

Alliance Project Database

Technical Specification

Third revision

Author: Marco Ridoni, Lunaria (marco.ridoni@virgilio.it)

Date: Feb. 12, 2009

For your consideration. Please see the list of changes with regard to the first draft (last page)

Introduction

This specification aims at defining a standard format for project (workcamp) data interchange among Alliance member and partner organizations.

Some key points are:

- Interoperability: the format chosen shall not be linked, or make assumptions of any kind, about the underlying database system or data management workflow. Only informations relevant to partners and members, or of general interest, should be considered for inclusion.
- Internationalization: in order to avoid formatting problems when importing/exporting data, informations like dates, boolean ("yes/no") values, accented/umlauted characters should be properly stored in a consistent and well-defined manner.
- Simplicity: the format must be simple enough for an average-skilled programmer to write a filter for, or for a database system to directly import the data; yet it shouldn't exclude any information deemed necessary just for the sake of simplicity.
- Extensibility: while the project database is currently meant as a repository for project data, to be seeded and used by each organization, in the future - and after a thorough discussion - it could also be used to provide more services. Different and backward-compatible format versions (e.g. 1.0, 1.1) should be possible, to account for specific needs.
- Integration: data fields should be clearly marked as "required" or "optional", to account for the different needs and possibilities of each organizations.

1) Basics

1.1) Transport

The "transport" format of choice is XML. While XML will be the reference format, a CSV (Comma Separated Values) export file, as well as a simple text output will be provided to make integration in existing project management systems easier. CSV and text versions of the data should always be "in sync" with the XML one. A specification regarding CSV/text output, with sample data, will follow.

1.2) Encoding

XML files should be encoded in the UTF-8 charset to preserve extended character information (accent/umlaut/diacritic, etc.). The appropriate "charset" tag should also be used at the top of the XML file.

1.3) Integration

Data fields should be clearly marked as "required" or "optional", no exception to this clause is allowed. When a field is marked as optional, empty values are allowed and indicate that the information is not available.

1.4) Non-compliance

Non-compliant files should be totally rejected, in order not to "poison" the shared project database

1.5) Verification

XML files should be checked against a supplied XSD (XML Schema) or DTD

2) Formats

2.1) Dates

Dates should be stored in ISO8601 "short" notation, with hyphens (e.g. January 31st, 2009 is "2009-01-31"). Existing applications are responsible for converting dates in this format to/from their internal standard.

2.2) Boolean values

Boolean values express a "true" condition when containing one of either "Yes", "On", "True" or "1"; "false" applies to "No", "Off", "False" or 0. Other values are not supported and are a cause for rejection. Strings such as "On" or "Off", when used in the context of boolean values, are to be considered case-insensitive.

2.3) External references

Countries are specified with their ISO Alpha-3 code. Languages should be encoded in ISO-639-1 or ISO-639-2 standard¹

Currencies should be identified by their ISO-4217 code

Airports, where indicated, should be identified by their ICAO code (e.g. FCO = Rome Fiumicino, LAX = Los Angeles, etc.).

A list of admissible "work types" should be compiled and made available.

3) Data fields/elements

Here follows a list of the data fields/elements in the format. Use the attached "alliance.xml" file for reference. Please note: the element names are just indicative and, with regard to this draft, a matter of personal preference. Elements marked with an asterisk (*) are required.

exportfile *

This is the root element, it defines as attributes: a *network* for data sharing (always "alliance" or whatever); a format *version* (1.0 in the current draft); an update date (*lastupdate*) to properly handle updates and/or incremental revisions. Can occur only once in a file.

workcamps *

All projects/workcamps organized/managed by a single organization are grouped under a "workcamps" element, which has a *organization* attribute specifying the code² of the organization itself. Can occur only once in a file.

workcamp *

This element comprises all the data fields pertaining to a single project. An indefinite number of *workcamp* elements is grouped under a *workcamps* element. Must occur at least once in a *workcamps* section.

code *

The code which uniquely identifies the camp in the Alliance project database. No camps sharing the same code, even if managed by different organizations, shall exist. While project codes formatting is left at the discretion of each organization, particular care should be taken in ensuring the consistency of such codes. For instance, given an organization whose code is ABC, codes could/should be "ABC-01", "ABC-02", "ABC-03", etc. but not "ABC/01", "ABC-02", "ABC 3".

