# Workshop on Standard Exchange Format (WoSEF)

**Susan Elliott Sim**
Dept. of Computer Science
University of Toronto
10 Kings College Rd.
Toronto, Ontario
Canada M5S 3G4
+1 (416) 978-4158
simsuz@cs.utoronto.ca

**Ric Holt**
Dept. of Computer Science
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1

+1 (519) 888-4567, x4671
holt@plg.uwaterloo.ca

**Rainer Koschke**
Institut für Informatik
Universität Stuttgart
Breitwiesenstraße 20-22
70565 Stuttgart
Germany
(+49) +711-7816-206
koschke@informatik.uni-stuttgart.de

**Keywords**
data interchange, exchange format, software tools

## 1 INTRODUCTION

A common exchange format for sharing data extracted from source code is necessary to advance the state of the art in many branches of software engineering, such as reverse engineering, software visualization, metrics, program comprehension, and testing. With a common exchange format researchers can more easily leverage each others' tools and take a "best of breed" approach when solving problems. To take an example from reverse engineering, one would be able to select the best parsing and analysis tools, send the output to a clustering program, display the results using a visualization tool from another group, and use the results to effect changes to the software. A standard exchange format would allow researchers to build a repository of example systems, or "guinea pigs", and compare or combine results from various tools.

There have been many efforts to create a standard data exchange format. Some of these are general-purpose exchange formats that can be adapted to data about software, while others are specifically for software. Some examples are XML (eXtensible Mark-up Language) [8], with a specialized form, XMI (XML Metadata Interchange format) [7] for UML diagrams, RDF (Resource Descriptor Format) [3], RSF (Rigi Standard Form) [4], TA (Tuple Attribute Language) [5], GraX [1], and CDIF (CASE Data Interchange Format) [1]. These formats vary in the amount of support and use they receive. While CDIF initially had a great deal of industry support, development on this format has since stopped and it has been absorbed by OMG as part of the effort to develop XMI. This proliferation of exchange formats underlines both the need for a standard format and the lack of consensus on one.


ICSE 2000 Limerick Ireland


## 2 ISSUES

In arriving at a standard exchange format, there are essentially three issues to be resolved: the representation format (syntax), the elements to be represented (schema and semantics), and the activities needed to foster adoption and tool support. Some of the questions associated with each of these issues are listed below.

**Representation Format**

- How should the data be represented, i.e. what syntax should be used?
- What are desirable properties of the representation?
- Will any of the existing exchange formats meet our needs?

**Elements to be Represented**

- What information about software should be included?
- What schema should the format have?
- What are the semantics of a particular element? How do we reconcile the concept of a module across programming languages and paradigms? How should we use similar terms such as function, procedure, method, and sub-routine?
- How do we represent common concepts such as line numbers? How do we resolve naming conflicts?
- How should we represent architectural information?

**Adoption Activities**

- How do we encourage a standard exchange format to be adopted?
- How should this information be disseminated?

- What tools, libraries, or APIs are needed to encourage the adoption?
- Who should be encouraged to adopt the exchange format?

## 3 WORKSHOP GOALS

The goal of this workshop is contribute to the evolution of a consensus on a standard exchange format. Holding this workshop at ICSE, the flagship conference of the software engineering community, allows researchers from all branches of the discipline to participate in the process. By raising the level of awareness, we hope to move the software engineering community towards consensus on a standard exchange format for software data.

This workshop will be discussion-oriented and the organization of the sessions will be based on the position papers and how they addressed the development of a standard exchange format. It is expected that the discussions will focus more on the first two issues, syntax and semantics, that the last one. However, by the end of the day we hope to reach some kind of consensus which will advance this last issue of adoption.

## REFERENCES


1. J. Ebert, B. Kullbach, and A. Winter. "GraX—An Interchange Format for Reengineering Tools", In *Sixth Working Conference on Reverse Engineering* (WCRE), pages 89-98, Atlanta, GA, October, 1999.
2. J. Ernst. *Introduction to CDIF.* Available at <http://www.eigroup.org/cdif/intro.html>
3. RDF Web Site. Available at <http://www.w3.org/RDF/>
4. RSF Web Site. Available at <http://www.rigi.csc.uvic.ca/rigi/manual/user.html>
5. TA Web Site. Available at <http://plg.uwaterloo.ca/~holt/papers/ta.html>
6. WoSEF Web Site. Available at <http://www.cs.utoronto.ca/~simsuz/wosef>
7. XMI Web Site. Available at <http://www.software.ibm.com/ad/features/xmi.html>
8. XML Web Site. Available at <http://www.w3.org/XML/>