Gplates and GPML: Open software and standards for linking data to geodynamic models

Dietmar Müller

and the GPlates Consortium

School of Geosciences, The Univ. of Sydney

EarthByte

www.earthbyte.org

Quest for discoverability

- Year upon year, computing power has increased
- At the same time, the price-to-performance ratio of computers – particularly those composed of commodity PC parts – has continued to fall.
- Advances in the implementation of general-purpose parallel and distributed computation systems have meant that powerful computational systems are no longer the exclusive domain of governments and large corporations.
- We harness these powerful computational systems to process large amounts of data or to simulate complex real-world phenomena.

Quest for discoverability

- But one topic in particular has received substantially less attention – the storage and management of scientific data.
- What goes into a number-crunching program? What comes out of it? The marvelous processing power of the software is wasted if the results cannot be adequately captured *and discovered*.
- To obtain the maximum benefit from a computational system, the **description of the data** employed by the storage and management system (called the " data model") must be **precise**, accurate, and **sufficient**.

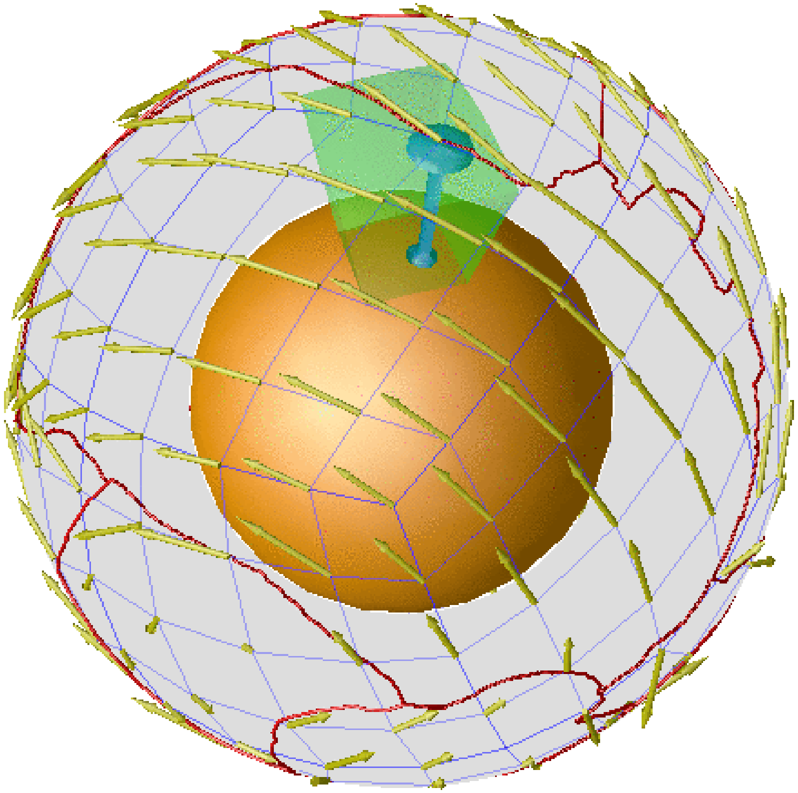
’Everything’ on Earth is controlled by Plate Tectonics

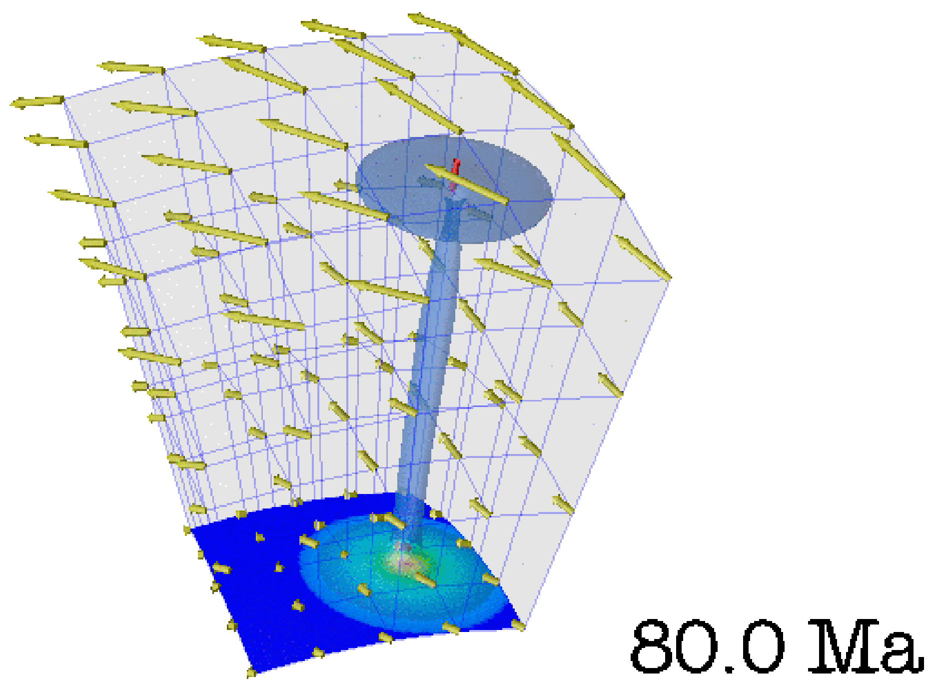
- **Resources** (hydrocarbons, minerals)
- **Geothermal energy** (mostly granites and active volcanism)
- **Tourism** (landscape, beaches, ocean)
- **Climate** past and present (distribution of continents and oceans)
- **Agriculture** (limestone, weathered basalt)
- **Wine & beer** (beer: magnesium limestone, terroir)
- **Civil engineering** (stability of slopes, tunnels, dams, hazards)
- **Evolution of life and biodiversity** (distribution of continents)
- Very important in **planetary research**

QuickTime™ and a
decompressor
are needed to see this picture.


Oil & Gas
EVOLUTION
and the Fossil Record
Wood
Religion/Politics
Climate
CARBON DIOXIDE
HOT SPRING
VOLCANO
S FB 1304
BONDI
BEACH
Building Materials
Colgate
Toothpastes
INTERIOR
Terroir

OpenDX visualization of CitcomS mantle convection simulation





From Mike Gurnis (2005)

Depends on interoperability/workflow between plate tectonic GIS, plate kinematic model, 3D parallel convection code and visualization tool

Quest for discoverability of geological/geophysical data in a plate tectonic context: Just buy off-the-shelf GIS software?

