"Internationalisation of SIF and harmonisation with other specs/standards"

Draft Version 1.0
June 20th 2003

Authors:

Luis Anido-Rifón (University of Vigo)

Niall Sclater (University of Strathclyde)

Contents

1 Foreword

This work is directly related to the European Commission's Mandate M/280 "Standardisation mandate to CEN,CENELEC and ETSI in the domain of 'Learning and Training Technologies & Educational Multimedia Software'", covering the development of a workplan for standards related activities in relation to Learning Technologies. The first agreed workplan was published as CWA 14040 in the year 2000. In a second step, the original workplan has been expanded. The development of Interoperability frameworks for exchange of information between diverse management systems in cooperation with OASIS [1] was one of the recommendations for further work.

This report was prepared by an appointed Project Team within the CEN/ISSS Learning Technologies Workshop. The CEN/ISSS Learning Technologies Workshop (WS-LT) [2] agreed at its meeting in Madrid on April 04/05 2002 to provide interoperability specifications for a range of information exchange, initially aimed at school systems, but also applicable in a wider lifelong learning context. This work item will support the output of the OASIS (Open Architecture and Schools in Society) project that enables different applications and systems to share information.

OASIS is an ambitious project to promote virtual communities in the school system, partly by harmonizing information exchange standards. It has been proposed in the framework of the User-friendly information society (the IST Programme) under the Fifth Framework Programme, which defines the European Commission activities in the field of RTD (Research Technological development and Demonstration). The work of OASIS will initially be based upon the SIF project (Schools Interoperability Framework) [3] but will be adapted by the partners to meet European needs. Some of the original specifications will be discarded and some additional ones will be created.

The work item will focus on collaborating with the OASIS partners in ensuring that the specifications they produce do not duplicate international standards activity, are appropriate for a broader audience and are suitable for e-Europe standardisation.

This CEN Workshop Agreement identifies those elements within the SIF Data Model that should be further studied in order to properly cope with a multi-cultural and multi-lingual environment like Europe. In addition, this document also identifies those parts of the SIF specification that overlap with other existing standards for the e-learning and other domains and whose adoption may be considered by SIF-based initiatives.

1.1 Report Structure

This report is structured, as follows:

- Introduction to SIF (section 2)

- Scope (section 3)

- Abbreviations (section 4)

- Stakeholders (section 5)

- SIF Data Model (section 6)

- Internationalisation of the SIF Data Model (section 7)

- Harmonisation of SIF with other standards/specifications (section 8)

- Detailed analysis of the SIF Data Model (section 9)

- Recommendations (section 10)

- References (section 11)

1.2 Project Team members contacts

Luis Anido Rifon (PT leader)
Email: lanido@ait.uvigo.es
University of Vigo (Spain)

Niall Sclater (PT expert)
Email: n.sclater@strath.ac.uk
University of Strathclyde (England)

2 Introduction to SIF

The Schools Interoperability Framework (SIF) [3] is an industry initiative to develop an open specification for ensuring that K-12 instructional and administrative software applications work together more effectively. SIF is not a product, but rather an industry-supported technical blueprint for K-12 software that will enable diverse applications to interact and share data seamlessly; now and in the future.

Much of today's educational software is proprietary, meaning that data cannot be easily shared and transferred, if at all.

The lack of interoperability to user means:

- Applications and their data are isolated from one another

- Redundant data entry is common

- Disconnected applications support costs

- Data reporting is costly and inefficient

- Data is inaccessible for decision makers

The lack of interoperability also forces administrators to spend large amounts of time and money to keep all of their institutions' various software programs up to date. With multiple sets of data, there is also an increased possibility that information being dispersed by a school to parents or others may not be entirely correct or current.

The goal of SIF is to eliminate redundant data entry and improve the schools' ability to gather and report data. This will allow schools to reduce support and maintenance costs because administrators can purchase software that enables data sharing without requiring costly add-ons to be built. This initiative will also make the process of purchasing software easier for administrators, because they will know that as long as new software is SIF certified, it will be able to interact with other SIF certified applications

A standard set of specifications used by all education software companies would:

- Ensure that data is entered only once in one application, and automatically propagates to other applications

- Allow applications to exchange data more effectively

- Enable schools to create powerful reports by accessing data from different applications

- Allow educators to deliver reports securely via the Internet to various organization

3 Scope

The purpose of this report is twofold:

1. Identify those elements within the SIF Data Model that should be extended in order to be used in a worldwide environment. For this, special attention was paid to the identification of issues related to multi-cultural and multi-lingual environments.

2. Harmonise the SIF Data Model with other existing specifications/standards in the e-learning domain. Also, for those aspects not directly related with e-learning (e.g. bus routes, cafeteria tickets, etc.) the PT has investigated whether standards in those areas exist or not.

Although the purpose of this CWA is to internationalise SIF in a worldwide context, special emphasis was put in those issues that may affect the multi-cultural and multi-lingual Europe.

3.1 Overview

The use of the SIF model in a multicultural and multilingual environment raises some issues about localisation and/or correct interpretation of attributes used in different contexts. For this reason it is mandatory to identify those SIF data elements that are cultural dependent or need to be further detailed or are not clearly explained.

The purpose is to support the global applicability and understanding of the SIF Data Model, investigating both on general culture topics (different calendars, languages, etc) or more specific technical issues (character sets, standards for dates, etc.).