¹ See note 3

² It should be discussed whether a list of organizations with updated contact informations (e-mail addresses, names of placement officers, etc.) should be included in the format or externally managed

Codes which are duplicates of other codes already present in the database, are a cause for rejection of the file.

work *

The work type of the project/workcamp³

start_date *

The starting date of the project

end-date *

The ending date of the project

name *

The name/title for the project

location *

The place where the project/workcamp will take place

country *

The country where the project/workcamp will take place

region

The region of a country where the project/workcamp will take place

languages *

A comma-separated list of languages requested for the project⁴

extrafee

The extrafee element specify the amount of the (possible) extra-fee for participation. Currency is expressed as an attribute and is mandatory. The value is an integer, no decimals allowed

min_age

This element defines the minimum age required for participation in the camp/project. When omitted, the standard Alliance rules apply. The value should be specified if the partner's rules concerning age (or the rules applicable to this specific project) don't follow the usual Alliance practices.

max_age

This element defines the maximum age allowed for participation in the camp/project. When omitted, the standard Alliance rules apply. The value should be specified if the partner's rules concerning age (or the rules applicable to this specific project) don't follow the usual Alliance practices.

disabled_vols

This elements indicates if the project/workcamp type/work/location allows for the participation of disabled volunteers. Notes should be used to give additional information. Boolean field

numvol *

Number of total (males + females) places for volunteers

³ An "official" list of work types should be compiled and available for reference

⁴ It should be discussed whether English, being the default "preferred" language, should be included in this list or not

vegetarian

Indicates if vegetarian food will be available.⁵ Boolean field

family

Indicates if whole families, with kids, are allowed. Boolean field

description *

A medium sized text with a description of the project

airport

The nearest airport⁶

train_station

The nearest train station

numvol_m

Number of places available for male volunteers

numvol_f

Number of places available for female volunteers

max_vols_per_country

Maximum number of volunteers per nationality

max_teenagers

Maximum number of teenager volunteers

max_national_vols

Maximum number of national (local) volunteers

notes

Anything that doesn't fit in the fields listed above

4) Recommendations

4.1) Data length

While the format mustn't make any assumption about the space available, in the underlying database system, for those data fields which tend to be of a certain length (project descriptions, notes, etc.) a recommendation not to exceed a certain length should be considered.

4.2) Language

All the informations contained in the import/export file should be in English or in the working language/s Alliance will choose.

⁵ It could also be interpreted as "Meals will be strictly vegetarian"; to be discussed furtherly

⁶ Is this to be interpreted as a geo-localization feature or as a arrival/meeting place for the project?

Third revision - Changes from second revision (Dec. 22, 2008)

- Added *numvol_m*, *numvol_f*, *max_vols_per_country*, *max_teenagers*, *max_national_vols*

Second revision - Changes from first draft (Nov. 20, 2008)

- Added a paragraph about CSV and text versions of the import/export file
- Dates are in ISO8601 format but with hyphens (not really important for XML, but to make them consistent with CSV export)
- Country codes are in ISO Alpha-3 format
- "Yes/No" values added as valid for boolean field elements
- More specific indication of boolean field elements
- Currency for extrafee must be an integer value, no decimals. Currency code is mandatory
- Non complying data files are rejected as a whole, no record should be imported or exported
- The *code* field was moved from being an attribute of the *workcamp* element to being an element of its own
- A recommendation regarding project codes formatting was moved outside of the recommendation to the main text
- The *age* element was split into two different elements: *min_age*, and *max_age*, both optional