- File formats and information models neither open, portable or useful
- Proprietary
- Expensive

The solution, step 1: build a suitable data model based on open standards

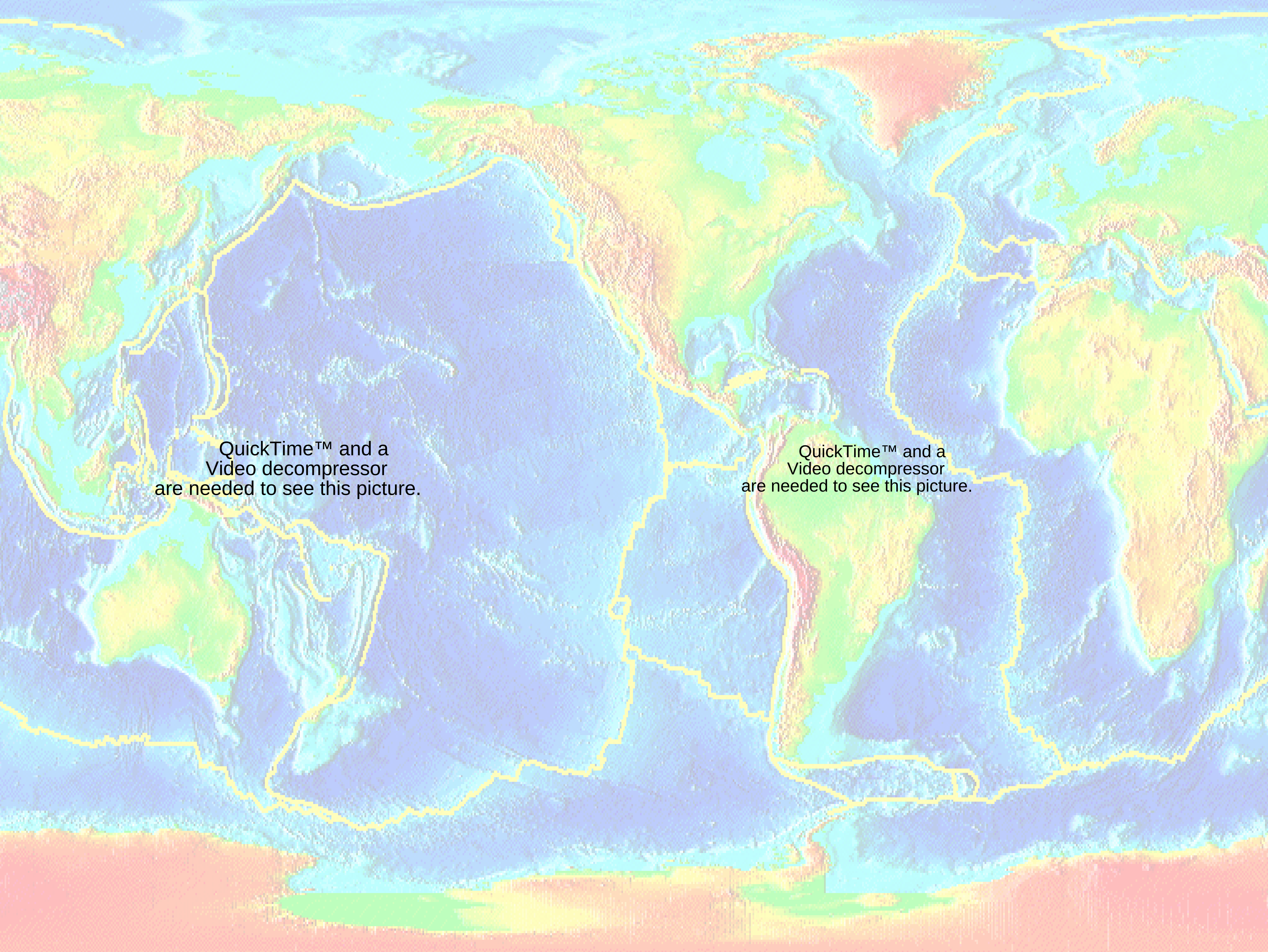
- XML: the eXtensible Markup Language.
- XML is...
  - open and portable:
    - XML is a non-proprietary, plain-text format.
    - many software tools exist for reading and writing XML.
  - extensible:
    - whitespace independent; only formatting is XML markup.
    - XML Schema enables precise, machine-comprehensible definition of what is “valid”.

XML/GML/GPML

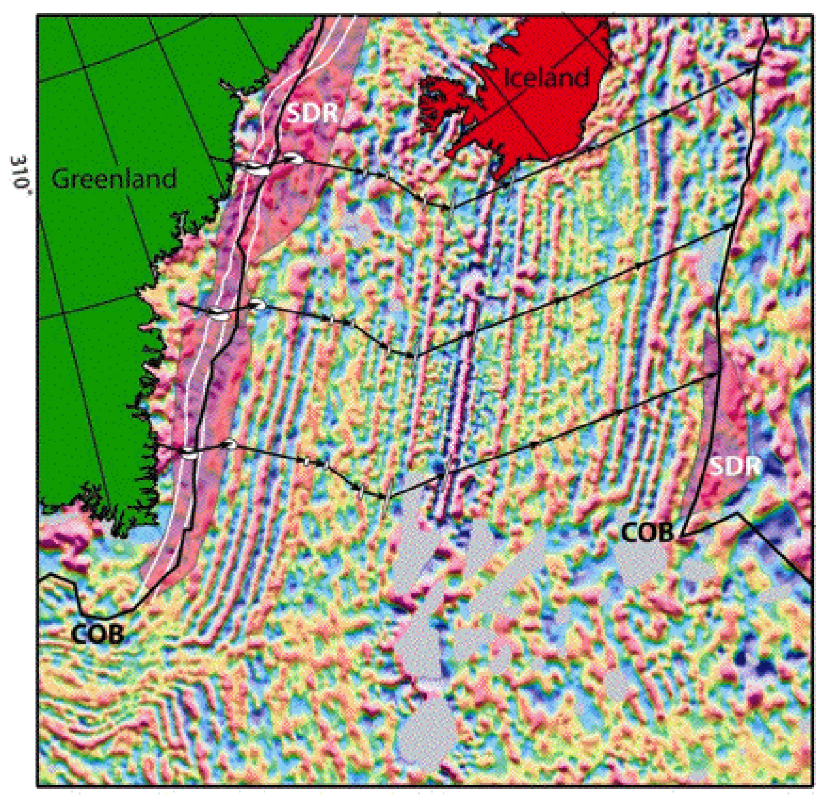
- GML: the Geography Markup Language:
  - an XML-based file format
  - defined by the Open Geospatial Consortium (OGC)
  - on its way to becoming an ISO standard.
- GPML: the GPlates Markup Language:
  - an XML-based “native” file format for geological/geophysical data attached to tectonic plates

