

API Reference

Harmony API

API Version: 2.0

Documentation for Harmony API.

Harmony is a tool using AI which allows you to compare items from questionnaires and identify similar content. You can try Harmony at app.harmonydata.org and you can read our blog at harmonydata.org/blog/.

CONTACT

NAME: Thomas Wood

URL: <https://fastdatascience.com>

INDEX

1. HEALTH CHECK	3
1.1 GET /health-check	3
2. INFO	4
2.1 GET /info/version	4
3. TEXT	5
3.1 POST /text/parse	5
3.2 POST /text/match	6
3.3 POST /text/examples	8
3.4 GET /text/cache	9

API

1. HEALTH CHECK

1.1 GET /health-check

Health Check

REQUEST

No request parameters

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

undefined

2. INFO

2.1 GET /info/version

Show Version

REQUEST

No request parameters

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

undefined

3. TEXT

3.1 POST /text/parse

Parse Instruments

Parse PDFs or Exceles or text files into Instruments, and identifies the language.

If the file is binary (Excel or PDF), you must supply each file with the content in MIME format and the bytes in base64 encoding, like the example RawFile in the schema.

If the file is plain text, supply the file content as a standard string.

REQUEST

REQUEST BODY - application/json

```
[{  
  Array of object:  
  file_id      string      Unique identifier for the file (UUID-4)  
  file_name    string      DEFAULT:Untitled file  
  file_type*   enum        The name of the input file  
  file_type*   enum        ALLOWED:pdf, xlsx, txt, docx  
  file_type*   enum        The file type (pdf, xlsx, txt)  
  content*     string      The raw file contents  
  text_content string      The plain text content  
  tables       [undefined]  
}]
```

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

```
[{  
  Array of object:  
  file_id      string      Unique identifier for the file (UUID-4)  
  instrument_id string      Unique identifier for the instrument (UUID-4)  
  instrument_name string     DEFAULT:Untitled instrument  
  instrument_name string     Human-readable name of the instrument  
  file_name    string      DEFAULT:Untitled file  
  file_name    string      The name of the input file  
  file_type    enum        ALLOWED:pdf, xlsx, txt, docx  
  file_type    enum        The file type (pdf, xlsx, txt)  
  file_section string      The sub-section of the file, e.g. Excel tab  
  study       string      The study  
  sweep       string      The sweep  
  metadata {  
    Optional metadata about the instrument (URL, citation, DOI, copyright holder)  
  }  
  language     enum        DEFAULT:en  
  language     enum        ALLOWED:de, el, en, es, fr, it, he, ja, ko, pt, ru,  
  language     enum        uk, zh, ar, la, tr, af, ak, am, as, ay, az, be, bg,  
  language     enum        bho, bm, bn, bs, ca, ceb, ckb, co, cs, cy, da, doi,  
  language     enum        dv, ee, eo, et, eu, fa, fi, fil, fy, ga, gd, gl, gn,  
  language     enum        gom, gu, ha, haw, hi, hmn, hr, ht, hu, hy, id, ig,  
  language     enum        ilo, is, jv, ka, kk, km, kn, kri, ku, ky, lb, lg,
```

ln, lo, lt, lus, lv, mai, mg, mi, mk, ml, mn, mni-
mtei, mr, ms, mt, my, ne, nl, no, nso, ny, om, or,
pa, pl, ps, qu, ro, rw, sa, sd, si, sk, sl, sm, sn,
so, sq, sr, st, su, sv, sw, ta, te, tg, th, ti, tk,
tl, ts, tt, ug, ur, uz, vi, xh, yi, yo, zh-tw, zu,
yue

The ISO 639-2 (alpha-2) encoding of the instrument language

```
questions* [{  
  Array of object:  
    question_no      string      Number of the question  
    question_intro   string      Introductory text applying to the question  
    question_text*   string      Text of the question  
    options           [string]  
    source_page      integer      DEFAULT:0  
                                The page of the PDF on which the question was located, zero-indexed  
    instrument_id    string      Unique identifier for the instrument (UUID-4)  
    instrument_name  string      Human readable name for the instrument  
    topics_auto      [undefined]  
    nearest_match_from_mhc_auto {  
      Automatically identified nearest MHC match  
    }  
  }  
}]  
}]
```