3.2 Focal Aspects

This work will support the outputs of the OASIS (Open Architecture and Schools in Society) project that enable different applications and systems to share information. This CWA is intended to be used by OASIS, and other initiatives worldwide that try to localise the SIF specification in their particular contexts, to develop their own data models taking into account existing standards/specifications that may be reused and internationalisation issues that may have been left out of the SIF specification. At the same time the SIF specification developers may consider this CWA for future updates of the specification.

3.3 Normative References

The following normative documents contain provisions which, through reference in this text, constitute provisions of this CWA.

ISO/IEC 10646-1: 2000. Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- Part 1: Architecture and Basic Multilingual Plane

ISO 8601: 2000. Data elements and interchange formats -- Information interchange -- Representation of dates and times

ISO 639: 1988. Code for the representation of names of languages

ISO 639-2: 1998. Codes for the representation of names of languages -- Part 2: Alpha-3 code

ISO 4217:2001. Codes for the representation of currencies and funds

ISO 31:1992. Quantities and Units. Part 0: General Principles, Units and Symbols.

ISO 31:1992. Quantities and Units. Part 1: Space and time.

ISO 3166-1:1997 Codes for the representation of names of countries and their subdivisions -- Part 1: Country codes

4 Abbreviations

For the purposes of the present document, the following abbreviations apply:

CEDEFOP	*Centre Européen pour le Développement de la Formation Professionnelle*
CEN	*European Committee for Standardization (Comitée Européen de Normalisation)*
CENELEC	*European Comittee for Electrotechnical Standardization*
CWA	*CEN Workshop Agreement*
ETB	*European Treasury Browser*
ETSI	*European Telecommunications Standards Institute*
EUN	*European SchoolNet*
IEC	*International Electrotechnical Comission*
IEEE	*Institute of Electrical and Electronics Engineers*
ISO	*International Organization for Standardization*
ISSS	*Information Society Standardization System*
LIP	*IMS Learner Information Package*
NCES	*National Center for Educational Statistics*
OASIS	*Open Architecture and Schools in Society*
PAPI	*Public And Private Information*
RTD	*Research Technological development and Demonstration*
SIF	*Schools Interoperability Framework*
SIG	*Special Interest Group*
SPEEDE	*Standardization of Postsecondary Education Electronic Data Exchange*
STUDENTHB	*Student Data Handbook*
W3C	*World Wide Web Consortium*
WS/LT	*CEN/ISSS Workshop on Learning Technologies*
XML	*eXtensible Markup Language*

5 Stakeholders

5.1 Introduction

Stakeholders are persons or organizations who can affect, or are affected by each recommendation made in this report, the main stakeholders are identified. It has to be recognized that there is a diversity between different stakeholders and that they may have competing interests. The various roles may be overlapping: a person or group may have several overlapping roles.

5.2 Direct Stakeholders

Direct stakeholders are persons or organizations who are directly affected by the recommendation made in this report:

- Developers of the SIF specification.

- Developers of SIF-based specifications (e.g. OASIS)

5.3 Indirect Stakeholders

Indirect stakeholders are persons or organizations who are indirectly affected by the recommendations made in this report:

- Schools

- Teachers

- Software enterprises developing SIF-compliant software

- Software managers at schools

5.4 Standards Organizations and Activities

A number of organizations working on standardisation are, or may become, particularly interested in this report. These include:

- CEN/ISSS [4] has several workshops involved with standards related to learning technologies, such as Metadata for Multimedia Information, E-commerce, and the Workshop in Learning Technologies (WS-LT).

- ISO [5] and IEC [6] have produced many globally accepted standards for a wide range of domains, the subcommittee ISO/IEC JTC1 SC36, Information Technology for Learning, Education, and Training, has several working groups and ad hoc committees. CEN/ISSS/LT-WS has a Category A liaison relationship with SC36.

- W3C [7] produces many generic and domain specific standards and specifications for the World Wide Web and the Internet.

There are several significant European activities connected with learning technologies, these include:

- ARIADNE [8] is a European foundation developing concepts and tools for computer-based and telematics-supported remote authoring, teaching and learning, with a strong emphasis on the sharing and reuse of electronic learning material. The foundation partners contribute to international standardisation activities, notably for metadata, and have developed an operational infrastructure.

- OASIS, Open Architecture and Schools Interoperability project aims to maintain public educational systems as leaders in Internet-based education by developing the concept of the small school virtual community. It mains objectives are to prepare a blueprint for interoperability that allow schools to share information in a virtual community and to develop a cost-effective school Zone Server.

- eEurope [9] is an initiative intended to accelerate positive change in the Union. It aims at ensuring this change towards the Information Society is cohesive, not divisive. eEurope also aims at bringing the benefits of the Information Society to the reach of all Europeans.

- EUN [10], European Schoolnet, has several work plans as part of a European network of national networks for school level ICT. Activities include examining issues around metadata and IPR.

- PROMETEUS [11] – PROmoting Multimedia access to Education and Training in EUropean Society, was co-founded with the WS/LT and has a membership of over 500 European corporate and academic organizations. There are several active Special Interest Groups (SIGs) that are producing recommendations and guidelines. Having built consensus, PROMETEUS may feed recommendations for standards into WS/LT. This is part of a two-way process, as WS/LT may identify topics that are not mature enough for standardisation but are appropriate for research and consensus building within PROMETEUS.