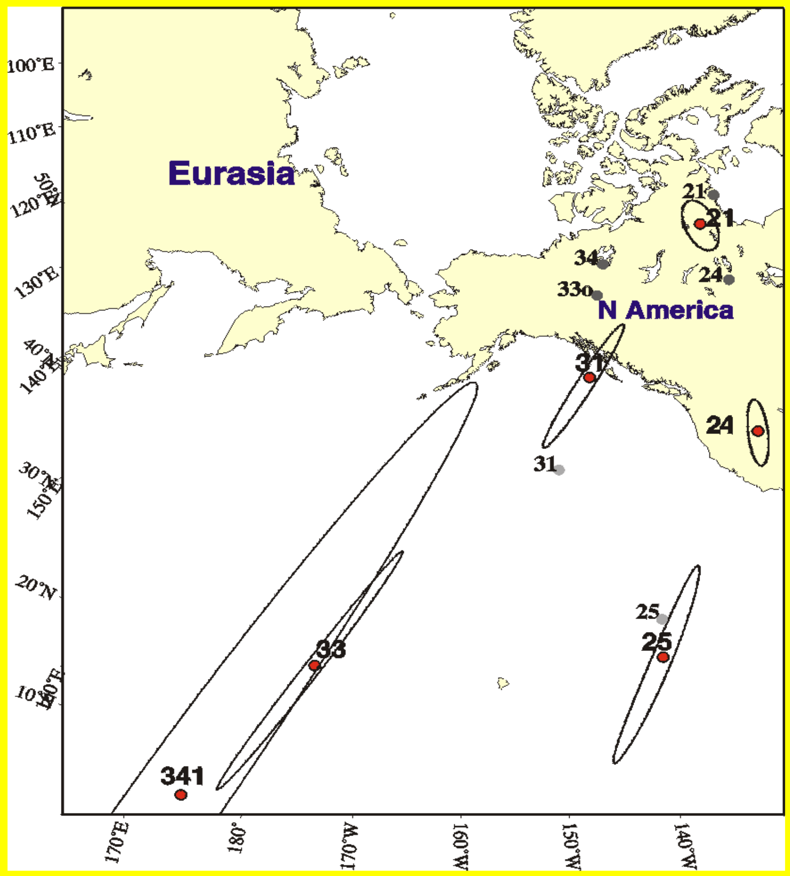
GML: a file format?... Not really.

- XML is the “file format”.
- GML defines “building blocks” for common use:
  - geometric primitives
  - temporal primitives
  - time scales, coordinate reference systems, etc.
- GPML is an “application schema” of GML:
  - combines and extends the GML building blocks.


QuickTime™ and a
Video decompressor
are needed to see this picture.
QuickTime™ and a
Video decompressor
are needed to see this picture.

Need to make plate rotations, uncertainties and metadata available





Atlantic magnetic anomaly grid. Plate vectors and their uncertainties show relative motion between Eurasia and Greenland.

GPML example: Continental crustal boundary

```
<gpml:PassiveContinentalBoundary>
        <gml:name/>
        <gml:description/>
        <gpml:identity>COB-11223344</gpml:identity>
        <gpml:revision>REV-20060413.6</gpml:revision>
        <gpml:deprecatedRevisions/>
        <gml:centerLineOf>
                <gpml:DirectedLineString>
                        <gml:posList srsName=”#WGS84” dimension=”2”>
                                95.3 67.77 180.8 15.36 ....
                        </gml:posList>
                </gpml:DirectedLineString>
        </gml:centerLineOf>
        <gml:validTime/>
        <gpml:featureReconstructionTimeline/>
        <gpml:type> Inner </gpml:type>
        <gpml:continentalSide> Left </gpml:continentalSide>
        <gml:metaDataProperty/>
</gpml:PassiveContinentalBoundary>
```

Here are some of the properties that define where the continental boundary is and what type of boundary it is.

Specifying that the centre line is a “Directed” LineString lets us unambiguously indicate what is on the “Left” and “Right” sides of the continental boundary

GPML example: Continental crustal boundary

```
<gpml:PassiveContinentalBoundary>
        <gml:name/>
        <gml:description/>
        <gpml:identity>COB-11223344</gpml:identity>
        <gpml:revision>REV-20060413.6</gpml:revision>
        <gpml:deprecatedRevisions/>
        <gml:centerLineOf/>
        <gml:validTime>
            <gml:TimePeriod>
                <gml:beginPosition frame=”#GeoTimeScale54” uom=”Ma”>110</gml:beginPosition>
                <gml:endPosition frame=”#GeoTimeScale54” uom=”Ma”>90</gml:endPosition>
            </gml:TimePeriod>
        </gml:validTime>
        <gpml:featureReconstructionTimeline/>
        <gpml:type> Inner </gpml:type>
        <gpml:continentalSide> Left </gpml:continentalSide>
        <gml:metaDataProperty/>
</gpml:PassiveContinentalBoundary>
```

There are also properties that define what period of time the continental boundary exists in, and which Plate IDs it is associated with during its lifetime.

GML lets us specify the time of appearance and disappearance of the Feature, as well as a reference to the geological time scale used to calculate that time.

