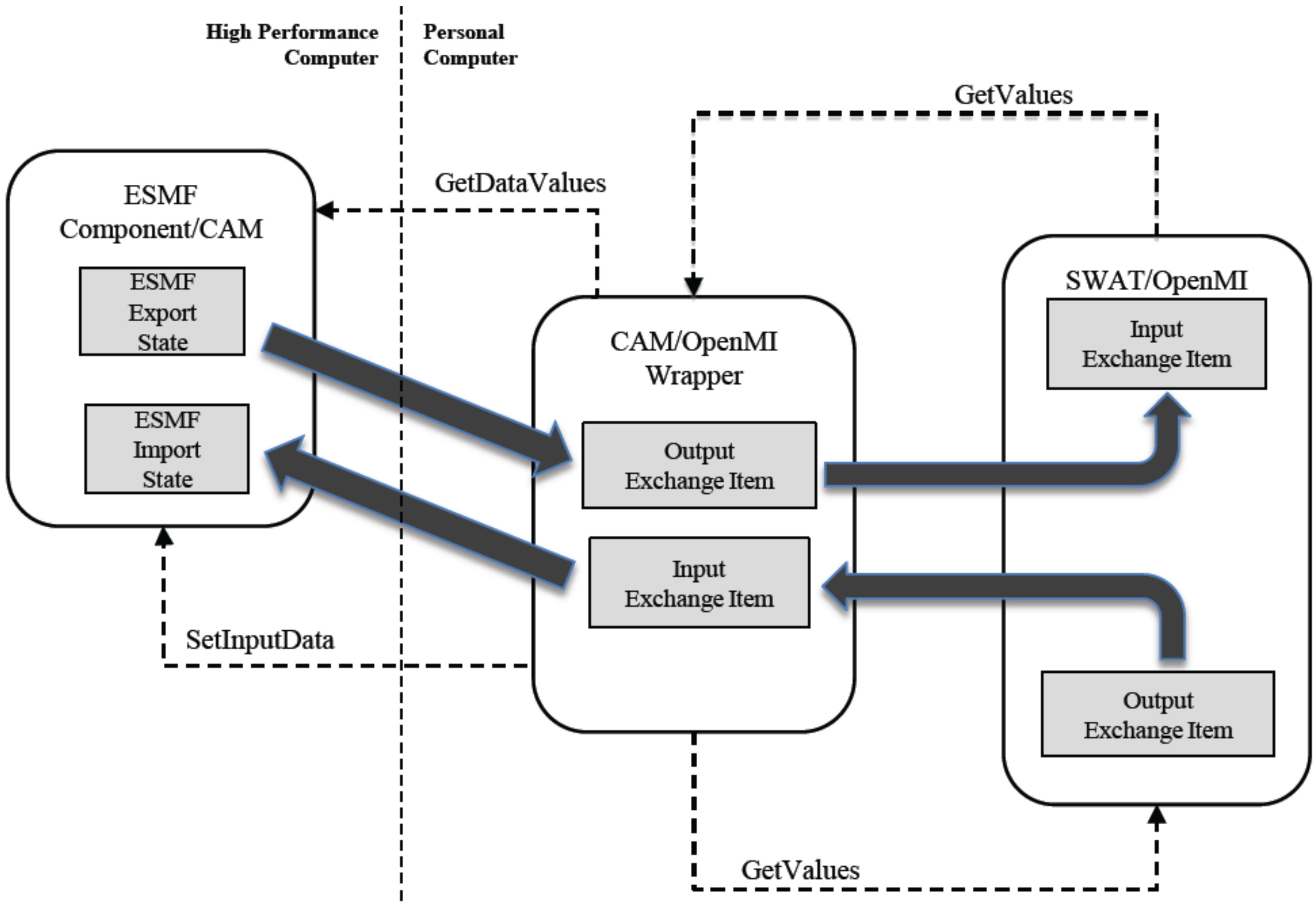
Nativi et al., 2013

Example of using model web concept for coupling earth system models