There are also significant global activities connected with learning technologies, like:

- The IMS Global Learning Consortium [12] has centres in several countries and has produced a wide range of specifications for learning technologies.

- The Advanced Distributed Learning (ADL) Initiative [13] , sponsored by the Office of the Secretary of Defence (OSD), is a collaborative effort between US government, industry and academia to establish a new distributed learning environment that permits the interoperability of learning tools and course content on a global scale.

6 SIF Data Model

The SIF Data Model describes the data objects and elements which may be transferred between agents via the ZIS. It is divided into 2 clear parts: the first one presents the common elements that will be used by other more complex objects and the second one where the objects which belong to each working group are described. Among the SIF working groups, some of them have not defined any data object because it's not its target (e.g. Customer Involvement, Requirements, Communications & Accords (CIRCA))

Common Elements

- **Address**: an address which will occur within objects and elements such as *StaffPersonal* and *StudentPersonal/StudentAddress*, etc.

- **Demographics**: this describes ethnicity, gender, country of birth, language etc.

- **Email:** email addresses of people within the system.

- **GridLocation**: a latitude and longitude used within other elements such as *BusStopInfo* and *Address.*

- **MeetingTime**: the meeting times and periods for a course.

- **OtherID**: lists other IDs associated with objects such as barcodes and identifiers of schools.

- **PhoneNumber**: of people within the system.

Data objects from specific working groups

Food services objects

- **Student meal**: communicates the current meal status of a pupil.

Human resources and financials objects

- **Billing**: specifies an amount to be billed.

- **Payment**: contains information about the payment of a billing object.

Library automation objects

- **LibraryPatronStatus**: this is a complex object detailing library books held, fines due etc.

Student information objects

- **AttendanceCodeInfo**: used for attendance records i.e. absences, reasons for absences etc.

- **RoomInfo**: information about rooms in a school – their size, capacity etc.

- **RoomType**: the type of room e.g. cafeteria, classroom etc.

- **SchoolCourseInfo**: information about courses.

- **SchoolInfo**: information about the school such as contact details and name of principal.

- **SectionInfo**: a section appears to be a part of a course.

- **StaffPersonal**: personal information relating to a staff member. Most of these elements are defined already in *Name, Email, Address*, etc.

- **StudentContact**: similar to *StaffPersonal* this contains contact details for a pupil's "contacts" (e.g. a parent).

- **StudentDailyAttendance**: time in and time out for a pupil on a particular day together with notes if necessary.

- **StudentPersonal**: all personal information related to a pupil. Most of this is contact details.

- **StudentPicture**: contains a picture or the URL of a picture of the pupil.

- **StudentSchoolEnrollment**: when a pupil enrolled at the school, their current academic level etc.

- **StudentSectionEnrollment**: similar to the above but relating to a student's enrolment in a section of a course.

- **TermInfo**: information about a term, its start and end date etc.

Transportation and geographic information objects

- **BusEquipment**: may refer to a wheelchair for instance.

- **BusInfo**: details about a bus, its capacity and any special equipment on it.

- **BusRouteDetail**: the schedule for a bus route, its stops and times.

- **BusRouteInfo**: all information about a bus route including the name of the driver, the distance and duration of the route.

- **BusStopInfo**: information about a bus stop – a description and location.

- **StudentTransportInfo**: transportation information about a student – their eligibility and a reference to a *BusRouteDetail* for instance.

7 Internationalisation of the SIF Data Model

This section analyses the Data Model presented in the Schools Interoperability Framework (SIF) specification version 1.1 [14]. Concrete actions are presented in section 10. These recommendations should be considered when adopting SIF in other settings different from the US, for example, Europe.

7.1 Objects that should be further studied

There are 2 data objects that show the need of adaptation of the SIF Data Model when used in a different cultural/educational setting: *StudentSectionEnrollment* and *TermInfo*.

This two objects refer to a particular period of time: section and term. Different educational settings use different durations for these periods (terms may have three or four months). A detailed analysis of this internationalisation needs as far as each particular object is concerned is presented in section 9.

The rest of this section focuses on those issues that can be applied in general to the SIF Data Model, i.e. two or more objects would be affected by these recommendations.

7.2 Repertoires for data coding

The SIF specification does not identify the character set repertoire that should be used to represent textual information. Nevertheless, this is an important issue when dealing with multi-lingual environments. There are several recommendations that could be done at this point.

1. The SIF specification may be updated to define a repertoire character set wide enough to cope with as many languages as possible, including Asian languages and those where the writing style (left to right or viceversa and top to bottom or viceversa) is different from the western style. An interesting starting point would be to analyse the ISO/IEC 10646 and UNICODE.

2. Those SIF-based specification that adapt the American specification to their particular cultural and lingual environment should take into account what the appropriate repertoire character set is. In this case, special attention should be paid to those cases where it is possible to have SIF zones or federation of SIF zones that cover different cultural areas with different needs as far as character set repertoires is concerned.

The correspondent resulting PT recommendation is **10.1**

7.3 Actions on *Language* elements

The value space for *Language* (Demographics object) and *LanguageOfInstruction* (SectionInfo object) is the ANSI/NISO Z39.53-2001: *Codes for the Representation of Languages for Information Interchange* [15]. This standard defines a set of three-letter codes for language identification. However, this language representation has several disadvantages for its use in a worldwide environment.