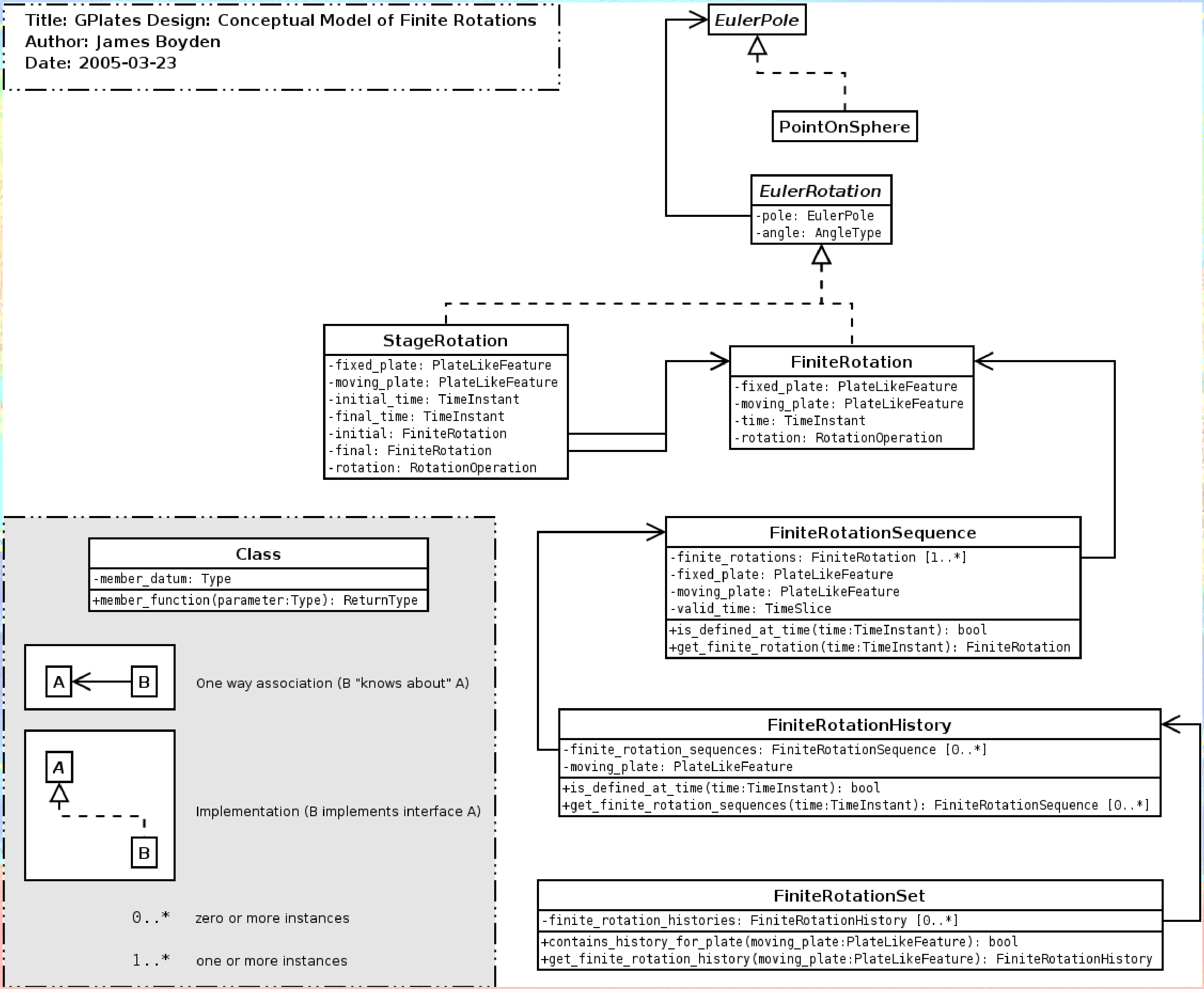
GPML example: Continental crustal boundary

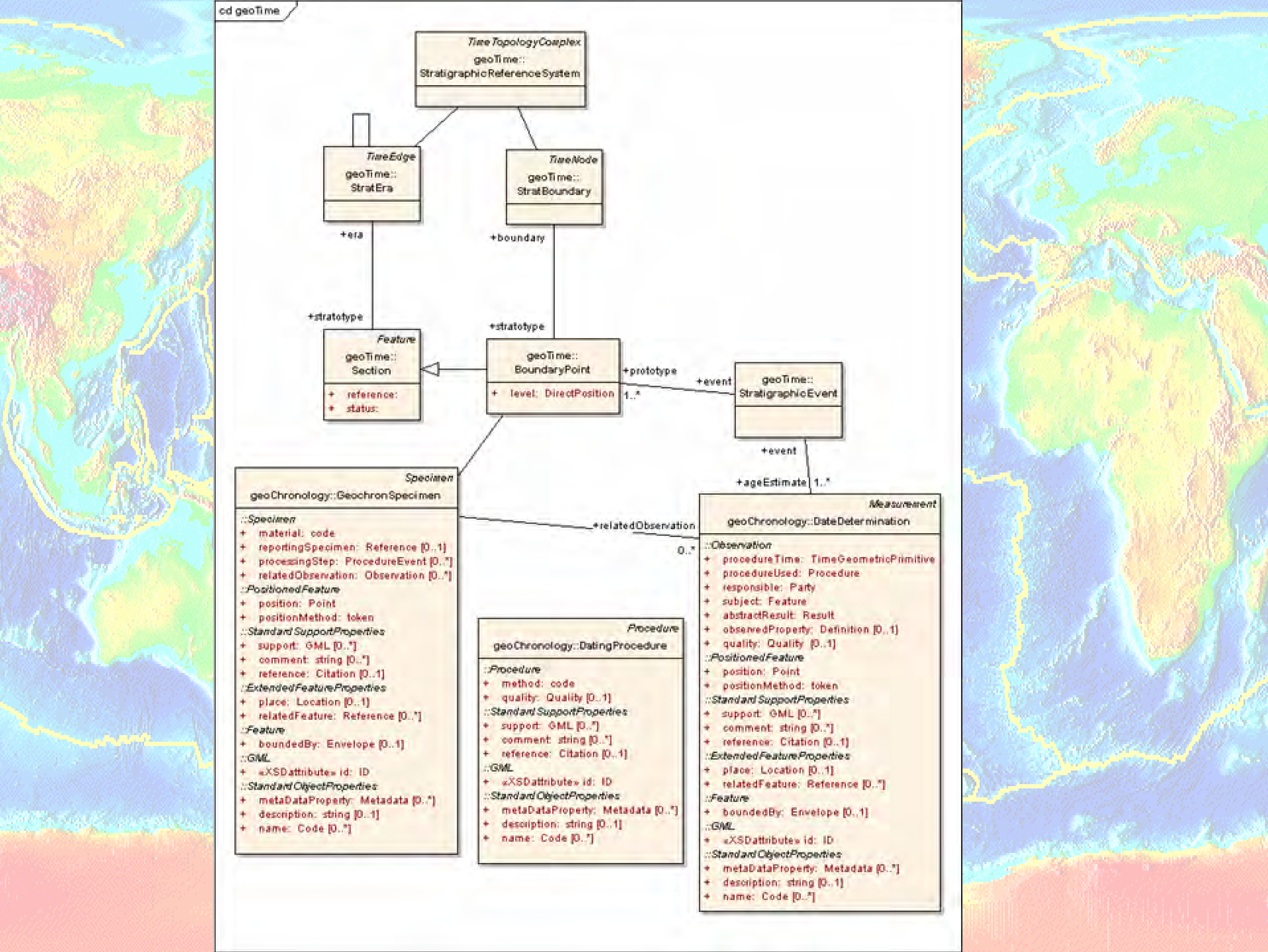
```
<gpml:PassiveContinentalBoundary>
        <gml:name/>
        <gml:description/>
        <gpml:identity>COB-11223344</gpml:identity>
        <gpml:revision>REV-20060413.6</gpml:revision>
        <gpml:deprecatedRevisions/>
        <gml:centerLineOf/>
        <gml:validTime/>
        <gpml:featureReconstructionTimeline>
                <gpml:RotationPlateIDSlice>
                        <gml:validTime>
                                <gml:TimePeriod>
                                        <gml:beginPosition>110</gml:beginPosition>
                                        <gml:endPosition>90</gml:endPosition>
                                </gml:TimePeriod>
                        </gml:validTime>
                        <gpml:plateID>999</gpml:plateID>
                </gpml:RotationPlateIDSlice>
        </gpml:featureReconstructionTimeline>
        …
</gpml:PassiveContinentalBoundary>
```

