Units Markup Language (UnitsML)

OASIS Technical Committee
First Official Teleconference
11:00 am, Wednesday
July 12, 2006

1

UnitsML – OASIS TC

- Introductions
- Why are we here?
- Brief background
- Goals
- Brief technical review
  - Markup
  - Units
- UnitsML overview

2

NIST UnitsML Committee Members

- Simon Frechette / MEL-SIMA
- Mark Carlisle / MEL-SIMA
- Bob Dragoset / PL
- Karen Olsen / PL
- Gary Kramer / CSTL
- Peter Linstrom / CSTL
- Kent Reed / BFRL
- Evan Wallace / MEL
- NIST Associates – Ismet Celebi, Ronny Jopp, Alexander Roth

3

Why are We Here?

- NIST customers have expressed the need for a system like UnitsML
- NIST provides the source for unit definitions in the U.S. – SP 811
- This project is providing computer-useable access to these definitions
- An OASIS TC is an appropriate mechanism for recommending an XML standard

4

Project Background

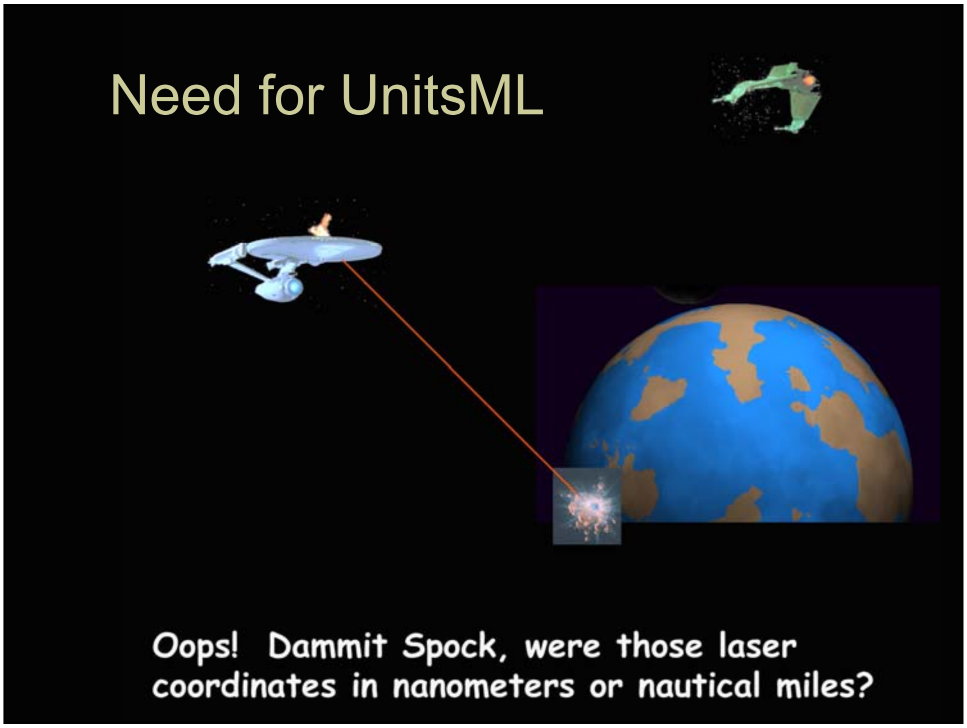
- Initial concept proposed by LBNL (F. Olken)
- NIST Physics Laboratory proposed to develop units markup language facilitating unambiguous exchange of units information
- Other NIST projects have shared needs for units markup (AnIML, AEX, MatML)
- Group formed to develop consensus for units representation

5

Goals

- To provide the U.S. national reference markup language for units and quantities
  - Provide traceability to NIST SP 811
  - Facilitate measurement data markup
  - Enable the interchange of data between markup languages in different domains
  - Prolong utility of marked-up archived data
- ~~To define the units and quantities~~

6



Need for UnitsML

- XML schemas are being created for specific communities, and many include markup for units—but different approaches are used
- MathML & GML
  - Mostly complete incorporation of units in MLs
  - Will probably not adopt UnitsML, but may reference UnitsML
- AnIML, AEX, & MatML
  - Incorporated units into MLs with placeholders for an authoritative source of units information

8

Units of Measure Overview

- Unit of measure
  - Any determinate amount or quantity adopted as a **standard of measurement** for other amounts or quantities of the same kind, *e.g.*, meter
- Quantity
  - The measurable, countable, or comparable **property** or aspect of a thing, *e.g.*, length

The *value* of a quantity is its magnitude expressed as the product of a number and a unit, and the number multiplying the unit is the *numerical value* of the quantity expressed in that unit.