This format does not allow the use of dialects or variations. For example, a variation/dialect of Spanish, Asturian, is not included in Z39.50. This is the reason to recommend the following actions:

1. Adopt a standard wider than Z39.50. For example ISO 639-2:1988, which is a three-letter code for the representation of languages. This standard covers, for example the above mentioned dialect of Spanish. In ISO 639-2 threee sets are provided, one for bibliographic applications (ISO 639-2/B), and one for terminology applications (ISO 639-2/T). The bibliographic version provides a three letter

code based on the English name of the language. The terminology version provides a three letter code based on the Native name of the language, and, at present, it is not available for all. The concrete standard to be used needs to be fixed.

2. Although ISO 639-2:1988 covers a wider set of languages than the currently used in SIF Z39.53, there are situations where this is not enough:

   a. Variations of the same language depending on the country where it is spoken cannot be managed using only the ISO standard. For example, there should be a mechanism to distinguish between the Mexican Spanish or the Spaniard Spanish, or the US English and the British English.

   b. Variations of the same language depending on the region where it is spoken cannot be managed using only the ISO standard. For example, the following French variants: Norman, Picard, Wallon, Angevin, Berrichon, Bourbonnais, Bourguignon, Franc-Comtois, Gallo, Lorraine, Poitevin, Santogeais) from http://www.ethnologue.com/show_language.asp?code=FRN do not have an ISO code.

The following format is proposed as an initial starting point:

Langcode("-"Subcode("-"Variant))

where

   Langcode: Three letter code according to the standard ISO 639-2:1988

   Subcode: Two letter code for identification of countries ISO 3166-1:1997

   Variant: Code for the variation of the language identified by the previous two codes.

Some examples are shown below:

   eng-US-philadelphia

   eng-GB-newcastle

   spa-ES-andalucia

   spa-MX-monterrey

   fre-FR-paris

The correspondent resulting PT recommendations are **10.2** and **10.3** .

7.4 Actions on *Date*-value elements and attributes

SIF Data Model does not specify how dates should be represented. The only reference in the specification to date format is in the section where encapsulation of SIF messages over HTTP is presented. In this case the format is, obviously, the format defined in the HTTP RFC (RFC 2616): CCYYMMDD. This format is followed by the SIF Data Model.

Nevertheless, this CWA recommends to explicitly define the date format in the conceptual data model description section within the SIF specification. For this, an initial proposal may be a widely used standard like ISO 8601. The proposed format by this PT is in the form YYYY-MM-DD, as ISO 8601:2000 recommends. The date portion only represents dates in the Common Era (CE). The date portion follows the

Gregorian calendar for dates after October 15, 1582, and the Julian calendar for dates prior to October 15, 1982, independent of locale. Other cases should be represented using the "Description" data item.

In addition, if SIF is to be used in multi-cultural environments a more elaborated proposal should be taken. Specially, if a SIF zone may cover a multi-cultural geographical area or if SIF federation may lead to exchange of data among heterogeneous cultural regions.

Although the format proposed above assures interoperability, other national formats for dates or Eras (reference points) are not taken into account. The suggestion is to include an additional local Date that corresponds to the DateTime in ISO 8601:2000 format whenever possible (e.g. year 100 in the Buddish era corresponds to a BCE year). Provided there exists a unique format/era reference in each country, localization would be identified using the country code (using ISO 3166-1:1997). This also supports different formats to represent dates (e.g. DD-MM-YYYY).

In order to assure interoperability in a multicultural environment the ISO 8601:2000 standard could be used as the canonical form to represent dates. The alternative representation proposed here should be used whenever providing a localized reference point in time is especially relevant (e.g. the first day of the Chinese year is not so clear pointed out in the Gregorian calendar).

Proposal for DateTime:

The PT proposes to add a new aggregate data element that can be used to allow DateTime localization. In this way, the new definition for the DateTime item would be, as follows (shaded rows correspond to LOM data elements, category 3 represents the new proposal):

Nr	Name	Explanation	Value Space
1	DateTime	A point in time with accuracy at least as small as second	ISO 8601
2	Description	Description of the date	
3	DateTimeLocale		
3.1	Locale	Identifier for the country where the LocalizedDateTime applies.	Country code from the code set ISO 3166-1:1999
3.2	Source	Identifier of the source (standard or recommendation) that defines the specific date format for the country	Repertoire of the ISO/IEC 10646 Note: A typical Source could be the URL of the document where the standard is specified
3.3	LocalizedDateTime	The same point in time as in the DateTime element but formatted according to the specific localization	Repertoire of the ISO/IEC 10646

Example of an XML instance using this approach:

```
<DATETIME>2003-12-25</DATETIME>

<DATETIMELOCALE>

    <LOCALE>US</LOCALE>

    <SOURCE>http://standards.org/us/calendarSpecs.pdf</SOURCE>

    <LOCALIZEDDATETIME>12/25/03</LOCALIZEDDATETIME>

</DATETIMELOCALE>

<DATETIMELOCALE>

    <LOCALE>UK</LOCALE>

    <LOCALIZEDDATETIME>25/12/03</LOCALIZEDDATETIME>

</DATETIMELOCALE>

<DATETIMELOCALE>

    <LOCALE>AE</LOCALE>

    <SOURCE>http://standards.org/ae/calendarNumSpecs.pdf</SOURCE>

    <LOCALIZEDDATETIME>1/11/1424</LOCALIZEDDATETIME>

</DATETIMELOCALE>

<DATETIMELOCALE>

    <LOCALE>AE</LOCALE>

    <SOURCE>http://standards.org/ae/calendarTextSpecs.pdf</SOURCE>

    <LOCALIZEDDATETIME>1 Dhu'l-Qa'dah 1424</LOCALIZEDDATETIME>

</DATETIMELOCALE>
```