There are also properties that define what period of time the continental boundary exists in, and which Plate IDs it is associated with during its lifetime.

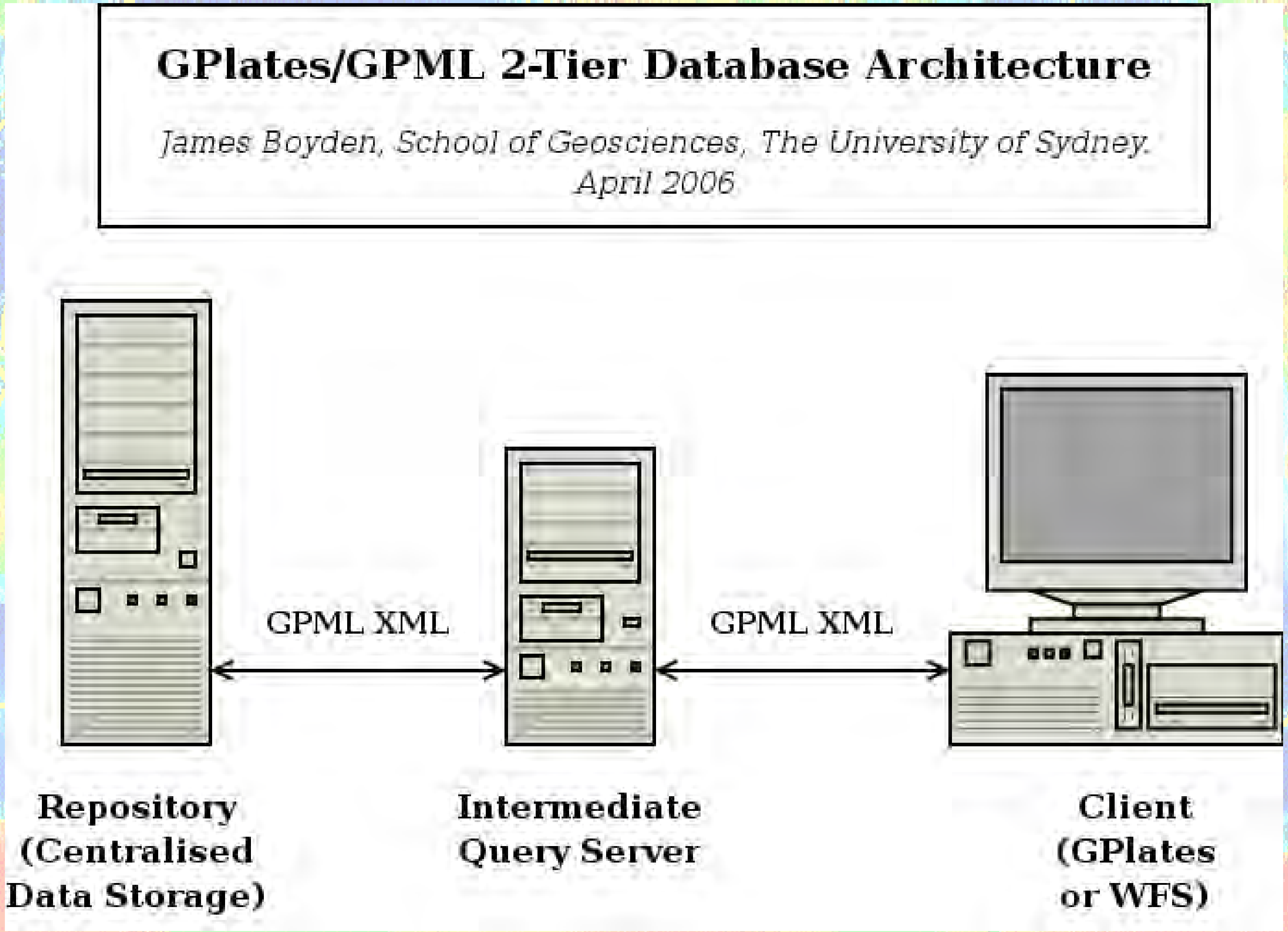
A Feature in GPML can be associated with more than one Plate ID during its lifetime, removing the need to have multiple feature entries which represent the same physical thing.

Title: GPlates Design: Conceptual Model of Finite Rotations
Author: James Boyden
Date: 2005-03-23




cd geoTime
TimeTopologyComplex
geoTime::StratigraphicReferenceSystem
TimeEdge
geoTime::StratEra
TimeNode
geoTime::StratBoundary
+era
+boundary
+stratotype
+stratotype
Feature
geoTime::Section
+ reference:
+ status:
geoTime::BoundaryPoint
+ level: DirectPosition
+prototype
1..*
+event
geoTime::StratigraphicEvent
+event
+ageEstimate
1..*
Specimen
geoChronology::GeochronSpecimen
::Specimen
+ material: code
+ reportingSpecimen: Reference [0..1]
+ processingStep: ProcedureEvent [0..*]
+ relatedObservation: Observation [0..*]
::PositionedFeature
+ position: Point
+ positionMethod: token
::StandardSupportProperties
+ support: GML [0..*]
+ comment: string [0..*]
+ reference: Citation [0..1]
::ExtendedFeatureProperties
+ place: Location [0..1]
+ relatedFeature: Reference [0..*]
::Feature
+ boundedBy: Envelope [0..1]
::GML
+ «XSDattribute» id: ID
::StandardObjectProperties
+ metaDataProperty: Metadata [0..*]
+ description: string [0..1]
+ name: Code [0..*]
+relatedObservation
0..*
Measurement
geoChronology::DateDetermination
::Observation
+ procedureTime: TimeGeometricPrimitive
+ procedureUsed: Procedure
+ responsible: Party
+ subject: Feature
+ abstractResult: Result
+ observedProperty: Definition [0..1]
+ quality: Quality [0..1]
::PositionedFeature
+ position: Point
+ positionMethod: token
::StandardSupportProperties
+ support: GML [0..*]
+ comment: string [0..*]
+ reference: Citation [0..1]
::ExtendedFeatureProperties
+ place: Location [0..1]
+ relatedFeature: Reference [0..*]
::Feature
+ boundedBy: Envelope [0..1]
::GML
+ «XSDattribute» id: ID
::StandardObjectProperties
+ metaDataProperty: Metadata [0..*]
+ description: string [0..1]
+ name: Code [0..*]
Procedure
geoChronology::DatingProcedure
::Procedure
+ method: code
+ quality: Quality [0..1]
::StandardSupportProperties
+ support: GML [0..*]
+ comment: string [0..*]
+ reference: Citation [0..1]
::GML
+ «XSDattribute» id: ID
::StandardObjectProperties
+ metaDataProperty: Metadata [0..*]
+ description: string [0..1]
+ name: Code [0..*]

