



KASABI[®]

The Kasabi Information Marketplace

knud.moeller@talis.com

@knudmoeller

19/04/2012, WWW2012, Lyon, France





A Place to...

- **publish** data
- **integrate** your data
- **monetize** your data

- **find** data
- **consume and use** data



- **web-based** platform
- **horizontal** market place
- RESTful APIs
- language bindings (Ruby, PHP, JS, Python)
- **pytassium**

Category

Government (53)

Geography (44)

Uncategorised (37)

Travel (28)

Media (21)

Publishing (19)

Linking (17)

Education (16)

Science (14)

Health (13)

Music (12)

Business (10)

Commerce (8)

Food (3)



What's so special?

- Kasabi is based on **linked data** principles
 - data in **graph** structure (RDF)
 - **URIs** identify data items
 - data **links** to other datasets (context)
 - linked data **views**



What's so special?

- Your data gets **APIs**
 - SPARQL endpoint
 - keyword search
 - lookup
 - reconciliation
 - custom APIs



Dashboard

My Dashboard | Kasabi

http://kasabi.com/dashboard

KASABI BETA

Browse **Knud Möller** Logout

User Dashboard

My Dashboard

Summary | My Datasets | My Profile

Activity Subscription plan [Free Bee](#)

5th April 2012 - 19th April 2012

API key usage Total credits used **6** [View my usage report](#)

My Datasets Total requests **15** [View my usage report](#)

11 datasets with access granted
My API Key - [REDACTED]

0 users
on 3 datasets

My most accessed datasets

Dataset	Credits
world-geography	3
Dbpedia	2
Food	1

My datasets

Dataset	Requests
Languages	15

[About Kasabi](#) [Getting Help](#) [Kasabi Standards](#) [Keep in touch](#)

Feedback



Creating a dataset

Create Dataset | Kasabi

http://kasabi.com/dataset/www2012/edit


KASABI BETA

Browse Knud Möller Logout

Edit

Create Dataset

Logo:



Remove

Name: *

WWW2012

Description:

Metadata for the 21st World Wide Web Conference, A

Web page addresses and e-mail addresses turn into Lines and paragraphs break automatically.
Allowed HTML tags: <a> <cite> <code> <th> <tr> <td> <p> <h1> <h2> <h3> <h4>
Use [gist:####] where #### is your gist number to embed the gist.

More information about formatting options

- Publishing
- Uncategorized
- Science
- Education
- Linking

Licence Agreement: *

Creative Commons CC0

Example Resource:

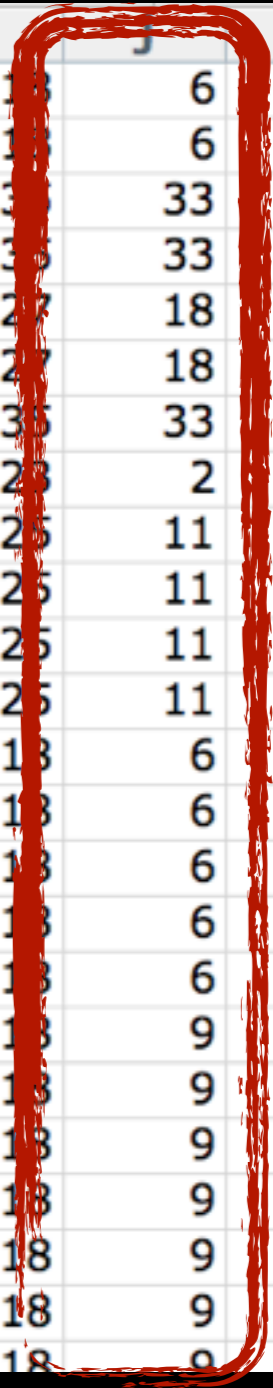
-
-

Add another item



CSV2RDF Conversion

◇	C	D	E	F	G	H	I	J	K
1	FR-R92B-58	Black	1059.31	1431.5	58	1016.04	18	6	00:00.0
2	FR-R92R-58	Red	1059.31	1431.5	58	1016.04	18	6	00:00.0
3	HL-U509-R	Red	13.0863	34.99			35	33	00:00.0
4	HL-U509	Black	13.0863	34.99			35	33	00:00.0
5	SO-B909-M	White	3.3963	9.5	M		27	18	00:00.0
6	SO-B909-L	White	3.3963	9.5	L		27	18	00:00.0
7	HL-U509-B	Blue	13.0863	34.99			35	33	00:00.0
8	CA-1098	Multi	6.9223	8.99			23	2	00:00.0
9	LJ-0192-S	Multi	38.4923	49.99	S		25	11	00:00.0
10	LJ-0192-M	Multi	38.4923	49.99	M		25	11	00:00.0
11	LJ-0192-L	Multi	38.4923	49.99	L		25	11	00:00.0
12	LJ-0192-X	Multi	38.4923	49.99	XL		25	11	00:00.0
13	FR-R92R-62	Red	868.6342	1431.5	62	1043.26	18	6	00:00.0
14	FR-R92R-44	Red	868.6342	1431.5	44	961.61	18	6	00:00.0
15	FR-R92R-48	Red	868.6342	1431.5	48	979.75	18	6	00:00.0
16	FR-R92R-52	Red	868.6342	1431.5	52	997.9	18	6	00:00.0
17	FR-R92R-56	Red	868.6342	1431.5	56	1016.04	18	6	00:00.0
18	FR-R38B-58	Black	204.6251	337.22	58	1115.83	18	9	00:00.0
19	FR-R38B-60	Black	204.6251	337.22	60	1124.9	18	9	00:00.0
20	FR-R38B-62	Black	204.6251	337.22	62	1133.98	18	9	00:00.0
21	FR-R38R-44	Red	187.1571	337.22	44	1052.33	18	9	00:00.0
22	FR-R38R-48	Red	187.1571	337.22	48	1070.47	18	9	00:00.0
23	FR-R38R-52	Red	187.1571	337.22	52	1088.62	18	9	00:00.0
24	FR-R38R-58	Red	187.1571	337.22	58	1115.83	18	9	00:00.0





CSV2RDF Conversion

```
<#weight> a :Resource ;
  :identity [
    :source_column 8 ;
    :process ( :regex ) ;
    :regex_match "^(.+)$";
    :regex_output "${1}gr";
    :base_uri "http://data.kasabi.com/dataset/adventure_works/weights/" ;
  ] ;
  :type gr:QuantitativeValueFloat ;
  :attribute
    [ :property gr:hasValue ; :source_column 8 ; :datatype xsd:float ] ,
    [ :property gr:hasUnitOfMeasurement ; :value "GRM" ; :datatype xsd:string ]
.

<#category> a :Resource ;
  :identity [
    :source_column 10 ;
    :base_uri "http://data.kasabi.com/dataset/adventure_works/product_categories/" ;
  ] ;
  :type owl:Class ;
.
```



CSV2RDF Conversion

```
gr:hasValue "1000.00"^^xsd:float ;  
a gr:QuantitativeValueFloat .  
  
<http://data.kasabi.com/dataset/adventure_works/weights/1006.97gr>  
gr:hasUnitOfMeasurement "GRM"^^xsd:string ;  
gr:hasValue "1006.97"^^xsd:float ;  
a gr:QuantitativeValueFloat .  
  
<http://data.kasabi.com/dataset/adventure_works/weights/1016.04gr>  
gr:hasUnitOfMeasurement "GRM"^^xsd:string ;  
gr:hasValue "1016.04"^^xsd:float ;  
a gr:QuantitativeValueFloat .  
  
<http://data.kasabi.com/dataset/adventure_works/weights/1025.11gr>  
gr:hasUnitOfMeasurement "GRM"^^xsd:string ;  
gr:hasValue "1025.11"^^xsd:float ;  
a gr:QuantitativeValueFloat .  
  
<http://data.kasabi.com/dataset/adventure_works/weights/1043.26gr>
```

<https://github.com/mmmmmrob/Vertere-RDF>



Datasets

WWW2012 | Kasabi

http://kasabi.com/dataset/www2012

KASABI BETA

Category - WWW2012



WWW2012

Creative Commons CC0

Science Education

Metadata for the [21st World Wide Web Conference](#), April 2012, Lyon, France, including authors, papers, organisations, etc.

