An Innovative Approach to Processing and Converting Environmental Data

27 March 2012

Simulation and Environment Reuse (SER) Forum

William W. Ferrell – wferrell@abamis.com
Warren Macchi – wmacchi@abamis.com
Anthony A. Barresi – abarresi@abamis.com
Farid Mamaghani – farid@sedris.org

Outline

- Simulation Systems and Environmental Data
- The Environmental Data Processing Pipeline
- Challenges in Processing Environmental Data
  - Many formats, differing data models, data handling, custom or extended structures, ...
- An Innovative and Unified Approach
  - Template-based data identification and conversion
  - Reusable components and algorithms

Simulation Systems and Environmental Data

Scenario

Other

Environment

- Multiple Data Representations
- Environmental Object Classification & Attribution
- Mapping Between Different Data Models
- Runtime & Target System Requirements
- Use of External Processing Libraries
- Many File Formats
- Differing Data Models
- Suitability of Data Sets
- Loss of Data/Precision
- Accessing Data Efficiently
- Undocumented Data Structures

Challenges in the Environmental Data Processing Pipeline

- Many file formats (source, interim, exchange, runtime, …)
- Undocumented data and data structures
- Differing data models
- Multiple data representations and organizations
- Suitability of data sets
- Mapping between different data models for processing
- Use of external processing libraries
- Environmental object classification and attribution
- Loss of data and/or precision
- Efficient access to contents of large data sets
- Runtime and target system database requirements

Issues Related to File Formats

- Variety of File Formats
  - Text: ASCII, CSV, XML, COLLADA, ...
  - Binary: Shapefile, OpenFlight, STF, JPG, ...
    - Custom runtime formats
- Considerations
  - Syntax and semantics
  - Storage/memory/processing requirements
  - File format (access, read/write) libraries

File Format Access Libraries: Considerations

- Develop a custom library or use off-the-shelf
- Language binding and implementation
- Licensing terms
- Closed or open source
- Memory and efficiency of data access
- API design and feature set
- Flexibility, robustness
- Documentation, technical support and community

Environmental Data Models: Impacts on Data Processing

- Data model(s) exposed by a library's API
- Mapping between external and internal data models
- Impact on interchange capabilities and knowledge reuse
- Multiple types of data and representations
- Efficiency of data structures
- Spatial reference frames and coordinate transformations
- Environmental object classification and attribution
- Relationships between environmental data entities
- Extensibility

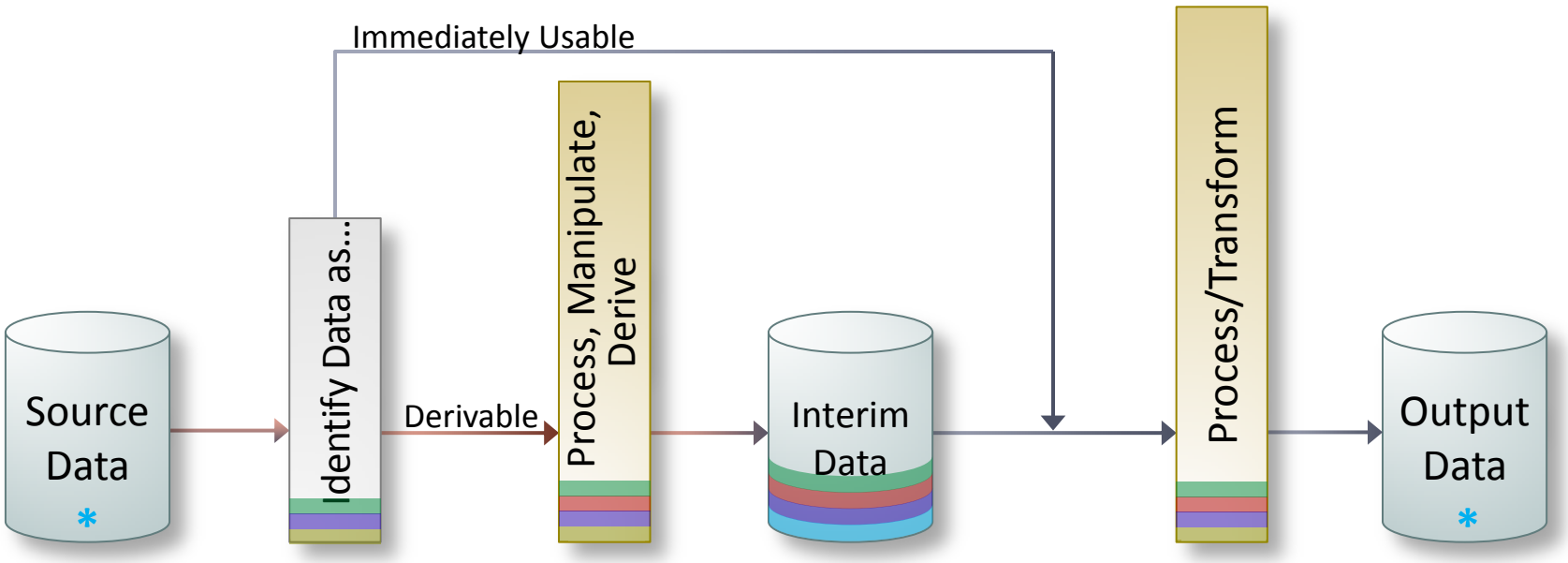
Handling the Data: Manipulation and Reformatting

