

Galileo, first steps

Luis Cárcamo-Ulloa

Technologies in *galileo*



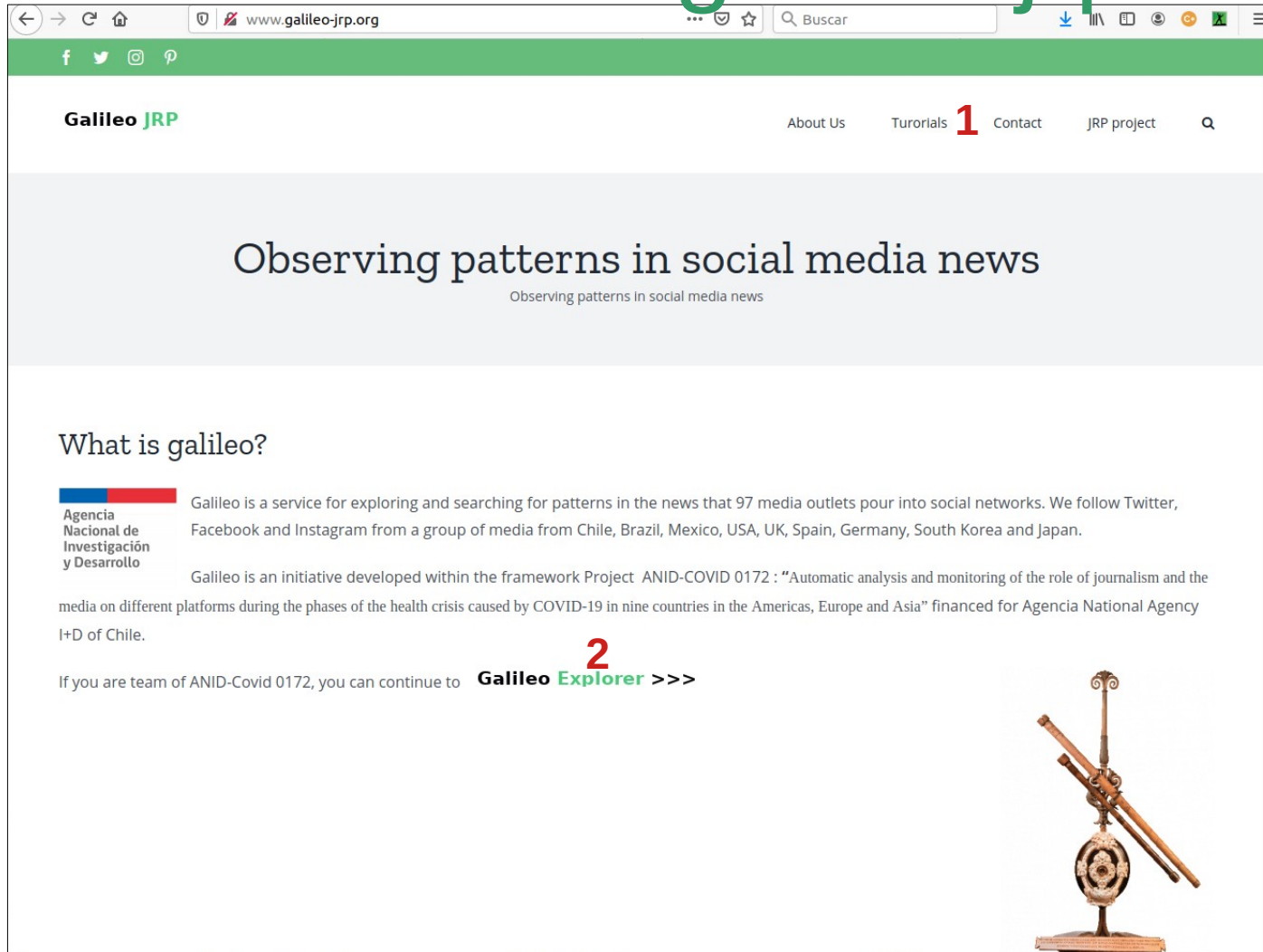
elastic → Engine search for no structured information



