Some journas retrived with the api from medpub central eu

Add quantities

Add it is difficutl to create the search with p.values it detedct ps

Jstore data request form: <http://info.ithaka.org/acton/form/10452/0054:d-0002/0/-/-/-/-/index.htm>

The information provide herein will become part of the agreement between ITHAKA and the recipient of the data. We are going to need to destroy the data upon completion of the project and notify Jstore. Completing a dataset agreement and processing of the dataset may take 2-5 weeks. Datasets are delivered via an online delivery mechanism.

Below, is the needed information about the person requesting the data. I can make the request, but I think is better to use your information since your academic position has more institutional prestige.

First Name

Last Name

Email Address

Title

Department

Institution

Project Description

I have added a tentative description of the project. Please, feel free to add, change, as you considerer it necessary.

The overarching purpose of this research is to understand how the use of p-value (probability value), as a method for quantifying statistical significance of a statistical model, has evolved historically. The project will draw from Institutionalization and Collective Attention theories to understand how the level of institutionalization of p-values leads to lesser degrees of specification, and the consequences of such declining levels of description. The methodology will make use of computational methods for data mining from large text corpora. This research project will provide and important contribution to the current debate about the misuse of p-values in fields were p-values are extensively used such as Life Sciences, Biomedical Sciences, and Psychology.

**3. Data Requested**

**We are requesting full texts.**

OCR Full Text: Depending on the journal and issue, this may be either text extracted from page scans via OCR or the full text as submitted digitally by publishers. There is not distinction between OCR and digitally-sourced full text in DFR datasets.

4. What is the scope and nature of the Data to be provided?

The scope of the data is articles for several journals since early 1900th to current dates.

JSTOR Search URL, if one has been constructed

I can construct one once we define the dates. The search will basically select the journals that we decide we want. I try to select articles that have p-values in them, but the jstore search algorithm selects articles that have p’s and p.s’s as well as p=’s, p>’s etc.

5. Please confirm the parameters of your dataset:

The dataset will be limited to these publication dates:

In jstore there are journals that go back to the 1800th, however some interesting journals such as American Sociological Review which starts at 1936. So, I was wondering if should we limit the data set to journals that go back to the 1800th or data sets that start in the 1930 will work for out purposes. What do you think?

The dataset will be limited to specific titles (please separate titles with commas):

From the list of journals (see Excel file attached), I attached to this email, I we can selected the following journals:

In JSTOR, we have the journal Sciences (204800 articles) and The Scientific American (203011). I like these two journals because they are similar sub-fields, number, and date ranges. We also have the Proceedings of the National Academy of Sciences of the United States of America (135225) which is similar to the two above, but this we can get it from the PMC Europe with the API you built in R. In the social sciences, I think we should stay with The American Journal of Psychology (13503) since Psychology tends to be more quantitative the Sociology and the journal goes back to the 1800th; What do you think? In JSTOR, we have The American Sociological Review (14172) which is nice, but it started in 1936.

In PMC Europe, we also have the American Journal of Pathology, which is a health or medical journal, which seem interesting. Which again we can get with you API.

Please see the attached Excel file for complete info about the journals.

So, from JSTOR we can get:

1. Science
2. The Scientific American
3. The American Journal of Psychology

And, using your API, I can get from PMC Europe the:

1. The Proceedings of the National Academy of Sciences of the United States of America
2. The American Journal of Pathology if we want.

Please, let me know your thought on this.

6. "Key researchers" and their affiliation:

Valentin Danchev, The Knowledge Lab, The University of Chicago

Camacho Jonathan, The University of Chicago

Others, if any, who are expected to have access to the Data (for example, graduate students or undergraduates)? Will other outside third parties require access to the Data? Please list their names and email addresses.

Not sure if you will want to allow other access to the data.

7. Use of the Data - check all that apply:

Research

Please explain:

9. Results of the research using the Data will be used as follows:

Publications

Conferences

Please explain:

12. Would you be willing to share progress on Project or method of research with JSTOR?

Yes