- Retrieval, refinement, and assembling of environmental data from multiple sources
- Determining suitability of source data sets:
  - Directly usable
  - Not directly usable but usable form is easily derivable
  - Not directly usable but usable form is **not** easily derivable
  - Absent entirely
- Strategies for handling varying representations:
  - Handle all representations by direct implementation
  - Handle a subset and reject all others
  - Handle a subset, pipe other representations to processing components that can transform them into usable ones
- Mechanisms for identifying suitability of data and for selecting algorithms for data transformations are needed

An Innovative and Unified Approach

- Relate internal & external data models
- Define the data semantics: relationships, organizations, hierarchies — Data Representation Model (DRM)
- Use appropriate spatial reference frames — Spatial Reference Model (SRM)
- Identify the environmental object classification and attribution — Environmental Data Coding Specification (EDCS)
- Express input and output requirements — Transmittal Content Requirements Specification (TCRS)
- Map each of the data models, once
- Invoke appropriate processing components for identified data
- Define control parameters to direct the process
- Capture requirements, parameters, etc. in reusable *templates* — TCRS

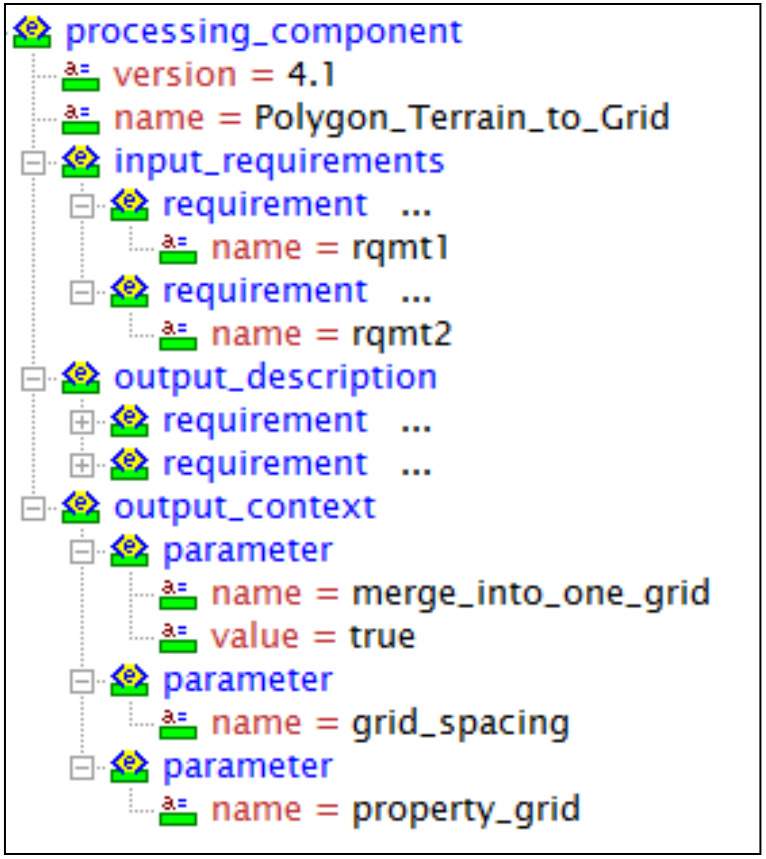
An Innovative and Unified Approach



* Source and Output data may be in the same or different formats, including SEDRIS Transmittal Format (STF).