More extensive conference data can be found at [SWDF](#).

Published 19th April 2012
Updated 19th April 2012
Resources 2920

  **702 Organisations (25%)**



Explore the dataset

View the [vocabularies and classes](#) used in the dataset, or see some [sample resources](#)
Read the [Developer Documentation](#) or [edit it](#)
Browse as [Linked Data](#)



Dataset Description

<http://data.kasabi.com/dataset/www2012>

WWW2012	
Property	Value
Type	Dataset
Title	WWW2012
Description	The WWW2012 dataset
Homepage	www2012
Date created	2012-04-18T23:39:47+00:00
Date modified	2012-04-19T00:35:48+00:00
Distinct subjects	2920
Status endpoint	status
Jobs endpoint	jobs
Attribution endpoint	attribution
Store endpoint	store
Example resource	1042 tom-heath

<http://xmlns.com/foaf/0.1/>



Dataset Description

<http://data.kasabi.com/dataset/www2012.ttl>

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix void: <http://rdfs.org/ns/void#> .
@prefix services: <http://labs.kasabi.com/ns/services#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

<http://data.kasabi.com/dataset/www2012> a void:Dataset ;
  dct:created "2012-04-18T23:39:47+00:00" ;
  dct:description "The WWW2012 dataset" ;
  dct:modified "2012-04-19T00:35:48+00:00" ;
  dct:title "WWW2012" ;
  void:classes "20" ;
  void:distinctSubjects "2920" ;
  void:exampleResource <http://data.semanticweb.org/conference/www/2012/paper/10
    <http://data.semanticweb.org/person/tom-heath> ;
  void:sparqlEndpoint <http://api.kasabi.com/dataset/www2012/apis/sparql> ;
  services:attributionEndpoint <http://api.kasabi.com/dataset/www2012/attributio
  services:jobsEndpoint <http://api.kasabi.com/dataset/www2012/jobs> ;
  services:reconciliationEndpoint <http://api.kasabi.com/dataset/www2012/apis/re
  services:searchEndpoint <http://api.kasabi.com/dataset/www2012/apis/search> ;
  services:statusEndpoint <http://api.kasabi.com/dataset/www2012/status> ;
  services:storeEndpoint <http://api.kasabi.com/dataset/www2012/store> ;
  void:uriLookupEndpoint <http://api.kasabi.com/dataset/www2012/apis/lookup> ;
```



Dataset Description

<http://data.kasabi.com/dataset/www2012.json>

```
    "value": "The WWW2012 dataset"
  }
],
"http://xmlns.com/foaf/0.1/homepage":
[
  {
    "type": "uri",
    "value": "http://kasabi.com/dataset/www2012"
  }
],
"http://purl.org/dc/terms/created":
[
  {
    "type": "literal",
    "value": "2012-04-18T23:39:47+00:00"
  }
],
"http://purl.org/dc/terms/modified":
[
  {
    "type": "literal",
    "value": "2012-04-19T00:35:48+00:00"
  }
],
"http://rdfs.org/ns/void#distinctSubjects":
[
```



APIs

Default APIs

Query

Search

Lookup

Reconcile

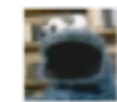
Attribute

SPARQL Endpoint

Use the SPARQL 1.1 query language to perform structured queries against a dataset. Useful for performing precise queries against a dataset whose structure you understand.

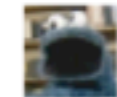
Sample queries

[Papers about "Online Communities"](#)



Knud Möller
19th Apr 2012

[All Paper Subjects](#)



Knud Möller
19th Apr 2012

Create a sample query

Contributed APIs



[Paper by Topic](#)

SPARQL Stored Procedure



Knud Möller
19th Apr 2012

Create an API



SPARQL API

http://api.kasabi.com/dataset/www2012/apis/sparql

API Key:

Output:

SPARQL Query

