External timesheet format

Thomas HOULLIER ⟨pro@houllier.net⟩

PRJ1-IRS1-v1.1 – November 02, 2024

Abstract

This document describes the external timesheet format used by the timesheeting program to export timesheet data.

Revision History

Revision	Date	Author(s)	Description
1.0	26OCT2024	TH	Creation
1.1	02NOV2024	TH	Added: IRS1-HED-075, IRS1-HED-085, IRS1-HED-150, IRS1-BDY-090, IRS1-BDY-100. Modified: IRS1-HED-150. Example file updated accordingly.

Reference documents

Index	Title	Reference	Revision	Author
RD1	timesheeting specification document	PRJ1-SPE1	v1.0	Thomas HOULLIER
RD2	List of tz database time zones	[2]	1246415064	Wikipedia

Document distribution

The present document is distributed under the *Creative Commons Attribution 4.0 International* license (https://creativecommons.org/licenses/by/4.0/) by its author Thomas HOULLIER.

Every document release is signed with the author's GPG key. A signature file is provided along with the released document.

Contents

Acronyms

CLI Command line interface
CSV Comma-separated values
DB Database
UTC Coordinated universal time

1 Introduction

1.1 Context

In support of the project requirement R-DEX-010 [RD1], we provide a format for timesheet data export to a file.

The format is meant to be interoperable, *ie* it is meant to be easily usable with a wide selection of external programs. The external programs typically targetted are spreadsheet programs (*Libreoffice Calc*, *Gnumeric*), Command line interface (CLI) text programs (*less*, *vi*), and Python libraries such as *pandas*.

We prioritize ease of use and readability for the exported format over compactness and efficiency.

1.2 Document structure

The document is structured as follows. First, definitions are given (Section 2), then the requirements for the exported file are listed (Section 3) and finally an example of a compliant export file is given (Section 4).

2 Definitions

The file refers to the external timesheet data file generated by the timesheeting program's export function.

3 Requirements

The requirements for the timesheet exported file format are listed here. They are split into general, header and body sections.

3.1 General

The general requirements apply to the *file* as a whole.

IRS1-GEN-010 – File type   The *file* shall be a text file.

IRS1-GEN-020 – File extension   The *file* extension shall be `.csv`.

IRS1-GEN-030 – File encoding   The *file* shall be encoded in `UTF-8`.

IRS1-GEN-040 – Line endings   The *file* shall use UNIX line endings (LF).

IRS1-GEN-050 – File ending   The *file* shall end with a newline character.

IRS1-GEN-060 – File structure   The *file* shall contain

- a header,
- a body.

IRS1-GEN-070 – Header location   The header of the *file* shall be a contiguous set of text lines at the beginning of the *file*.

IRS1-GEN-080 – Body location   The body of the *file* shall be a contiguous set of text lines immediately following the header lines (*ie* without gap or empty lines in between).

3.2 Header

The header requirements apply to the header part of the *file*.

IRS1-HED-010 – Header format   Every line in the header shall begin with a `#` character followed by a whitespace.

IRS1-HED-020 – Header export date   The header shall contain the date at which the *file* was generated.
    Note the time reference is the local system clock, as is the case in the rest of the program.

IRS1-HED-030 – Period start date   The header shall contain the start date for the time period specified during the export.

IRS1-HED-040 – Period stop date   The header shall contain the stop date for the time period specified during the export.

IRS1-HED-050 – Header dates timezone   The dates in the header shall be expressed in the timezone set in the program at the time of export.

IRS1-HED-060 – Header timezone   The timezone used to generate the dates in the header shall be indicated with a TZ identifier string (*eg* `Europe/Paris`). The list of current TZ identifiers may be found at [RD2].

IRS1-HED-070 – Header program version   The version of the program at the time of export shall be written in the header.

IRS1-HED-075 – Header database version   The Database (DB) version of the program at the time of export shall be written in the header.

IRS1-HED-080 – Program version string   The program version (R-REL-010 [AD1]) shall be indicated by a string `XX.YY`. With,

- `XX`: The major program version (*eg* `03`),

- `YY`: The minor program version (*eg* `26`).

A dot character separates the major version and minor version. Note the program version string has a fixed size.

IRS1-HED-085 – Database version string   The program DB version shall be indicated by an integer string of variable length. An example is `23`.

IRS1-HED-090 – Header date format   The dates in the header shall conform to the format `DDMMMYYYY HH:MM:SS`. Where,

- `DD` is the number of the day of the month (*eg* `09`),

- `MMM` is the abbreviated month name, with first letter capitalized and remaining letters in lower case (*eg* `Oct`),

- `YYYY` is the year number (*eg* `2024`),

- `HH` is the hours number (*eg* `07`),

- `MM` is the minutes number (*eg* `39`)

A `:` separator is present between hours and minutes, and between minutes and seconds.