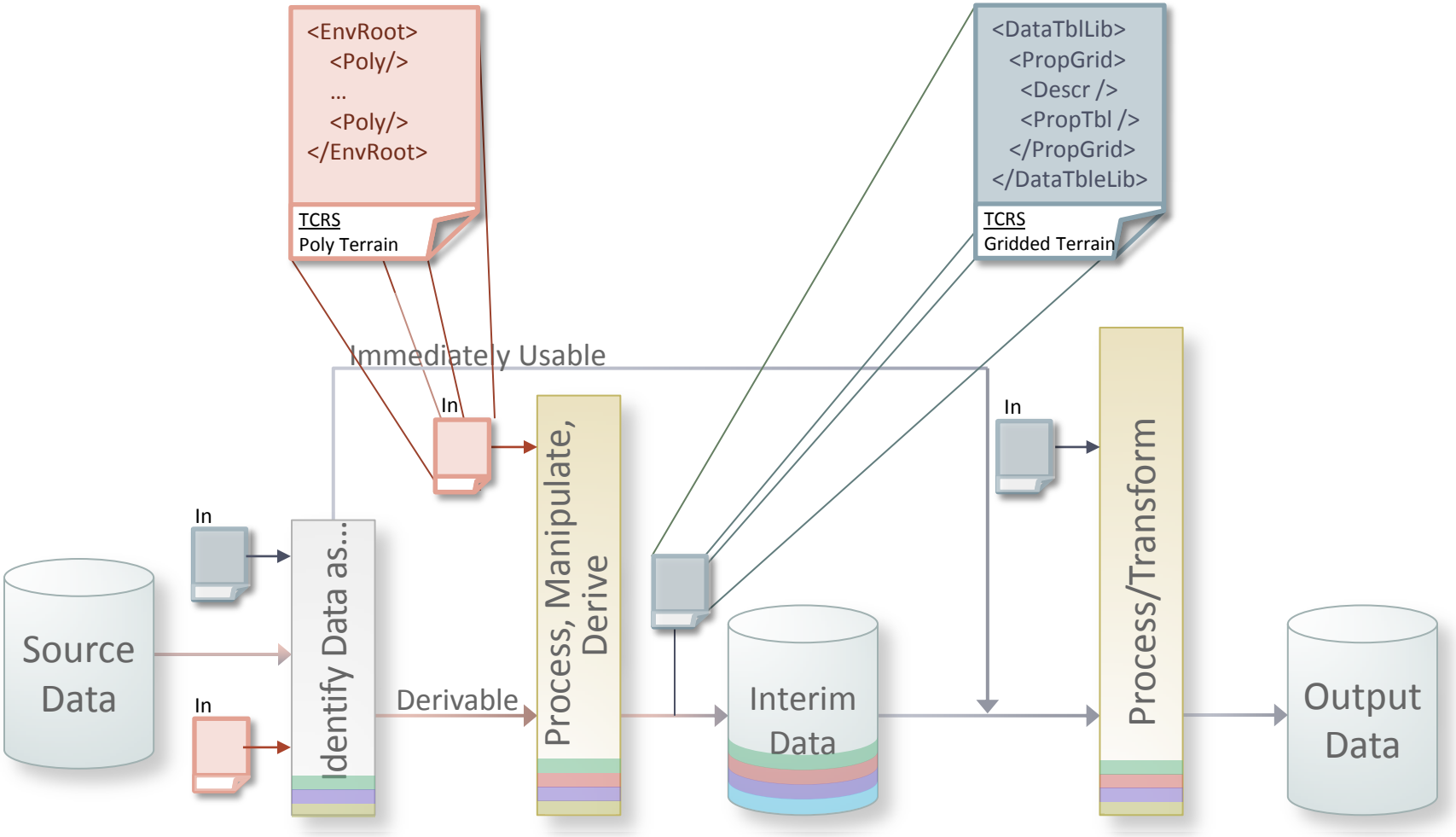
Template Driven Data Identification and Processing

- Using TCRS to identify different data representations
- Using templates to choose suitable processing components for transformation of data to other representations
- Template composition:
  - TCRS to express input requirements
  - TCRS to express output data organization
  - Parameters to control processing



Sample Template

Template Driven Data Identification and Processing (cont.)



Applying the Approach
Example: Updating *STF to DTED Converter*

- **Goal:** Support conversion of polygonal terrain in addition to gridded terrain
- Updating *STF to DTED Converter*
  - Specify templates for both polygonal and gridded terrain representations
  - Develop an external processing component, reusing *STF to CTDB Converter*'s polygon-to-grid algorithms
  - Identify suitable data using the templates
  - Invoke the external processing component to derive gridded terrain data from polygonal terrain data
- Results
  - *STF to DTED Converter* has expanded source data support
  - Support added without major changes to the tool
  - Polygonal terrain conversion algorithm available as an external processing component for other tools to reuse

Benefits of the Standards-Based Template Approach

- Provides knowledge and algorithm encapsulation
- Promotes software reusability
- Can be non-intrusive to existing tools
- Speeds development of new tools
- Encourages better documentation of data representations and requirements
- Aids in building a comprehensive processing algorithm library

Summary

- Challenges in processing environmental data:
  - Mechanics: file formats, libraries, APIs
  - Semantics: data models, representations, data understanding
- Standards-based template-driven processing of environmental data offers many benefits
- Value-adding and R&D opportunities emerge by exploring and applying this approach
- Future work

More Information

- SEDRIS technology components
  - *Available at http://www.sedris.org*
- Tools, utilities and converters
  - *Available at http://tools.sedris.org*
- Data samples
  - *Available at http://data.sedris.org*
- Questions
  - *Contact help@sedris.org*
- ISO/IEC standards activities
  - *Information at http://wg8.sedris.org*
    http://standards.sedris.org