An alternative approach of the XML binding for the conceptual model would be to use, whenever possible, attributes instead of sub-elements:

```
<DATETIME>2003-12-25</DATETIME>

<DATETIMELOCALE LOCALE="US" SOURCE=" http://standards.org/us/calendarSpecs.pdf">

12/25/03

</DATETIMELOCALE>

<DATETIMELOCALE LOCALE="UK">25/12/03</DATETIMELOCALE>

<DATETIMELOCALE LOCALE="AE" SOURCE=" http://standards.org/ae/calendarNumSpecs.pdf">

1/11/1424

</DATETIMELOCALE>

<DATETIMELOCALE LOCALE="AE" SOURCE=" http://standards.org/ae/calendarTextSpecs.pdf">
```

1 Dhu'l-Qa'dah 1424

</DATETIMELOCALE>

The correspondent resulting PT recommendation is **10.4**

7.5 Actions on *currency-based* values

There are several elements in the SIF Data Model that are defined to encapsulate numbers representing monetary amount. In particular:

-   BilledAmount in the object BillingObject

-   ReceivedAmount in the object PaymentObject

-   CircTx/FinInfo/Amount, FineAmount, RefundAmount in the object LibraryPatronStatus

In order to cope with multi-currency environments, the following recommendations are proposed:

1. The previously presented objects should be extended to allow the specification of the particular currency being used to indicate each amount (e.g. US dollars, European euros, Mexican pesos, etc.). A straightforward solution, which may be considered as an initial approach, is to create an aggregate data element (e.g. monetaryamount) with two sub-elements (e.g. amount, currency) being the latter the identifier of the currency being used. Also, the different representations for amounts may have to be taken into account (e.g. for the amount one thousand and fifty six cents the representation may be 1.000,56 or 1,000.56)

2. The identifier for the currency must follow a widely used standard. An initial approach may be to use ISO 4217:2001. *Codes for the representation of currencies and funds*, which includes a three letter code for each currency. A further study is needed to check if this standard also supports different formats for the representation of amounts.

A proposal for a conceptual model for currency-type values can be:

Name	Explanation	Value Space
Currency	Identifier of the currency unit	ISO 4217:2001
Amount	Number indicating the amount of money	

Below are shown two possible XML bindings of this conceptual model.

In the first two examples it is shown a aggregate-data element approach.

<!-- Example 1: Currency element with Euro currency type-->

<MonetaryAmount>

```
    <Currency>EUR</Currency>

    <Amount>45.96</Amount>

</MonetaryAmount>
```

```
<!-- Example 2: Currency element with US Dollar currency type-->

<MonetaryAmount>

    <Currency>USD</Currency>

    <Amount>52.00</Amount>

</MonetaryAmount>
```

The same result can be achieved with a different approach using attributes instead of sub-elements, as it is done in SIF specification (cf. OtherId element):

```
<!-- Example 1: Currency element with Euro currency type (attribute approach)-->

<MonetaryAmount Type="EUR">45.96</MonetaryAmount>
```

```
<!-- Example 2: Currency element with US Dollar currency type (attribute approach)-->

<MonetaryAmount Type="USD">52.00</MonetaryAmount>
```

The correspondent resulting PT recommendation is **10.5**

7.6 Actions on *measurement-type* values

There are several objects with elements whose value represent measurements (e.g. RouteDistance in object BusRouteInfo or Size in object RoomInfo). The SIF Data Model makes no reference to the possibility of using these objects in an environment where several measurement formats are used. Two recommendations are made:

1. The previously presented object should be extended to allow the specification of the particular measurement type being used (e.g. miles, feet, pounds, kilometres, kilos). A straightforward solution, which may be considered as an initial approach, is to create an aggregate data element (e.g. measurement) with two sub-elements (e.g. amount, unit) being the latter the identifier of the measurement type being used.

2. The identifier for the measurement must follow a widely used standard. An initial approach may be to use ISO 31:1992, *Quantities and Units. Part 0: General Principles, Units and Symbols. Part 1: Space and time.*

A proposal for a conceptual model for measurement-type values can be:

Name	Explanation	Value Space
Unit	Identifier of the measurement unit	ISO 31:1992
Amount	Number indicating the measure	

Below are shown two possible XML bindings of this conceptual model.

In the first two examples it is shown a aggregate-data element approach.

