Technical details of the P2Pvalue directory

David Rozas (University of Surrey)

This work was partially supported by the Framework programme FP7-ICT-2013-10 of the European Commission through project P2Pvalue (grant no.: 610961).

Who am I?

- Free software enthusiast
- Working with Drupal for the past 4 years
- Background in CS (@URJC [Spain], @NTNU [Norway])
- Currently studying a PhD in Sociology @UniS [UK]: analysing the Drupal community from a sociological perspective (CBPP case study)
- Collaborating with the P2Pvalue project
- +info|contact:
  - @drozas 
  - www.davidrozas.com

Outline

- Drupal and why using it?
- Some key aspects of the architecture
- How to access the data
- Future work
- Questions

Disclaimer

- Still a beta version!
- Any issues, please contact us!
  - Contact form: http://www.directory.p2pvalue.eu/contact
  - Github: https://github.com/P2Pvalue/cbpp-directory-code/issues?state=open

What is Drupal?

- A free software framework to build web applications
- A community of +1M users and +30k developers (Drupal.org stats)
- +1M websites: whitehouse.gov, data.gov.uk, louvre.fr, mtv.co.uk, blogs.aljazeera.com, etc.
- +info: https://drupal.org/about

Why using Drupal?

- **Robust** architecture
- **Flexibility** to perform changes
- **Modularity** to extend with new functionalities: +25k contributed (non-core) modules
- Flexibility quite critical, taking into account the research nature of the initiative

3 key aspects of the architecture

- Flexible way to:
  - Manage permissions (who access what?): roles module
  - To classify (and update) the cases. E.g.: add a new type of license: taxonomy module
  - Display the contents in several formats: views module

Permissions and roles system

- Flexible way to manage permissions, workflow and set a “wiki-like” system
- Currently 3 main roles:
  - **Authenticated users** (open registration): permissions to add new cases, edit details about them, see different revisions (versions) of the case, etc.
  - **Moderators**: permissions to edit any kind of content, revert revisions in case of vandalism (post-moderation policy), add new terms to the vocabularies, etc.
  - **Administrators**: full permissions to change all the configuration parameters

Taxonomies

- Taxonomy: the practice of “classifying things or concepts”
- Using Drupal taxonomy system: a flexible way to classify and update the CBPP cases
- Currently using 12 vocabularies
- Each vocabulary has a set of “terms” that can be indexed and filtered by
  - Via URLs: http://www.directory.p2pvalue.eu/explore/typologies/type-activity/open-education
  - Via API (more details soon)

Views

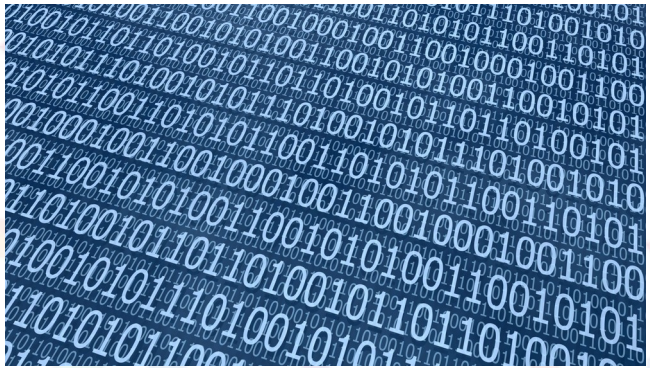
- A GUI to a SQL builder: provides us a flexible way to fetch content from the database and display it in several formats:
  - **HTML**

    E.g.: Use to expose the taxonomies as fields to filter by (Explore - http://directory.p2pvalue.eu/explore/cbpp-communities)
  - **JSON, XML**

    E.g.: Use to create the API (http://directory.p2pvalue.eu/api-instructions)

Accessing the data

- “How-to” contribute covered on “Tour” and other presentations
- Let's then take a look at “How-to” access the data



Accessing the data

- Data gathered is licensed under a CC0 1.0 license (dedicated to Public domain - http://creativecommons.org/publicdomain/zero/1.0/)
- The “one-click” way
- The “geeky” way

The “one-click”way

- A batch process to create an “up-to-date” version of the data in CSV
- Via http://www.directory.p2pvalue.eu/download/all-cbpp
- Probably the most useful way to import it for statistical analysis (R, Calc, SPSS, etc.)

The “geeky”way

- REST (Representational state transfer) service: an architectural style designed by W3C applied to the development of web services
- Currently supporting XML, JSON and PHP objects
- Probably the most useful way to be used by other applications

The “geeky”way: some examples

- Accessing the data in different formats:
  - XML (default): http://directory.p2pvalue.eu/rest/api (or .xml)
  - JSON: http://directory.p2pvalue.eu/rest/api.json
  - PHP: http://directory.p2pvalue.eu/rest/api.php

The “geeky”way: some examples

- You can also perform queries using the different values of the taxonomies by adding parameters to the URL
- E.g.:http://directory.p2pvalue.eu/rest/api?languages=%22Spanish%22

  “All the communities which include 'Spanish' as a *Language of the platform*”

The “geeky”way: some examples

- Several parameters can be combined: “AND” logic
- E.g.:

http://directory.p2pvalue.eu/rest/api?languages=%22Spanish%22&legal_entity=%22public%20institution%22

“All the communities which include 'Spanish' as a *Language of the platform* AND which have a value of "public institution" for the field *Legal entity*”

The “geeky”way: some examples

- A concrete CBPP case can be selected using the NID (Node ID – Unique identifier)
- E.g.: http://directory.p2pvalue.eu/rest/api.json?nid=250

  All the information about the case 250 (Drupal) in JSON
- +info : http://www.directory.p2pvalue.eu/api-instructions

Future work

- Data visualisation
  - Maps of the communities
  - Maps of the user actions
  - Graphical statistics
- Moderation policy
- User engagement mechanisms
- Let us know if you are playing with the data:
  http://www.directory.p2pvalue.eu/contact

Thanks!

… questions [?]

Slides: https://github.com/drozas/p2pvalue_directory_slides