

## White Paper

### Language Tags

---

#### Preliminary Remark

Usually openfunds fields are available in one language only. Most commonly they are in English. As there are few exceptions to this rule, the core of openfunds does not differentiate between different languages used in a specific field, i.e. by assigning different OF-IDs to the same field depending on the contents' language.

However, there are fields in openfunds that may exist in multiple languages. For example "Fund Name" (OFST010110 Fund Name), "Investment Objective" (OFST010300 Investment Objective), diverse company or organisation names as well as several other fields. Some of these fields, for example one describing a company or organisation name, may only occur in a few foreign languages and scripts, while others, for example "Investment Objective" could possibly exist in one or two dozen languages.

To avoid having to create new OF-IDs for each language version of an existing field, openfunds has adopted a language tag extension that fulfils three criteria:

**1) The OF-ID of each field in openfunds should be language independent**

This ensures that new language versions of any field can be introduced without the openfunds committee having to assign a new OF-ID.

**2) The system should allow a default language while permitting levels of granulation, for example American English, British English, etc.**

In this way, two aims are achieved. First it allows the handling of language variants and secondly it guarantees that database systems without language tags implemented can communicate with database systems that use language tags.

**3) Commonly used language codes should be implemented**

By using a commonly accepted language code convention, additional languages can be introduced without involving the openfunds committee.

A solution that allows a specific field to be populated by content in different languages without having to assign a new OF-ID must also include an extension of the OF-ID. This extended OF-ID option includes a "language tag" in parenthesis at the end of the OF-ID number.

In general, the extended OF-ID has the following format:

OFXXnnnnnn(**lang**)[rrrr]

- „OF“ stands for openfunds,
- XX for the values „ST“ for „static“, „AS“ for „Assets“ und „DY“ for „dynamic“ and nnnnnn represents a 6 digit numerical code,
- (lang) indicates the language based on the explanation below (Language Tag Foundation), and
- [rrrr] represents an optional recipient code

# openfunds

## Language Tag Foundation

The foundation of the openfunds language tag convention is built on the IETF language tag. IETF stands for „Internet Engineering Task Force“, a voluntary organization that develops and promotes Internet Standards. The IETF publishes “best current practice” (BCP) Papers, which are based on “Request for Comments” (RFC) articles. The IETF also integrates existing standards such as those published by ISO.

The following excerpts are based on article [RFC 5646](#) of the BCP Series 47 (as of January 2016). At its core, this article proposes the following structure, which for our purposes only the main four groups are presented. (Compare also [https://en.wikipedia.org/wiki/IETF\\_language\\_tag](https://en.wikipedia.org/wiki/IETF_language_tag)):

OFXXnnnnnn(pp-eee-Ssss-RR)[rrrr], whereby the four language tag-groups are defined as follows:

- pp: primary language subtag (two characters; ISO 639-1), e.g. en, de
- eee: extended language subtags (three characters; ISO 639-3), e.g. swg, cmn
- Ssss: script subtag (four characters, first letter capitalized; ISO 15924), e.g. Hans, Hant
- RR: region subtag (two characters, all letters capitalized, ISO 3166-1), e.g. US, CH

## Recommendation for implementation in openfunds

A full integration of the IETF language tag syntax proposal can be ruled out due to the relatively few languages and scripts that would actually be used, as well as the small number of affected fields. Therefore, the openfunds committee proposes the following convention be used in openfunds version 1.0:

### Only one or two language tag groups

IETF BCP47 provides the possibility to combine language group tags. The openfunds committee therefore recommends to implement either of these two formats:

#### (pp) or (pp-Ssss)

Thus individual fields can be easily adapted to accommodate major languages such as English (en), French (fr), Spanish (es) or German (de) by the use of the appropriate two character language tag.

In a database, the OF-ID OFST010300(en) would be easily recognizable as „Investment Objective“ in English (OFST010300 is the openfunds OF-ID for Investment Objective). For the Investment Objective in Spanish, the OF-ID would be OFST010300(es).

In the case of script variations of a base language, for example Chinese where you may have Traditional Chinese or Simplified Chinese, the language tag form (pp-Ssss) would be used. The Investment Objective in Simplified Chinese is represented as OF-ID OFST010300(zh-Hans), while the OF-ID OFST010300(zh-Hant) corresponds to Traditional Chinese.

Of course, everyone is free to implement openfunds in such a way that all possible combinations of the four most commonly used language tag groups are present. This, however, does not yet appear necessary from today’s perspective.

For maximum compatibility, openfunds also allows the complete omission of language tags. In this case fields should only contain content defined by the fund house. To facilitate the successful transfer of fund data from systems that implement language tags to systems that do not, openfunds recommends the following procedure if a field includes a language tag in either (pp) or (pp-Ssss) Format:

Store the field at least twice in sequence using the following convention:

- The first time without using a language tag with the field containing content in the language based on the fund houses’ choice.

# openfunds

- A second time with the same content in the same language, including a language tag corresponding to the language used.
- A third or more time(s) depending on additional language variations of the field.

## Language tags are only relevant for a limited number of data fields

As mentioned before, the use of language tags only makes sense for some fields. Therefore, openfunds lists the following fields where an optional language tag should be considered (see below). This list is not comprehensive, and will likely be expanded in the future. Additionally, in the corresponding field description that can be found in the [Fields section](#), it is mentioned when a language tag option should be used.

A list of allowed languages is not included in this whitepaper as experience has shown that such a list continuously expands with time.

## Fields with language tag functionality

OFST001000 Fund Group Name  
OFST001020 ManCo  
OFST001100 Fund Promotor Name  
OFST001300 Fund Administrator Name  
OFST001400 Custodian Bank Name  
OFST001450 Portfolio Managing Company Name  
OFST001500 Fund Advisor Name  
OFST001600 Auditor Name  
OFST001900 Collateral Manager Name  
OFST002000 Marketmaker Name  
OFST002700 Transfer Agent Name  
OFST005010 Umbrella  
OFST010020 Legal Fund Name  
OFST010110 Fund Name  
OFST010300 Investment Objective  
OFST020060 Share Class Name

## Character Encoding

For maximum compatibility, UTF-8 (8-bit Universal Character Set Transformation Format) should be used.

## Joining openfunds

If your firm has a need to reliably send or receive fund data, you are more than welcome to use the openfunds fields and definitions free-of-charge. Interested parties can contact the openfunds association by sending an email to: [businessoffice@openfunds.org](mailto:businessoffice@openfunds.org)

### **openfunds.org**

Staffelstrasse 12

CH-8045 Zürich

Switzerland

Tel.: +41 44 286 80 20

Email: [businessoffice@openfunds.org](mailto:businessoffice@openfunds.org)

Website: <https://www.openfunds.org>