STATUS CODE - 422: Validation Error

RESPONSE MODEL - application/json

```
{  
  detail [{  
    Array of object:  
      loc*  
        ANY OF  
        prop0  
          string  
        prop1  
          integer  
      msg* string  
      type* string  
    }  
  }  
}
```

3.2 POST /text/match

Match

Match instruments

REQUEST

REQUEST BODY - application/json

```
{  
  instruments* [{  
    Array of object:  
      file_id        string      Unique identifier for the file (UUID-4)  
      instrument_id  string      Unique identifier for the instrument (UUID-4)  
      instrument_name string      DEFAULT:Untitled instrument  
                                Human-readable name of the instrument  
      file_name      string      DEFAULT:Untitled file
```

```

file_type          enum          The name of the input file
                                ALLOWED:pdf, xlsx, txt, docx
                                The file type (pdf, xlsx, txt)
file_section       string        The sub-section of the file, e.g. Excel tab
study              string        The study
sweep              string        The sweep
metadata {
Optional metadata about the instrument (URL, citation, DOI, copyright holder)
}
language           enum          DEFAULT:en
                                ALLOWED:de, el, en, es, fr, it, he, ja, ko, pt, ru, uk,
                                zh, ar, la, tr, af, ak, am, as, ay, az, be, bg, bho, bm,
                                bn, bs, ca, ceb, ckb, co, cs, cy, da, doi, dv, ee, eo,
                                et, eu, fa, fi, fil, fy, ga, gd, gl, gn, gom, gu, ha,
                                haw, hi, hmn, hr, ht, hu, hy, id, ig, ilo, is, jv, ka,
                                kk, km, kn, kri, ku, ky, lb, lg, ln, lo, lt, lus, lv,
                                mai, mg, mi, mk, ml, mn, mni-mtei, mr, ms, mt, my, ne,
                                nl, no, nso, ny, om, or, pa, pl, ps, qu, ro, rw, sa, sd,
                                si, sk, sl, sm, sn, so, sq, sr, st, su, sv, sw, ta, te,
                                tg, th, ti, tk, tl, ts, tt, ug, ur, uz, vi, xh, yi, yo,
                                zh-tw, zu, yue
                                The ISO 639-2 (alpha-2) encoding of the instrument language

questions* [{
Array of object:
  question_no      string        Number of the question
  question_intro   string        Introductory text applying to the question
  question_text*   string        Text of the question
  options          [string]
  source_page      integer       DEFAULT:0
                                The page of the PDF on which the question was located, zero-indexed
  instrument_id    string        Unique identifier for the instrument (UUID-4)
  instrument_name  string        Human readable name for the instrument
  topics_auto      [undefined]
  nearest_match_from_mhc_auto {
Automatically identified nearest MHC match
  }
}]
}
query              string        Search term
parameters {
Parameters on how to match
  framework string DEFAULT:huggingface
                                The framework to use for matching
  model           string DEFAULT:sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2
                                Model
}
}

```

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

```

{
  questions* [{
Array of object:
    question_no      string        Number of the question
    question_intro   string        Introductory text applying to the question

```

```

question_text*  string      Text of the question
options        [string]
source_page    integer      DEFAULT:0
                                     The page of the PDF on which the question was located, zero-indexed
instrument_id   string      Unique identifier for the instrument (UUID-4)
instrument_name string      Human readable name for the instrument
topics_auto    [undefined]
nearest_match_from_mhc_auto {
  Automatically identified nearest MHC match
}
}]
matches* [ {
  Array of object:
} ]
query_similarity [undefined]
}

```

STATUS CODE - 422: Validation Error

RESPONSE MODEL - application/json

```

{
  detail [ {
    Array of object:
    loc*
      ANY OF
      prop0
        string
      prop1
        integer
    msg*  string
    type* string
  } ]
}

```

3.3 POST /text/examples

Get Example Instruments

Get example instruments

REQUEST

No request parameters

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

```

[ {
  Array of object:
  file_id          string      Unique identifier for the file (UUID-4)
  instrument_id    string      Unique identifier for the instrument (UUID-4)
  instrument_name  string      DEFAULT:Untitled instrument
                                     Human-readable name of the instrument
  file_name        string      DEFAULT:Untitled file
                                     The name of the input file
  file_type        enum        ALLOWED:pdf, xlsx, txt, docx
                                     The file type (pdf, xlsx, txt)
} ]