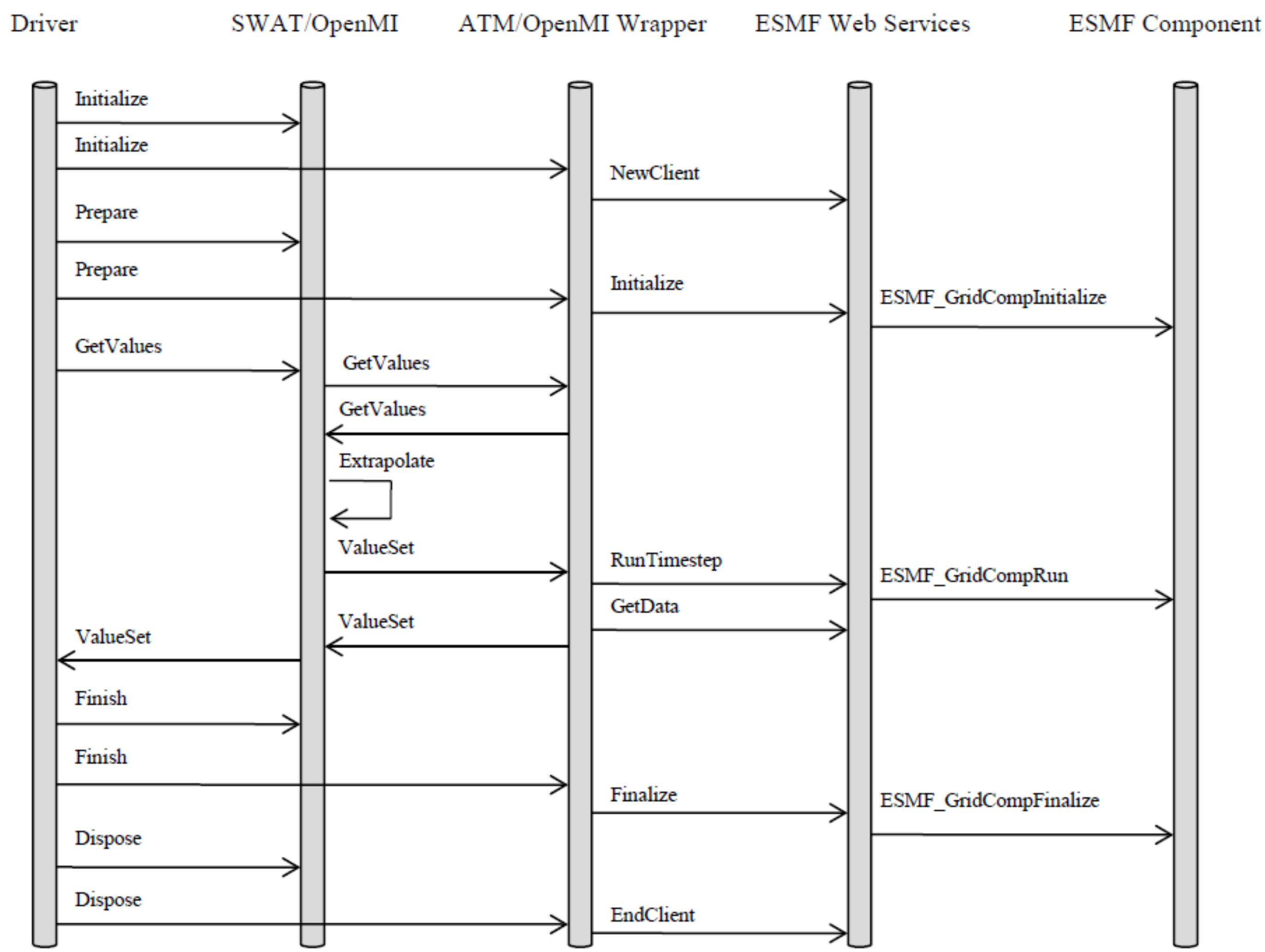
Goodall et al., 2013

OpenMI/ESMF Interoperability via Web Services