<!-- Example 1: Longitude measurement using Kilometers as unit-->

<Measurement>

    <Unit>KTM</ Unit>

    <Amount>6.7</Amount>

</Measurement>

<!-- Example 2: Size measurement using square inches as unit-->

<Measurement>

    <Unit>INK</ Unit>

    <Amount>12.00</Amount>

</Measurement>

The same result can be achieved with a different approach using attributes instead of sub-elements, as it is done in SIF specification (cf. OtherId element):

<!-- Example 1: Currency element with Euro currency type (attribute approach)-->

<Measurement unit="KTM">6.7</Measurement>

<!-- Example 2: Currency element with US Dollar currency type (attribute approach)-->

<Measurement unit="INK">12.00</Measurement>

The correspondent resulting PT recommendation is **10.6**

7.7 Actions on *Vocabularies*

Many SIF data elements use vocabularies, defined by the SIF specification or externally defined in the set of codes SPEEDE (Standardization of Postsecondary Education Electronic Data Exchange) [16] and NCES's STUDENTHB (Nacional Center for Education Statistics) [17]. These codes are composed of an abbreviation, which, in turn, may be a two-number code, one, two or three-letter code or a combination of numbers and letters. The code is accompanied by a textual description of its meaning in the English language.

There are two actions that may be taken on SIF vocabularies as far as internationalisation is concerned:

1. Elements within each vocabulary are described using a textual description in the English language. These descriptions should be translated into other languages. This action may be taken by the own SIF community following a gradual translation process. Translations should be carried out in the short term for those languages more widely spoken worldwide. An alternative to this option may be to put this responsibility on each SIF-based initiative, which will be responsible for translating SIF vocabularies into those languages relevant for its geographical context.

2. Many SIF vocabularies may provide a set of values that are not suitable or do not cover completely the specific scope out of a US school environment. For example, the attribute *type* in the common object *OtherId* has as its value space: *District-assigned number, State-assigned number*, *Migrant number, US government VISA number*. This has to be extended/modified for each particular cultural/political setting. An exhaustive analysis of the SIF Data Model is needed to identify those vocabularies that need to be further developed. An initial proposal is presented in section 9.

The correspondent resulting PT recommendations are **10.7** and **10.8**

8 Harmonisation of SIF with other standards/specifications

8.1 Introduction

Due to the main aim of the SIF specification, the main information involved in message exchange is related to student management data and administration services. This means that there are few objects dealing with educational (e.g. performance, preference) information. The next section will analyse each object and element identified in the SIF specification. From this analysis it will be clear that harmonisation will be needed with those standards and specifications related to learner information. This subsection introduces some of them:

- ***IMS Learner Information Package (LIP)*** [18]

Learner Information is a collection of information about a Learner (individual or group learners) or a Producer of learning content (creators, providers or vendors).

IMS LIP is a structured information model. The model defines fields into which the data can be placed and the type of data may be put into this fields. Typical data might be the name of the learner, a course or training complete, a learning objective, a preference of a particular type of technology, and so on.

- ***Public And Private Information (PAPI) Learner*** [19]

The PAPI Learner Standard describes a particular subset of all possible types of learner information. Learner information is considered a subset of general information about learning technology.

- ***vCard*** [20]

vCard is the electronic business card. It is a powerful new means of Personal Data Interchange (PDI) that is automating the traditional business card. Some important features are:

  - vCards carry vital directory information such as name, addresses (business, home, mailing, parcel), telephone numbers (home, business, fax, pager, cellular, ISDN, voice, data, video), email addresses and Internet URLs (Universal Resource Locators).

  - All vCards can also have graphics and multimedia including photographs, company logos, audio clips such as for name pronunciation

  - Geographic and time zone information in vCards let others know when to contact you.

  - Of course, vCard support multiple languages

8.2 Main harmonisation issues identified

1. Specifications introduced above define standardised descriptions for learner personal information. This is the main area where SIF may be harmonized with external specifications/standards. Most SIF objects contain elements with personal data (e.g. *Address*, *Email*, *StudentContact*, etc.) This CWA recommends to harmonise personal data information included in SIF specification using any of the standards/specifications introduced above. The specific standard that should be used needs to be decided after a deeper analysis by the SIF and SIF-based inititiatives specification developers.

2. Many SIF data elements use vocabularies, defined by the SIF specification or externally defined in the set of codes SPEEDE (Standardization of Postsecondary Education Electronic Data Exchange) [16] and NCES's STUDENTHB (Nacional Center for Education Statistics) [17]. These codes are composed of an abbreviation, which, in turn, may be a two-number code, one, two or three-letter code or a combination of numbers and letters. The code is accompanied by a textual description of its meaning in the English language. In the internationalisation section (section 7) it was mentioned the need to describe these meanings in alternatives languages other than English, with special

attention to those languages in the context of each particular SIF-based specification. Also, additional elements may be needed to cover the specific geographic and cultural needs of each SIF-based environment. For the latter, SIF and SIF-based specification developers must take into account existing initiatives to develop taxonomies and vocabularies for the educational domain within their context. For example, in Europe there exist several well-known providers of educational vocabularies in its wider sense (e.g. ETB (*European Treasury Browser*) [21] or CEDEFOP (European Centre for the Development of Vocational Training) [22]).

3. The common data element *MeetingTime* defines a time slot for a specific course. The use of the vCalendar specification in this object should be further analysed.

The correspondent resulting PT recommendations are **10.9** and **10.10**

9 Detailed analysis of SIF Data Model

This section will be developed to include a detailed analysis of the SIF Data Model in order to identify:

- Data objects and data elements that should be extended to cope with a multi-cultural and multi-lingual environment (e.g. Europe)

- Vocabularies that SIF uses for a concrete data element that do not cover properly non-US school systems.