```



```

file_section      string      The sub-section of the file, e.g. Excel tab
study            string      The study
sweep           string      The sweep
metadata {
Optional metadata about the instrument (URL, citation, DOI, copyright holder)
}
language         enum        DEFAULT:en
                                ALLOWED:de, el, en, es, fr, it, he, ja, ko, pt, ru, uk, zh,
                                ar, la, tr, af, ak, am, as, ay, az, be, bg, bho, bm, bn,
                                bs, ca, ceb, ckb, co, cs, cy, da, doi, dv, ee, eo, et,
                                eu, fa, fi, fil, fy, ga, gd, gl, gn, gom, gu, ha, haw,
                                hi, hmn, hr, ht, hu, hy, id, ig, ilo, is, jv, ka, kk, km,
                                kn, kri, ku, ky, lb, lg, ln, lo, lt, lus, lv, mai, mg,
                                mi, mk, ml, mn, mni-mtei, mr, ms, mt, my, ne, nl, no,
                                nso, ny, om, or, pa, pl, ps, qu, ro, rw, sa, sd, si, sk,
                                sl, sm, sn, so, sq, sr, st, su, sv, sw, ta, te, tg, th,
                                ti, tk, tl, ts, tt, ug, ur, uz, vi, xh, yi, yo, zh-tw,
                                zu, yue
                                The ISO 639-2 (alpha-2) encoding of the instrument language

questions* [{
Array of object:
  question_no      string      Number of the question
  question_intro   string      Introductory text applying to the question
  question_text*   string      Text of the question
  options          [string]
  source_page      integer     DEFAULT:0
                                The page of the PDF on which the question was located, zero-indexed
  instrument_id    string      Unique identifier for the instrument (UUID-4)
  instrument_name  string      Human readable name for the instrument
  topics_auto      [undefined]
  nearest_match_from_mhc_auto {
Automatically identified nearest MHC match
  }
}]
}]

```

3.4 GET /text/cache

Get Cache

Get all items in cache

REQUEST

No request parameters

RESPONSE

STATUS CODE - 200: Successful Response

RESPONSE MODEL - application/json

```

{
  instruments* [{
Array of object:
    file_id        string      Unique identifier for the file (UUID-4)
    instrument_id  string      Unique identifier for the instrument (UUID-4)
    instrument_name string     DEFAULT:Untitled instrument
                                Human-readable name of the instrument
  }
}

```

```

file_name      string      DEFAULT:Untitled file
                The name of the input file
file_type      enum        ALLOWED:pdf, xlsx, txt, docx
                The file type (pdf, xlsx, txt)
file_section   string      The sub-section of the file, e.g. Excel tab
study         string      The study
sweep         string      The sweep
metadata {
Optional metadata about the instrument (URL, citation, DOI, copyright holder)
}
language       enum        DEFAULT:en
                ALLOWED:de, el, en, es, fr, it, he, ja, ko, pt, ru, uk,
                zh, ar, la, tr, af, ak, am, as, ay, az, be, bg, bho,
                bm, bn, bs, ca, ceb, ckb, co, cs, cy, da, doi, dv, ee,
                eo, et, eu, fa, fi, fil, fy, ga, gd, gl, gn, gom, gu,
                ha, haw, hi, hmn, hr, ht, hu, hy, id, ig, ilo, is, jv,
                ka, kk, km, kn, kri, ku, ky, lb, lg, ln, lo, lt, lus,
                lv, mai, mg, mi, mk, ml, mn, mni-mtei, mr, ms, mt, my,
                ne, nl, no, nso, ny, om, or, pa, pl, ps, qu, ro, rw,
                sa, sd, si, sk, sl, sm, sn, so, sq, sr, st, su, sv,
                sw, ta, te, tg, th, ti, tk, tl, ts, tt, ug, ur, uz,
                vi, xh, yi, yo, zh-tw, zu, yue
                The ISO 639-2 (alpha-2) encoding of the instrument language

questions* [{
Array of object:
    question_no      string      Number of the question
    question_intro   string      Introductory text applying to the question
    question_text*   string      Text of the question
    options           [string]
    source_page       integer     DEFAULT:0
                                The page of the PDF on which the question was located, zero-indexed
    instrument_id     string      Unique identifier for the instrument (UUID-4)
    instrument_name   string      Human readable name for the instrument
    topics_auto       [undefined]
    nearest_match_from_mhc_auto {
Automatically identified nearest MHC match
    }
}
}]
vectors* [{
Array of object:
}
}
}

```