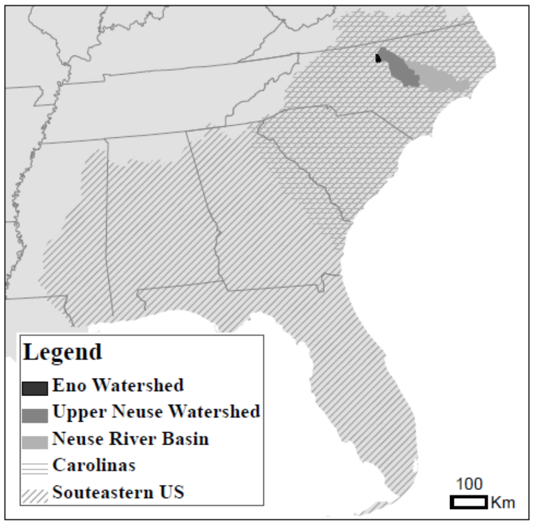
Goodall et al., 2013

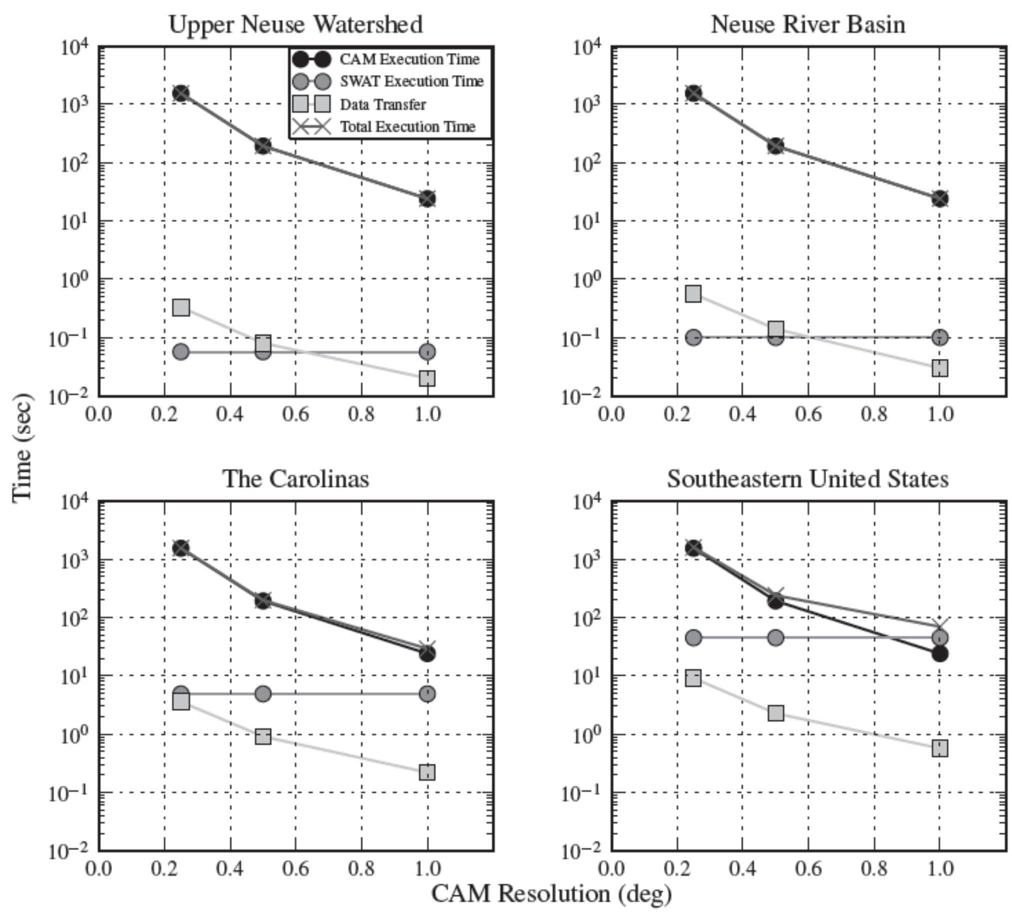
Data flow for OpenMI/ESMF Interoperability