- Those recommendations presented in section 10 that can be applied to each data element or data object.

The next table will be further developed in the eventual version of this section.

Address	Street/StreetPrefix	To be further discussed for internationalisation
	Street/StreetSuffix	To be further discussed for internationalisation
Demographics	Ethnicity	To be further discussed for internationalisation
	BirthDateVerification	To be further discussed for internationalisation
EnglishProficiency	To be further discussed for internationalisation	
Name	A new last name element is needed for certain countries (e.s. Spain)	
Student meal	To be further discussed for internationalisation	
SchoolCourseInfo	StateCourseCode	To be further discussed for internationalisation
	DistrictCourseCode	To be further discussed for internationalisation
SchoolInfo	To be further discussed for internationalisation	
StudentSchoolEnrollment	StaffAssigned	To be further discussed for internationalisation
	ExitDate	To be further discussed for internationalisation
BusRouteInfo	To be further discussed for internationalisation	

To be further developed.

10 Recommendations

This section contains concrete recommendations based on the analysis made in the previous three sections.

10.1 Recommendation 1 – Data coding

The need   The SIF specification does not identify the character set repertoire that should be used to represent textual information and this is an important issue when dealing with multi-lingual environments.

Besides there is no way to specify the writing styles (left to right/right to left or horizontal/vertical).

Stakeholders   SIF specification developers, SIF-based specification developers

Action   Update the SIF specification to define a repertoire character set wide enough to cope with as many languages as possible. A possible starting point would be to analyse the ISO/IEC 10646 and UNICODE.

An alternative would be to define, for each cultural environment where a SIF-based specification is to be used, which repertoire set is needed to properly represent the concrete language/s used in that context.

Outcome   An update on the SIF specification containing the selected repertoire character set or an indication to take this in mind when defining a localised version of SIF

Action priority   High

Organization   SIF, SIF-based initiatives

Timescale   6 months

10.2 Recommendation 2 – *Language* elements

The need   The value space for *Language* (Demographics object) and *LanguageOfInstruction* (SectionInfo object) is the ANSI/NISO Z39.53-2001: *Codes for the Representation of Languages for Information Interchange*.

This language representation has several disadvantages for its use in a worldwide environment. This format does not allow the use of dialects or variations. For example, a variation/dialect of Spanish, Asturian, is not included in Z39.50.

Stakeholders   SIF specification developers, SIF-based specification developers

Action   To specify the use of a standard wider than Z39.50. For example ISO 639-2:1988, which is a three-letter code for the representation of languages.

Outcome   To formally explain and present in the SIF specification the use of ISO 639-2.

Action priority High

Organization   SIF, SIF-based initiatives

Timescale   6 months

10.3 Recommendation 3 – *Variations* of the language

The need	Variations of the same language depending on the country where it is spoken cannot be managed using the current format to represent languages in SIF.
	Variations of the same language depending on the region where it is spoken cannot be managed using only the ISO standard.
Stakeholders	SIF specification developers, SIF-based specification developers
Action	To adopt the following format:

Langcode("-"Subcode("-"Variant))

where

Langcode: Three letter code according to the standard ISO 639-2:1988

Subcode: Two letter code for identification of countries ISO 3166-1:1997

Variant: Code for the variation of the language identified by the previous two codes.

Outcome	A new format for those elements representing the language in order to identify the language variations.
Action priority	Medium
Organization	SIF, SIF-based initiatives
Timescale	1 year

10.4 Recommendation 4 – *Date*-value elements

The need   SIF Data Model does not specify how dates should be represented. The only reference in the specification to date format is in the section where encapsulation of SIF messages over HTTP is presented. Specification of dates formats should be done at the conceptual data level.

Stakeholders   SIF specification developers, SIF-based specification developers

Action   To explicitly define the date format in the conceptual data model description section within the SIF specification. For this, an initial proposal may be a widely used standard like ISO 8601. The proposed format by this PT is in the form YYYY-MM-DD, as ISO 8601:2000 recommends.

To add an extension to the specification of the DateTime item in order to give the possibility to use a Localized Date-Time when it can be relevant for a particular cultural context. The extension recommended by this CWA is reported at section 7.4 of this document.

Outcome   A clearly specified way of representing dates.

A specification of the DateTime format including the possibility to indicate localized dates and times.

Action priority Medium

Organization   SIF, SIF-based initiatives

Timescale   1 year

10.5 Recommendation 5 – *Currency-based* values

The need　　There are several elements in the SIF Data Model that are defined to encapsulate numbers representing money (e.g. BilledAmount in the object BillingObject)

Stakeholders　SIF specification developers, SIF-based specification developers

Action　　The previously presented objects should be extended to allow the specification of the particular currency being used to indicate each amount.

An initial solution could be to create an aggregate data element (e.g. monetaryamount) with two sub-elements (e.g. amount, currency) being the latter the identifier of the currency being used. The identifier for the currency must follow a widely used standard. An initial approach may be to use ISO 4217:2001. *Codes for the representation of currencies and funds*.

Outcome　　A new data element as described above

Action priority　Medium

Organization　SIF, SIF-based initiatives

Timescale　　1 year

10.6 Recommendation 6 – *Measurement-type* values

The need The SIF Data Model makes no reference to the possibility of using objects whose value represents measurements in an environment where several measurement formats are used (e.g. kilometres and miles).