```
PREFIX swrc: <http://swrc.ontoware.org/ontology#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>

SELECT DISTINCT ?subject
WHERE {
  ?paper a swrc:InProceedings ;
         dc:subject ?subject .
}
ORDER BY ?subject
```

Tips

Use the SPARQL API to match patterns in the graph to extract data

Click to add prefix to query

- [Bibliographic Ontology](#)
- [Dublin Core Elements](#)
- [Dublin Core Terms](#)
- [FOAF](#)
- [W3C Geo Schema](#)
- [Music Ontology](#)
- [Programmes Ontology](#)
- [ORG](#)
- [OWL](#)
- [RDF Namespace](#)
- [RDF Schema](#)
- [Relationship](#)
- [Reviews](#)
- [RSS 1.0](#)
- [SIOC Core Ontology](#)
- [SKOS Core](#)
- [VCard](#)
- [XML Schema](#)

Read the [API Docs](#) for more information.

Request | **Response**

RESPONSE BODY

```
{
  "head": {
    "vars": [ "subject" ]
  },
  "results": {
    "bindings": [
      {
        "subject": { "type": "literal", "value": "3D Graphics" }
      }
    ]
  }
}
```



Search API

Keyword Search

http://api.kasabi.com/dataset/www2012/apis/search

API Key:

Query:

Sort criteria:

Number of results:

Offset:

Output:

Request Response

RESPONSE BODY

```
{
  "head": {
    "query": "linked",
    "startIndex": 0,
    "totalResults": 14
  },
  "results": [
    {
      "uri": "http://data.semanticweb.org/conference/www/2012/phd/26",
      "title": "From Linked Data to Linked Entities: A Migration Path",
      "score": 3.0511758
    },
    {
      "uri": "http://data.semanticweb.org/conference/www/2012/dev/17",
      "title": "LDIF - A Framework for Large-Scale Linked Data Integration",

```

Tips

Read the [API Docs](#) for more information.



Lookup API

Lookup API Explorer

Use this form to test out the [Lookup API](#) for the [WWW2012](#) dataset

<http://api.kasabi.com/dataset/www2012/apis/lookup>

API Key:

Resource URI:

<http://data.semanticweb.org/person/tom-heath>

Output:

Lookup

Tips

The Lookup API returns a description of a single resource.

The description will contain all properties of a resource, including its relationships to other resources

Read the [API Docs](#) for more information.

Request

Response

RESPONSE BODY

```
{
  "http://data.semanticweb.org/person/tom-heath": {
    "http://swrc.ontoware.org/ontology#affiliation": [
      {
        "value": "http://data.semanticweb.org/organization/talis-education-ltd",
        "type": "uri"
      },
      {
        "value": "http://data.semanticweb.org/organization/talis-information-limited",
        "type": "uri"
      }
    ],
    "http://xmlns.com/foaf/0.1/name": [
      {
        "value": "Tom Heath",
        "type": "literal"
      }
    ]
  }
}
```



Reconciliation API

Reconciliation API Explorer

Use this form to test out the [Reconciliation API](#) for the [WWW2012](#) dataset

<http://api.kasabi.com/dataset/www2012/apis/reconciliation>

API Key:

Query:

Tom Heath

Type URI:

Limit:

Reconcile

Tips

The reconciliation API allows labels and simple identifiers to be looked up in a dataset to find the URI of an item. This is useful to link together datasets.

The API is supported in Google Refine so [can be used when tidying up data using that tool](#).

Read the [API Docs](#) for more information.

Request

Response

RESPONSE BODY