Goodall et al., 2013

Scaling analysis: Data transmission time vs. model execution time



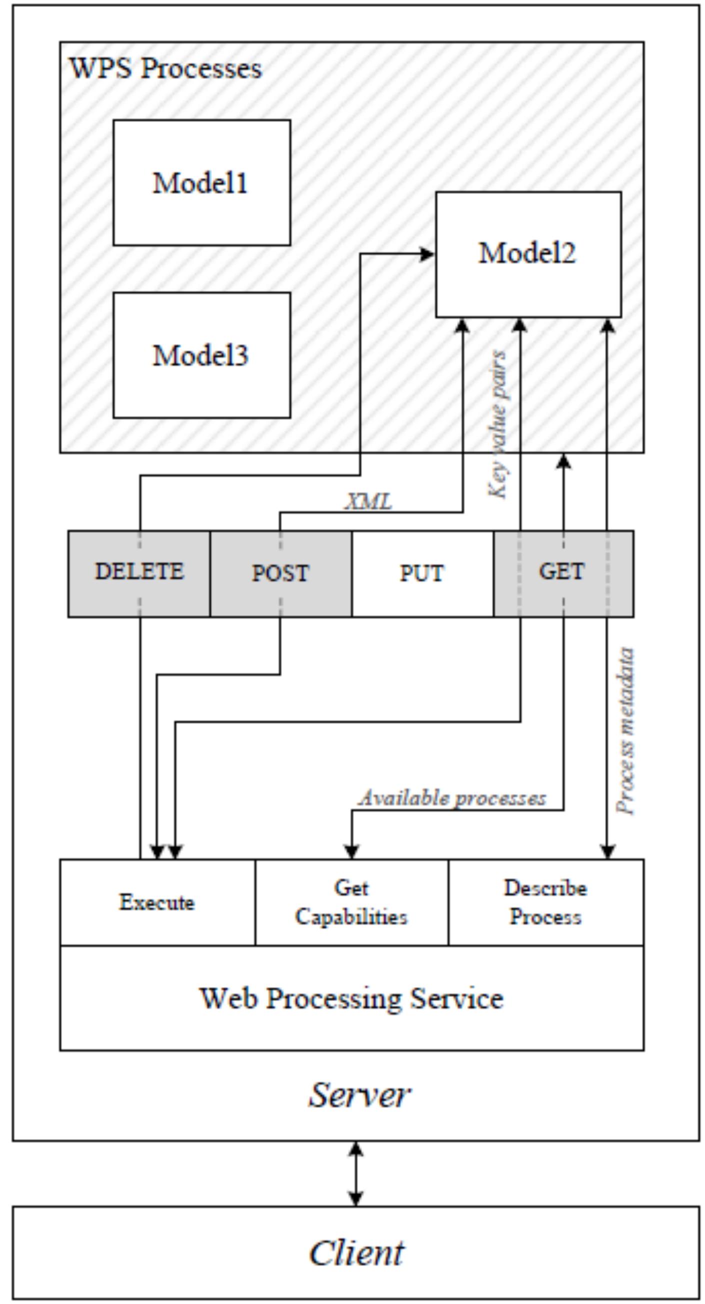


Goodall et al., 2013

Modeling Web using REST/ OGC Web Processing Service

- Models exposed as services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard.
- *Get Capabilities* gives metadata about models on server
- *Describe Process* gives metadata about a specific model service
- *Execute* is more complex that a typical WPS to support model time stepping (e.g., session state with users must be maintained)

Castonova et al., 2013



Outline

- Introduction
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- Example modeling frameworks
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- Interface Standards
  - OpenMI
  - BMI
- Current topics
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- **Tutorials (time permitting)**
  - Getting Started with WMT
- Summary

WMT Tutorial

- Time permitting, I will briefly show how integrated modeling concepts are used by the CSDMS WMT.

- I will use the integrated model described in the coastal evolution lab available here: https://csdms.colorado.edu/wiki/Labs_WMT_CEM

Outline

- Introduction
  - Motivation for integrated modeling/model coupling
  - Terminology and key concepts
- Example modeling frameworks
  - CSDMS
  - FRAMES
  - ESMF
  - OMS
  - FluidEarth
- Interface Standards
  - OpenMI
  - BMI
- Current topics
  - Semantics, ontologies, metadata controlled vocabularies
  - Model web
- Tutorials (time permitting)
  - Getting Started with WMT
- **Summary**

Take home points

- *What is integrated modeling?* IM seeks to provide methods and tools that allow for the combination of data, models, assessment methods from multiple disciplines into a single simulation system.

- *How is this typically accomplished now?* By defining modular components (or services) that encapsulate disciplinary knowledge and then coupling components using a modeling framework to simulate a system.

- *What are some key challenges that require attention by researchers?* Creating general interface standards and semantic mediation across disciplines are two key challenges.

- *Where is this all heading?* To a model web: Both web-based configuration of HPC models and Models as Services. Why? To overcome hardware and software dependencies; to advance reproducible, reusable, and extensible of models; to take advance of cloud infrastructures.