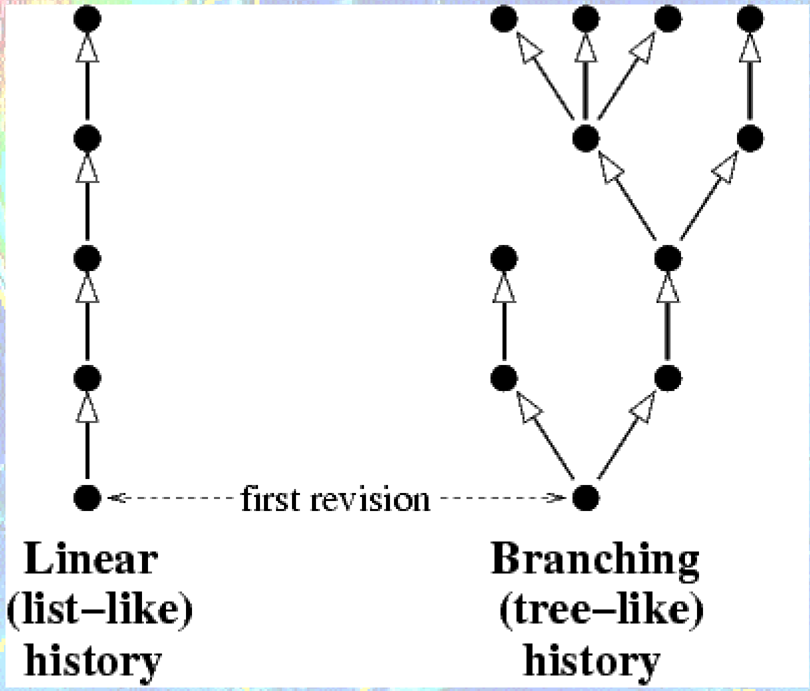
The solution, step 2: Design a suitable data base