Stakeholders SIF specification developers, SIF-based specification developers

Action To extend those data objects to allow the specification of the particular measurement type being used (e.g. miles, feet, pounds, kilometres, kilos).

An initial approach, is to create an aggregate data element (e.g. measurement) with two sub-elements (e.g. amount, unit). The identifier for the measurement unit must follow a widely used standard. An initial approach may be to use ISO 31:1992, *Quantities and Units. Part 0: General Principles, Units and Symbols. Part 1: Space and time.*

Outcome A new data element as described above

Action priority Medium

Organization SIF, SIF-based initiatives

Timescale 1 year

10.7 Recommendation 7 – Internationalisation of Vocabularies

The need   Many SIF data elements use vocabularies defined by the SIF specification or externally defined in two sets of codes. These codes are composed of an abbreviation and a textual description of its meaning in the English language

Stakeholders   SIF specification developers, SIF-based specification developers

Action   To translate descriptions into other languages. This action may be taken by the own SIF community following a gradual translation process. Translations should be carried out in the short term for those languages more widely spoken worldwide. An alternative to this option may be to put this responsibility on each SIF-based initiative

Outcome   Translations of the vocabularies descriptions into other languages besides English

Action priority   Low

Organization   SIF, SIF-based initiatives

Timescale   2 year

10.8 Recommendation 8 – *Proposals for Vocabularies*

The need   Many SIF vocabularies may provide a set of values that are not suitable or do not cover completely the specific scope out of a US school environment

Stakeholders   SIF specification developers, SIF-based specification developers

Action   To extend/modify vocabularies for each particular cultural/political setting.

Outcome   New vocabularies suitable for each SIF zone

Action priority   Medium

Organization   SIF, SIF-based initiatives

Timescale   1 year

10.9 Recommendation 9 – Harmonization of *Personal Information*

The need   The main information involved in message exchange is related to student management data and administration services.

Most SIF objects contains elements with personal data

Stakeholders   SIF specification developers, SIF-based specification developers

Action   To harmonise personal data information included in SIF specification using any of the following standards/specifications: IMS LIP (*Learner Information Package*), PAPI (*Public And Private Information Learner*) or vCard

Outcome   Standardized learner information instead of defining another model

Action priority   Medium

Organization   SIF, SIF-based initiatives

Timescale   1 year

10.10 Recommendation 10 – *Harmonization of vocabularies*

The need   Many SIF data elements use vocabularies, defined by the SIF specification or externally defined in two sets of codes. The space value covered by this vocabularies may have been previously defined in other existing initiatives.

Stakeholders SIF specification developers, SIF-based specification developers

Action   To take into account existing initiatives to develop taxonomies and vocabularies for the educational domain within their context.

            In Europe there exist several well-known providers of educational vocabularies in its wider sense (e.g. ETB (*European Treasury Browser*) [9] or CEDEFOP (European Centre for the Development of Vocational Training)[10])

Outcome   Identification of existing initiatives with existing taxonomies and vocabularies for the educational domain that may apply to the SIF Data Model.

Action priority Low

Organization SIF, SIF-based initiatives

Timescale   2 year

11 References

[1] OASIS, *Open Architecture and Schools in Society* http://oasis.cnice.mecd.es/index.htm
[2] CEN ISSS WS-LT, CEN *Information Society Standardization System Learning Technologies Workshop* http://cenorm.be/isss/Workshop/LT/Default.htm
[3] SIF http://www.sifinfo.org
[4] CEN/ISSS, *CEN Information Society Standardization System* http://www.cenorm.be/isss
[5] ISO, *International Organization for Standardization* http://www.iso.ch
[6] IEC, *International Electrotechnical Commission* http://www.iec.ch
[7] W3C, World Wide Web Consortium http://www.w3.org
[8] ARIADNE http://www.ariadne-eu.org/
[9] eEurope http://prometeus.org
[10] European Schoolnet EUN http://www.eun.org
[11] PROMETEUS PROmoting Multimedia access to Education and Training in EUropean Society http://prometeus.org
[12] IMS Project http://www.imsproject.org
[13] ADL, *The Department of Defense's (DoD) Advanced Distributed Learning* http://www.adlnet.org
[14] Schools Interoperability Framework (SIF) specification version 1.1 http://www.sifinfo.org/spec.html
[15] ANSI/NISO Z39.53-2001: *Codes for the Representation of Languages for Information Interchange* http://www.niso.org/standards/std_info_retrieval.html#Z39.53
[16] SPEEDE, *Standardization of Postsecondary Education Electronic Data Exchange* http://nces.ed.gov/edi/speedeExp.asp
[17] NCES STUDENTHB, *National Center for Educational Statistics Student Data Handbook* http://nces.ed.gov/pubs2000/studenthb/
[18] LIP, *Learner InformationPackage* http://www.imsproject.org/accessibility/acclipv1p0pd/imsacclip_infov1p0pd.html
[19] PAPI, *Public And Private Information* http://edutool.com/papi
[20] vCard, *virtual Card* http://www.imc.org/pdi/vcardoverview.html
[21] ETB, *European Treasury Browser* http://www.en.eun.org/eun.org2/eun/en/etb/content.cfm?lang=en&ov=7208
[22] CEDEFOP, *European Centre for the Development of Vocational Training* http://www.cedefop.eu.int/