9

Units of Measure Overview

- SI base units (7)
- SI derived units
  - Special derived (22)
  - Derived
- SI multiples and submultiples
- Units outside the SI
  - Acceptable for use within the SI
  - Temporarily acceptable for use within the SI
  - Unacceptable for use within the SI

10

SI Base Units & Quantities

Unit Name	Unit Symbol	Quantity Name	Quantity Symbol	Dimension Symbol
meter	m	length	l	**L**
kilogram	kg	mass	m	**M**
second	s	time	t	**T**
ampere	A	electric current	I	**I**
kelvin	K	thermodynamic temperature	T	**Θ**
mole	mol	amount of substance	n	**N**
candela	cd	luminous intensity	I_v	**J**

Examples: $l = 7.24$ m $v = 6.32$ m/s

$m = 1.53$ kg dim $v = \mathbf{L}\,\mathbf{T}^{-1}$

11

Some SI Special Derived Units

Derived quantity	Special name	Special Symbol	Expression in terms of	
			other SI units	SI base units
plane angle	radian	rad		$m \cdot m^{-1} = 1$
solid angle	steradian	sr		$m^2 \cdot m^{-2} = 1$
frequency	hertz	Hz		s^{-1}
force	newton	N		$m \cdot kg \cdot s^{-2}$
pressure, stress	pascal	Pa	N/m^2	$m^{-1} \cdot kg \cdot s^{-2}$
energy, work, quantity of heat	joule	J	$N \cdot m$	$m^2 \cdot kg \cdot s^{-2}$
power, radiant flux	watt	W	J/s	$m^2 \cdot kg \cdot s^{-3}$
electric charge	coulomb	C		$s \cdot A$
electric potential	volt	V	W/A	$m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$
capacitance	farad	F	C/V	$m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$

12

SI Derived Units and Multiples/Submultiples

There are essentially an unlimited number of derived units.

quantity	unit
area	square meter
velocity	meter per second
	millimeter per microsecond
parameters of virial equations of state $(\text{volume/quantity})^n$	(cubic meter per mole) to the nth power
fifth order virial coefficients (n=5)	$m^{15} \cdot mol^{-5}$

13

Approach to UnitsML & Related Activities

- XML schema (UnitsML) – **OASIS TC**
  - For incorporating into other MLs; and
  - Describing database contents
- Units database (UnitsDB)
  - Units, prefixes and quantities with dimensions
  - Unique naming (indexing) scheme
  - Representation in terms of other units, including conversion factors
- Tools for using UnitsML
  - Web Services access to UnitsDB
  - Off-line units file tool (selection, user-defined units, updating)
  - Examples for using UnitsML in other MLs

14

The UnitsML "Meta" Approach

- Direct approach

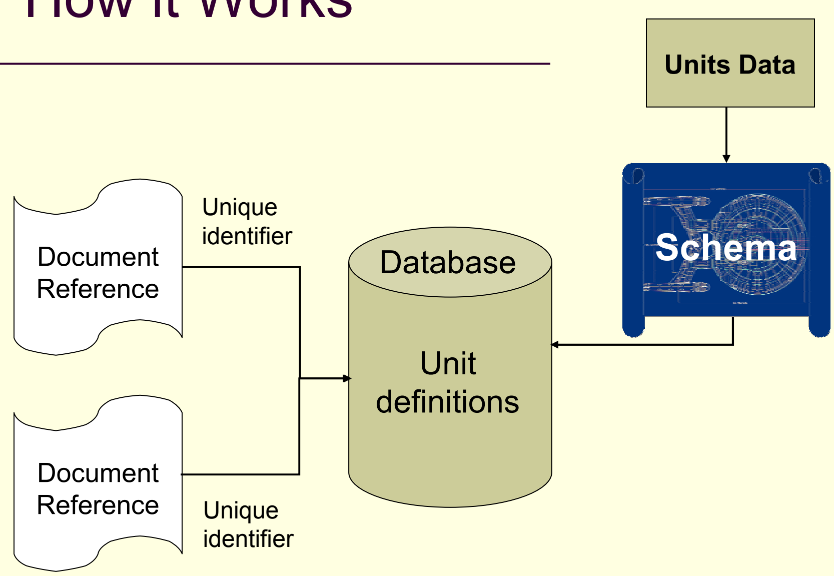
```
<physQuant name="tableMass" value="18.5" units="kilogram" />
<physQuant name="myAccel" value="2.1" units="mm.μs^−2" />
```

- UnitsML "Meta" approach

```
<UnitsML:Unit numericID="User1" symbolicID="mm.μs-2">
    …
    <UnitRepresentation prefixReference="m"
                unitReference="m" power="1" />
    <UnitRepresentation prefixReference="μ"
                unitReference="s" power="−2" />
    …
</UnitsML:Unit>
```

15

How it Works



16

Unit Identifiers

Unambiguous method for identifying each unique unit of measure – Numeric String vs. ASCII symbol

Numeric string: NISTu6

- Pros:
  - No ambiguities
- Cons:
  - Contains no taxonomic information (*i.e.*, not recognizable to humans—possibly prone to human error)
  - Requires a "look-up" table for display

17

Unit Identifiers

Unambiguous method for identifying each unique unit of measure – Numeric String vs. ASCII symbol

ASCII symbol: cm.s^-1

- Pros:
  - Contains taxonomic information (*i.e.*, recognizable to humans—less likely for human errors)
  - Internationally recognized/understood
  - A parser could display this symbol for presentation (*i.e.*, cm ⁄ s or cm · s^{-1})
- Cons:
  - Ordering is not unique (*e.g.*, N.m vs. m.N)
  - Not all units have accepted ASCII symbols (*e.g.*, nautical mile, knot)
  - Some unit symbols are not unique (*e.g.*, Btu, rad)

18