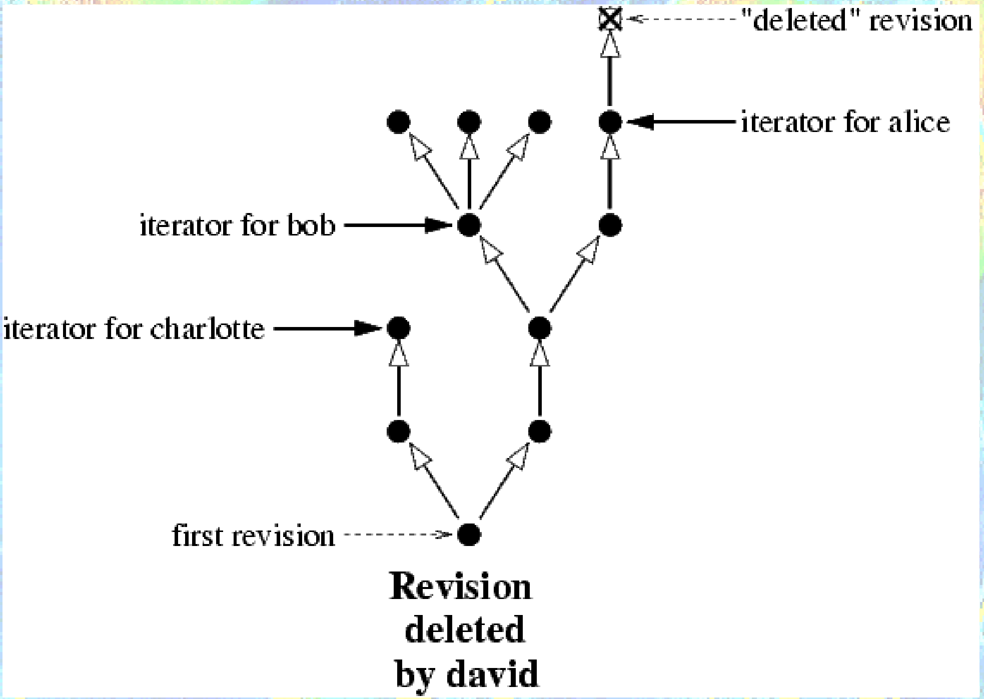


Five Principles of a Scientific Data Repository

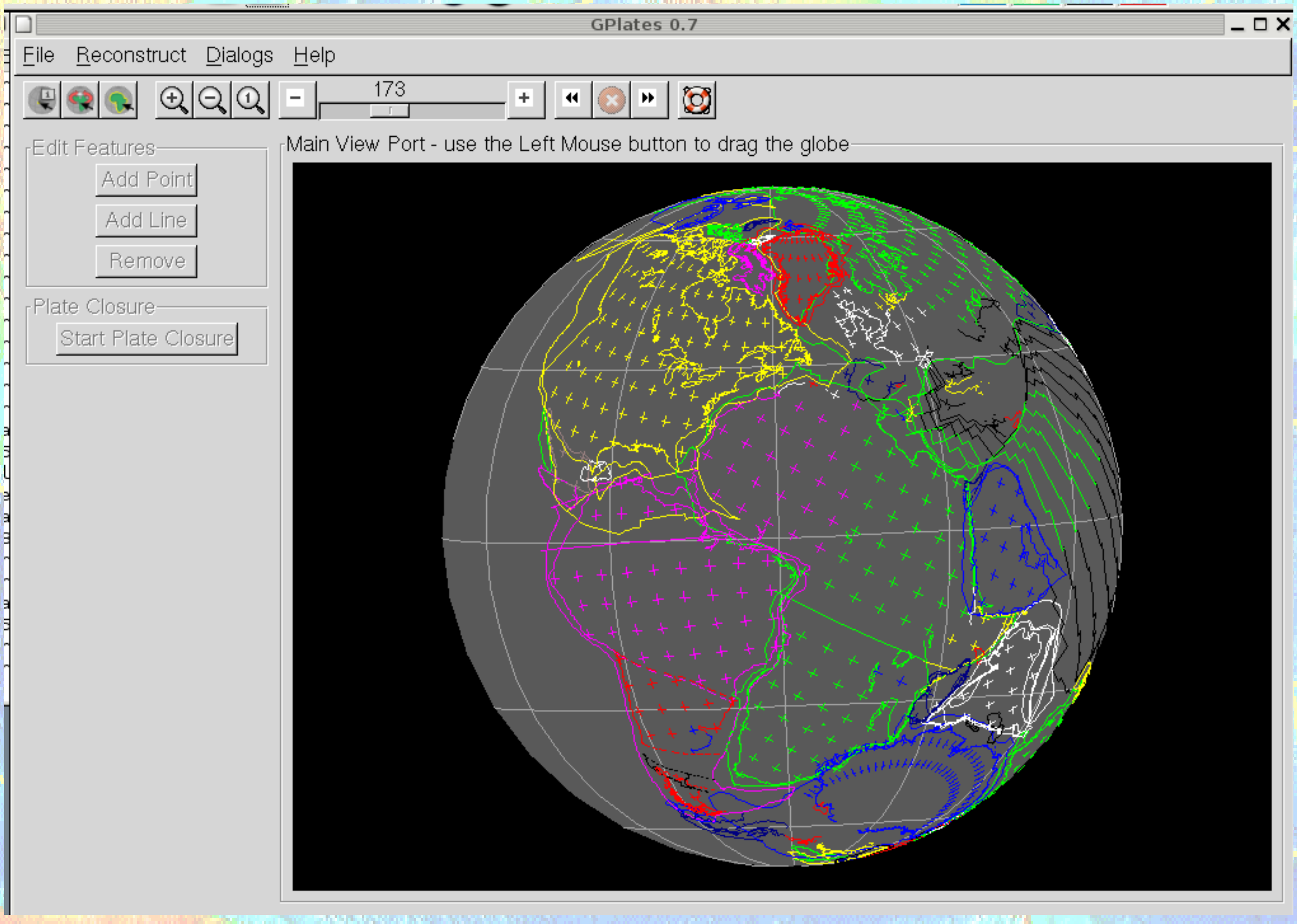
- It should enable precise, accurate and sufficient description of the data
- It should enable collaboration and data-sharing, while still protecting the privacy of non-public data
- It should perform revision control of the data
- It should perform revision control at a per-datum granularity, and enable retrieval of arbitrary sets of (arbitrary revisions of) data
- It should enable branching of data revision histories

Database requirements: Revision history





The solution, step 3: client software

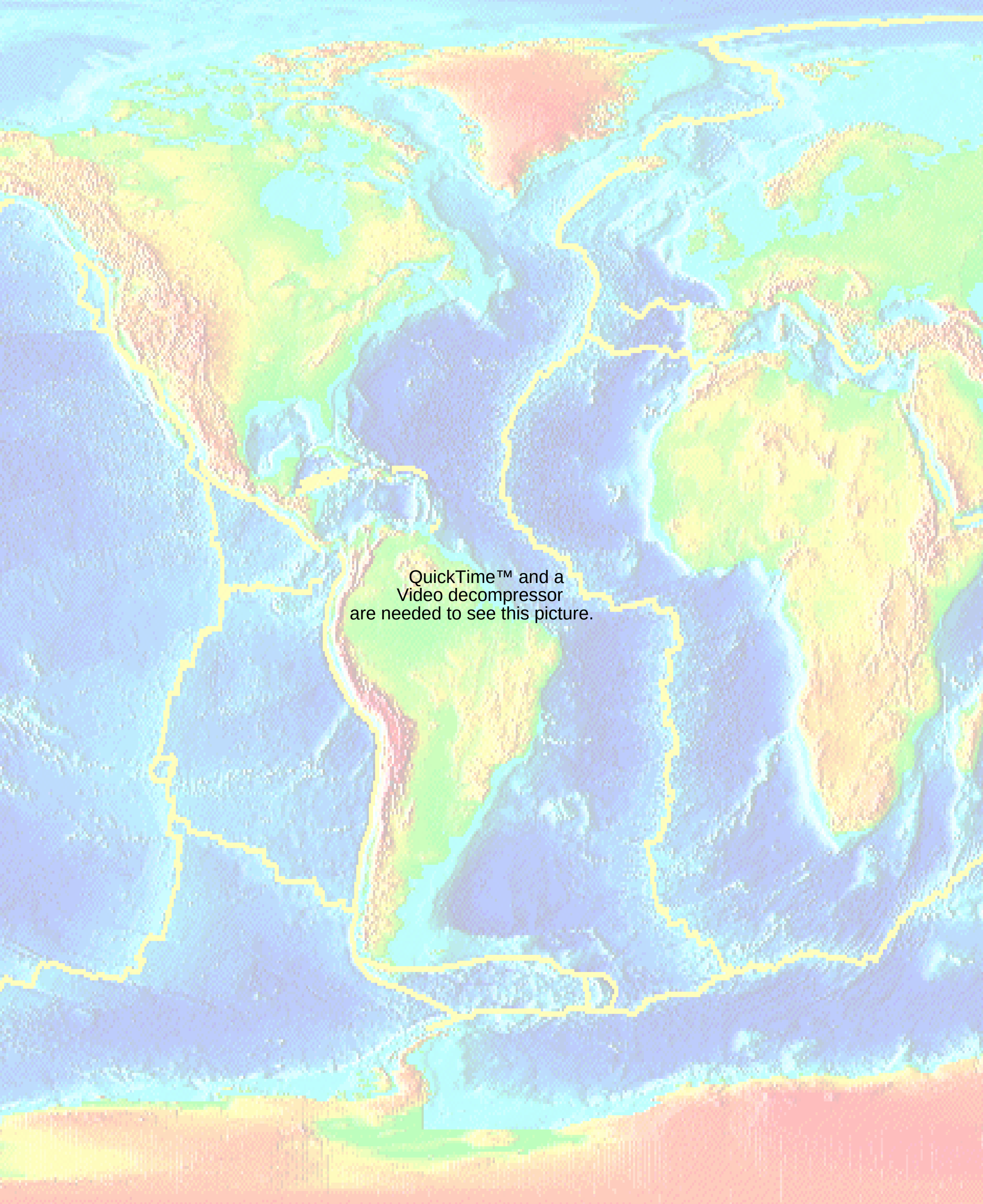


Interactive software enabling the linking of plate kinematics to geodynamics

step 4: Establish workflows to link geodata to simulations

Mantle convection 60Ma to present - density anomaly slice at 500 km

Depends on workflow to link a kinematic model, 3D convection code, and visualization combining model output with reconstructed data