kibana → Visualize plugin



Go to galileo-jrp



The screenshot shows the homepage of the Galileo JRP website. The browser address bar displays 'www.galileo-jrp.org'. The website has a green header with the Galileo JRP logo on the left and navigation links for 'About Us', 'Tutorials 1', 'Contact', and 'JRP project' on the right. A search icon is also present. The main content area features a large heading 'Observing patterns in social media news' with a subtitle 'Observing patterns in social media news'. Below this, there is a section titled 'What is galileo?' which includes a logo for the 'Agencia Nacional de Investigación y Desarrollo' and text describing the project's purpose and funding. At the bottom of this section, there is a link to 'Galileo Explorer >>>' with a large red number '2' above it. An image of a Galilean telescope is visible in the bottom right corner of the page.

0 go to
www.galileo-jrp.org

1 In this section you
will find Tutorial and
FAQ

2 Click to Explorer for
work in Galileo.

Define a query

The screenshot shows a data discovery tool interface with the following components:

- Query Bar:** Contains the query "Covid coronavirus cuarentena pandemia" (1), a KQL language selector, a date range filter "Jan 1, 2020 @ 01:22:09.53 → Nov 30, 2020 @ 01:37:38.96", and a "Refresh" button.
- Filter Panel:** On the left, it shows "instagram-espana" as the selected source and a list of available fields like "message", "_id", and "index".
- Histogram:** A bar chart showing the count of results over time from February to July 2020. A dropdown menu is open over the histogram, showing options for "Auto", "Millisecond", "Second", "Minute", "Hour", "Day", "Week", "Month" (3), and "Year".
- Calendar:** A calendar view for November 2020 is shown on the right, with a date selected (2).
- Raw Data:** Below the histogram, two rows of raw data are displayed (4). The first row shows a message about a world map of coronavirus cases. The second row shows a news article about a deadly day in Argentina.

1 You query can be defined for a) set of keywords, b) "exact expression" and c) combined booleans [AND, OR, NOT]

2 Define date in and out.
3 Define histogram bar periodicity

4 Raw data. Text in disorder, but you can select fields and order data.



Organizing fields

The screenshot shows the Discover tool interface with several key elements:

- Search Bar:** Contains the query "covid-barcelona_search".
- Filter Options:** A dropdown menu is open, showing various index patterns like "facebook-uk", "instagram*", and "instagram-alemania".
- Selected Fields:** A list of fields is shown, including "Description", "ImageText", "MediaOrigin", "MediaOutlet", "URL", and "Country". The "Country" field is circled in red.
- Save Search Dialog:** A dialog box is open, asking to save the search query. The "Save" button is highlighted.
- Search Results:** A bar chart shows the count of results over time, with a peak in late 2020. Below the chart, there are several search results from various sources, including "La Vanguardia" and "El País".

Red arrows and numbers 1-4 indicate the following steps:

1. Add filter
2. Filter by type
3. Add Country field
4. Save search query

1 Select index (Facebook, Twitter or Instagram) or sub-index (every social media and every country).

2 Incorporate relevant fields (variable) for every social media. The fields have different names in every social media and different order.

3 If you make comparative studies, you can use one index (example, Twitter) and add the Country field.

4 Save search query

Download Corpus

The screenshot shows the Kibana search interface. At the top, there are menu options: New, Save, Open, Share, Inspect. Below that, the search criteria are: 'Covid corona' in the 'instagram-espana' index. A 'SHARE THIS SEARCH' button is visible. The search results show 232 hits. A bar chart displays the count of results per week from January 2020 to August 2020. A 'Generate CSV' button is highlighted with a red '1'. Below the chart, there is a 'Copy POST URL' button with a red '2'. A table of search results is visible, with columns for Time, MediaOrigin, Description, ImageText, and URL. A dialog box titled 'Abriendo covid-barcelona_search.csv' is open, showing the file name and a 'Guardar archivo' button with a red '4'. A 'Download report' button is also visible with a red '3'.

1 Share and download corpus

2 You can generate CSV or spreadsheet [; = colum separator]

3 Download report

4 Save in download directory in your PC