Note the corresponding `strftime` [1] format string for the date format is `%d%b%Y %H:%M:%S`. Note the date is always represented with a fixed number of characters.

IRS1-HED-100 – Export date format   The export date shall be written on a single text line beginning with `Export date:` followed by a whitespace and the export date formatted according to IRS1-HED-090.

An example line is,

```
# Export date: 26Oct2024 14:52:28
```

IRS1-HED-110 – Period start date format   The export period start date shall be written on a single text line beginning with `Export start date:` followed by a whitespace and the period start date formatted according to IRS1-HED-090.

An example line is,

```
# Export start date: 01Jan2024 00:00:00
```

IRS1-HED-120 – Period stop date format   The export period stop date shall be written on a single text line beginning with `Export stop date:` followed by a whitespace and the period stop date formatted according to IRS1-HED-090.

An example line is,

```
# Export stop date: 31Dec2024 23:59:59
```

IRS1-HED-130 – Header timezone format   The header timezone shall be written on a single text line beginning with `Header timezone:` followed by a whitespace and the TZ identifier string (IRS1-HED-060).

An example line is,

```
# Header timezone: Europe/Paris
```

IRS1-HED-140 – Program version format   The program version shall be written in the header on a single text line beginning with `timesheeting version:` followed by a whitespace and the program version string (IRS1-HED-080).

An example line is,

```
# timesheeting version: 03.26
```

IRS1-HED-150 – Header ordering   The elements of the header shall be ordered as follows,

1. Export date (IRS1-HED-020),

2. Period start date (IRS1-HED-030),

3. Period stop date (IRS1-HED-040),

4. Header timezone (IRS1-HED-060),

5. Program version (IRS1-HED-070),

6. Program DB version (IRS1-HED-075).

3.3 Body

IRS1-BDY-010 – CSV format   The body of the *file* shall be in Comma-separated values (CSV) format.

IRS1-BDY-020 – CSV delimiter   The delimiter character used by the CSV format shall be a comma followed by a whitespace (,).

IRS1-BDY-030 – Body structure   The body of the *file* shall contain,

- A line of column names at the top,
- Timesheet data entries.

IRS1-BDY-040 – Column list   The body CSV columns shall be, in order,

1. Entry ID,
2. Project ID,
3. Project name,
4. Task ID,
5. Task name,
6. Location ID,
7. Location name,
8. Start date,
9. Stop date.

IRS1-BDY-050 – No empty fields   The timesheet entries in the *file* body shall not contain any empty fields.

IRS1-BDY-060 – Id format   The ID fields in timesheet entries shall be represented as a number string of variable length.

IRS1-BDY-070 – Timesheet date format   The date fields in timesheet entries shall be represented as a Coordinated universal time (UTC) UNIX timestamp in seconds.

IRS1-BDY-080 – Timesheet name format   The name fields in timesheet entries shall be represented as a string which may contain whitespace.

IRS1-BDY-090 – Timesheet entries ordering   The timesheet entries in the *file* body shall be ordered by increasing entry start date.

IRS1-BDY-100 – Timesheet entries period The timesheet entries in the *file* body shall have a start date chronologically contained within the period specified by the header start date and stop date.

4 Example file

We provide an example of compliant file format in Listing 1.

```
 1 # Export date: 26Oct2024 14:52:28
 2 # Export start date: 01Jan2024 00:00:00
 3 # Export stop date: 31Dec2024 23:59:59
 4 # Header timezone: Europe/Paris
 5 # timesheeting version: 03.26
 6 # timesheeting DB version: 23
 7 Entry ID, Project ID, Project name, Task ID, Task name, Location ID
     , Location name, Start date, Stop date
 8 4, 1, Project1, 3, Task3, 1, Location1, 1729952454, 1729953654
 9 8, 1, Project1, 11, Task11, 1, Location1, 1729953659, 1729953789
10 9, 15, Project15, 5, Task5, 3, Location3, 1729953888, 1729953988
```

Listing 1: Compliant exported timesheet file.

References

[1] CPP Reference. *std::strftime*. [Online; accessed 26-October-2024]. 2024. URL: https://en.cppreference.com/mwiki/index.php?title=cpp/chrono/c/strftime&oldid=161518.

[2] Wikipedia contributors. *List of tz database time zones — Wikipedia, The Free Encyclopedia*. [Online; accessed 26-October-2024]. 2024. URL: https://en.wikipedia.org/w/index.php?title=List_of_tz_database_time_zones&oldid=1246415064.