EarthByte System

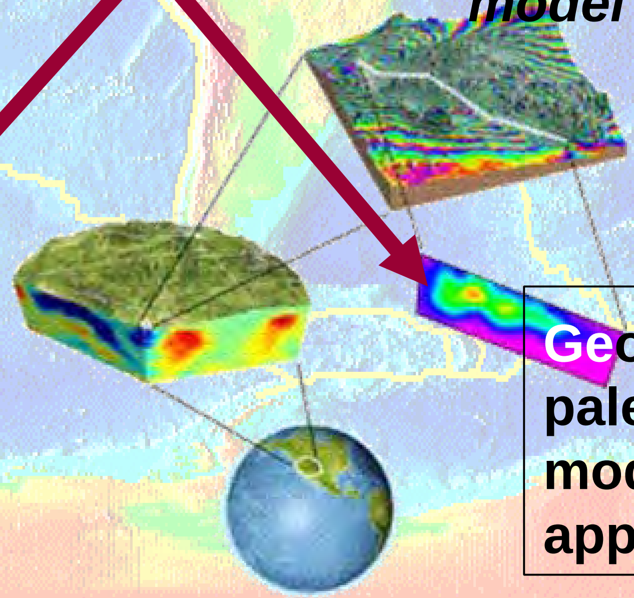
Data to plate encoder via global plate polygons to GPML

Plate Tectonic GIS

GPlates

Interactive manipulation of data/geodyn model outputs

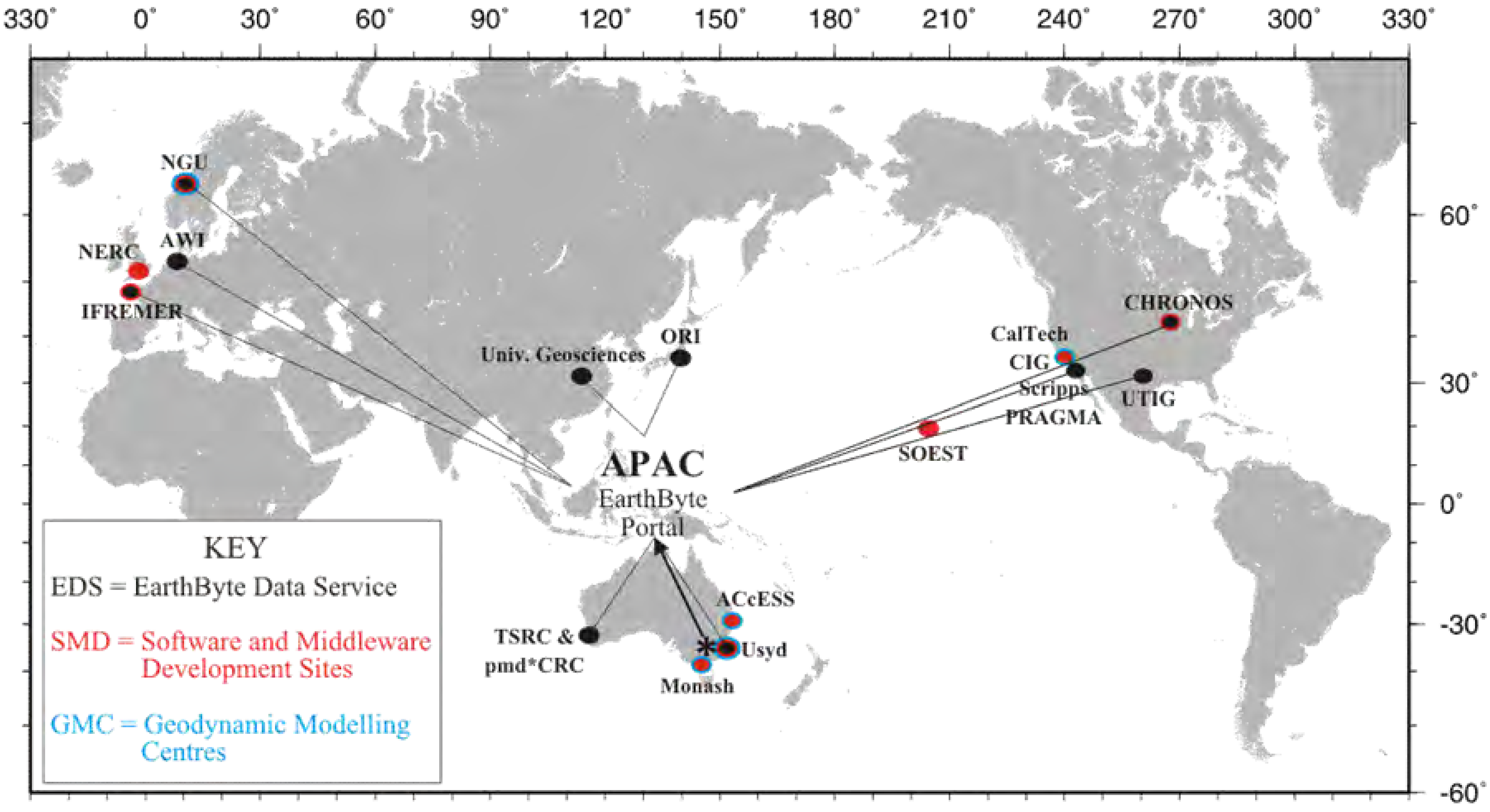
QuickTime™ and a Cinepak decompressor are needed to see this picture.



GPlates map making module (based on GMT software), interactive or scripting-based

Geodynamic/ paleoclimate modelling applications

Currently funded through EarthByte e-research pilot project and APAC



Key (inter)-national connections

- Norwegian Geological Survey (Trond Torsvik)
- Caltech GeoFramework (Mike Gurnis)
- CSIRO (Simon Cox and Rob Woodcock)
- CHRONOS