```
{
  "result": [
    {
      "id": "http://data.semanticweb.org/person/tom-heath",
      "name": "Tom Heath",
      "type": [
        "http://xmlns.com/foaf/0.1/Person"
      ],
      "score": 1,
      "match": true
    }
  ]
}
```



Custom APIs

Create an API in WWW2012

You can create the following types of services within this dataset

Create SPARQL Stored Procedure

This type of API allows you to bind a SPARQL query to a URL so it will be automatically executed on request. Parameters can be passed from the query string and transformations can be applied to create custom data formats.

You'll need:

- A SPARQL query that extracts the data. You can specify how parameters can be included from the request URL.
- Stored Procedures have a default limit of 10 results. This can be changed by using the limit keyword in your SPARQL query.
- Optionally, an XSLT transform to create custom output formats, but XML and JSON is available by default. We've created [some generic ones you can use](#).

Setting up the API is quite easy and shouldn't take more than a few minutes. Read the [SPARQL Stored Procedure](#) documentation for more details.

Create Linked Data API

A Linked Data API allows you to define a custom RESTful API for extracting data from a dataset by defining the graph patterns, e.g. types of entities and their relationships, that are of interest.

You'll need:

- familiarity with Turtle for creating and editing the configuration file
- a good working knowledge of the Linked Data API vocabulary and/or a starter template that you can customize

Read the [Linked Data API](#) documentation to find pointers to get you started



Custom APIs

Query:

```
PREFIX swrc: <http://swrc.ontoware.org/ontology/swrc/>
PREFIX dc: <http://purl.org/dc/terms/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?paper ?title
WHERE {
  ?paper a swrc:InProceedings ;
         rdfs:label ?title ;
         dc:subject ?subject .
}
```

Parameters:

Parameter:

Default Value:

Required:

Type:

Base:

GET <http://api.kasabi.com/dataset/www2012/apis/7hf?apikey=&subject=online+communities&output=json>



There is more...

- <http://kasabi.com/doc/api>
- data management APIs (update, jobs, status, ...)



Does it cost anything?

Worker Bee	Swarm	Hive
<p>Best for individuals</p> <p>£15/mo</p> <p>All the features of the Free Bee plan with:</p> <ul style="list-style-type: none">60,000 credits per monthpublish up to 20 datasetsaccess to premium databulk data loadingown domain for data pagesemail support <p>Sign Up</p>	<p>Best for teams</p> <p>£120/mo</p> <p>All the features of the Worker Bee plan with:</p> <ul style="list-style-type: none">400,000 credits per monthpublish up to 100 datasetsset royalty rate for databrandable data pagesstatistics API <p>Sign Up</p>	<p>Best for enterprise</p> <p>£1500/mo</p> <p>All the features of the Swarm plan with:</p> <ul style="list-style-type: none">5,000,000 credits per monthpublish up to 500 datasetsprioritised data loadingcustom licensesprivate datasetstelephone supportservice level agreement <p>Sign Up</p>

<p>Free Bee</p> <p>Explore the benefits of Kasabi</p> <p>FREE</p> <p>Sign Up</p>	<ul style="list-style-type: none">9,000 credits per monthaccess any standard datasetspublish up to 5 datasetsstandard licensesbrowsable data pagesdata update APIlive data updatesdata usage statistics
--	--



Summary

- Kasabi is a platform to **publish, link, find** and **consume** data
- based on linked data principles
- Linked Data as a Service
- **APIs** over your data
- data in different **flavours** (turtle, json, rdf/xml)



Keep in touch!

- <http://kasabi.com>
- <http://blog.kasabi.com/>
- Twitter: @kasabi
- IRC: #kasabi (freenode.net)
- this presentation:

<http://www.slideshare.net/dunken69/the-kasabi-information-marketplace>



This work is licensed under a [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/).



Under the Hood

- **Cohodo**
- (used to be Talis Platform)
- distributed data platform
- load balancing, data replication, etc.