References (1 of 2)

- Argent, R.M., 2004. An overview of model integration for environmental applications - Components, frameworks and semantics. Environ. Model. Softw. 19, 219–234. doi:10.1016/S1364-8152(03)00150-6
- Boettiger, C., 2015. An introduction to Docker for reproducible research. ACM SIGOPS Oper. Syst. Rev. 49, 71–79. doi:10.1145/2723872.2723882
- Castronova, A.M., Goodall, J.L., Ercan, M.B., 2013. Integrated modeling within a Hydrologic Information System: An OpenMI based approach. Environ. Model. Softw. 39, 263–273. doi:10.1016/j.envsoft.2012.02.011
- David, O., Ascough, J.C., Lloyd, W., Green, T.R., Rojas, K.W., Leavesley, G.H., Ahuja, L.R., 2013. A software engineering perspective on environmental modeling framework design: The Object Modeling System. Environ. Model. Softw. 39, 201–213. doi:10.1016/j.envsoft.2012.03.006
- Elag, M.M., Goodall, J.L., Castronova, A.M., 2011. Feedback loops and temporal misalignment in component-based hydrologic modeling. Water Resour. Res. 47, n/a–n/a. doi:10.1029/2011WR010792
- Goodall, J.L., Saint, K.D., Ercan, M.B., Briley, L.J., Murphy, S., You, H., DeLuca, C., Rood, R.B., 2013. Coupling climate and hydrological models: Interoperability through Web Services. Environ. Model. Softw. 46, 250–259. doi:10.1016/j.envsoft.2013.03.019
- Gregersen, J.B., Gijsbers, P.J.A., Westen, S.J.P., 2007. OpenMI: Open modelling interface. J. Hydroinformatics 9, 175. doi:10.2166/hydro.2007.023
- Hill, C., DeLuca, C., Suarez, M., Da Silva, A., 2004. The architecture of the earth system modeling framework. Comput. Sci. Eng. 6, 18–28. doi:10.1109/MCISE.2004.1255817
- Laniak, G.F., Olchin, G., Goodall, J.L., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R., Hughes, A., 2013. Integrated environmental modeling: A vision and roadmap for the future. Environ. Model. Softw. 39, 3–23. doi:10.1016/j.envsoft.2012.09.006

References (2 of 2)

- Liu, Y., Gupta, H., Springer, E., Wagener, T., 2008. Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. Environ. Model. Softw. 23, 846–858. doi:10.1016/j.envsoft.2007.10.007
- Moore, R. V., Tindall, C.I., 2005. An overview of the open modelling interface and environment (the OpenMI). Environ. Sci. Policy 8, 279–286. doi:10.1016/j.envsci.2005.03.009
- Morsy, M.M., Goodall, J.L., Bandaragoda, C., Castronova, A.M., Greenberg, J., 2014. Metadata for Describing Water Models, in: International Environmental Modelling and Software Society (iEMSs) 7th International Congress on Environmental Modelling and Software. doi:10.13140/2.1.1314.6561
- Nativi, S., Mazzetti, P., Geller, G.N., 2013. Environmental model access and interoperability: The GEO Model Web initiative. Environ. Model. Softw. 39, 214–228. doi:10.1016/j.envsoft.2012.03.007
- Peckham, S.D., 2014. The CSDMS Standard Names: Cross-Domain Naming Conventions for Describing Process Models, Data Sets and Their Associated Variables. Int. Environ. Model. Softw. Soc. (iEMSs), San Diego, California, USA.
- Peckham, S.D., Hutton, E.W.H., Norris, B., 2013. A component-based approach to integrated modeling in the geosciences: The design of CSDMS. Comput. Geosci. 53, 3–12. doi:10.1016/j.cageo.2012.04.002
- Villa, F., Athanasiadis, I.N., Rizzoli, A.E., 2009. Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. Environ. Model. Softw. 24, 577–587. doi:10.1016/j.envsoft.2008.09.009
- Whelan, G., Kim, K., Pelton, M.A., Castleton, K.J., Laniak, G.F., Wolfe, K., Parmar, R., Babendreier, J., Galvin, M., 2014. Design of a component-based integrated environmental modeling framework. Environ. Model. Softw. 55, 1–24. doi:10.1016/j.envsoft.2014.01.016

Acknowledgements

- Support for the development of this clinic material was provided by the National Science Foundation under award number CBET-0846244.

Questions?

Contact info:
Jon Goodall
University of Virginia
Dept. of Civil and Environmental Engineering
goodall@virginia.edu

download slides: http://goo.gl/VRgzc9