P-01

Extended Abstract

Data mining in synchrotron radiation

W. Dwiecki[1]*

[1]*Faculty of Chemistry, Adam Mickiewicz University in Poznan, Poznan, Poland*

Keywords: synchrotron, data mining, Google Scholar

*e-mail: wojciechdwiecki@wp.pl

INTRODUCTION

Growing popularity and availability of synchrotron radiation-based techniques in recent years results in increase in the number of publications and articles which can be easily found in scientific databases or popular web browsers [1]. Due to the amount and variety of publications it is a serious problem to predict how issues related to synchrotron will develop in future. However, analysing the data from past 30 years collected with commonly available tools for example Google Scholar and appropriate combination of queries can provide some information about trends and research areas where synchrotron radiation will develop and be the most frequently applied [2].

DATA SOURCE SELECTION

The most important issue in data exploration is a selection of the right data source. In this presentation there was compared with one another: Google Scholar and two databases: Scopus and Academic Search Complete (ASC) [3].

Searched results presented in *Figure 1.* and in *Table 1.* confirm that Google Scholar is definitely the largest "mine" of knowledge about the synchrotron as was predicted. Such large differences in results are because: (a) Google Scholar searches queries in entire document, Scopus and ASC searches only in title, abstract and keywords of publication; (b) Google Scholar is in fact a collection of results from all existing databases.
Despite the fact that Scopus and ASC allows for searching with more advanced options than Google Scholar, number of publications current in databases are much smaller and require from user very precise typing of queries which can be time-consuming.

MEASUREMENT TECHNIQUES USING SYNCHROTRON RADIATION

Using Boolean operators (such as OR and AND) number of results in Google Scholar could be limited to the most accurate. During data exploration *synchrotron* was combined with abbreviations of most common measurement techniques which were inserted between quotation marks (equivalent to searching WITH EXPRESSION). In this way number of publications related to the following techniques was investigated: Extended X-ray absorption fine structure (EXAFS), X-ray absorption fine structure (XAFS), X-ray absorption near edge structure (XANES), Energy-dispersive X-ray diffraction (EDXRD), Small-angle X-ray scattering (SAXS) and Wide-angle X-ray scattering (WAXS). Results given in *Figure 2.* show that most popular measurement technique is EXAFS. This kind of queries combined with another words could be a useful tool in evaluation the best method for the measurement of specific material.

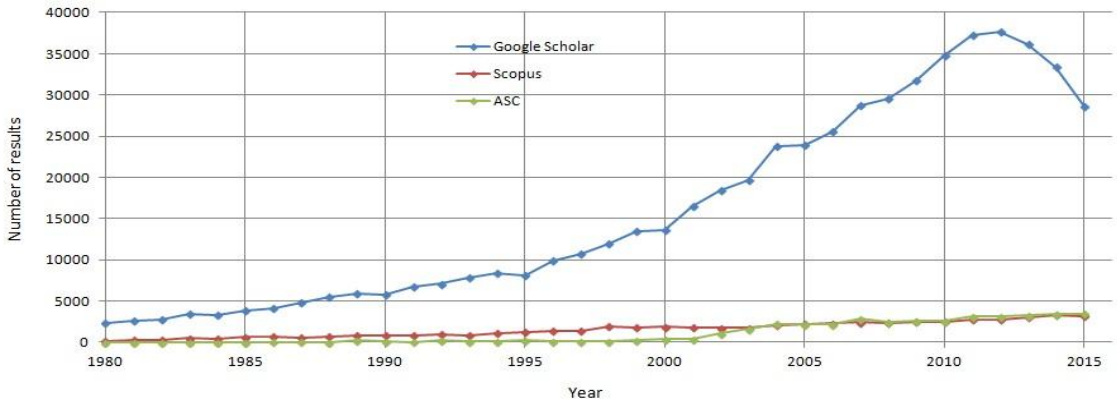

Figure 1. Comparison of number of results *synchrotron* collected in different sources in the period 1980 – 2015.

	Google Scholar	Scopus	ASC
Number of results	567530	56058	39191
Fraction of all results (%)	85,63%	8,46%	5,91%

Table 1. Number of results *synchrotron* in different sources in the period 1980 – 2015.

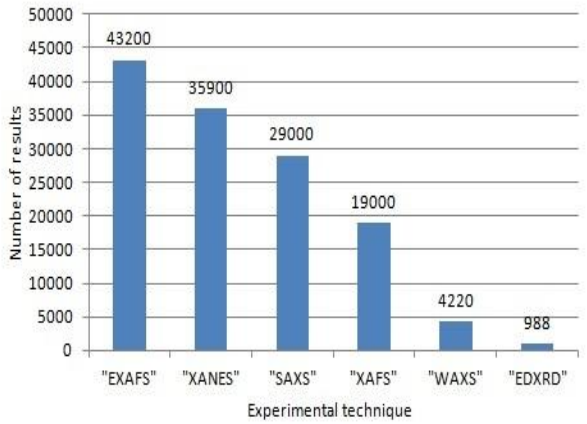

Figure 2. Number of results containing names of measurement techniques received in Google Scholar.

SYNCHROTRON RADIATION IN CRYSTALLOGRAPHY

Investigations focused on crystallographic structure of material can be divided into two groups: diffraction of single crystals and powder diffraction. Google Scholar gives the answer which of these two techniques are used more commonly in the synchrotron radiation (*Figure 3.).*

SYNCHROTRON POPULARITY IN JOURNALS

Advanced search in Google Scholar allows determining the frequency with which the word *synchrotron* appears in the most prestigious scientific journals. One of Google Scholar built-in function is a list of journals which are divided into different areas of science. In data analysis was used journals titles from bookmark "Chemistry and material science" with the highest h5 index [4].

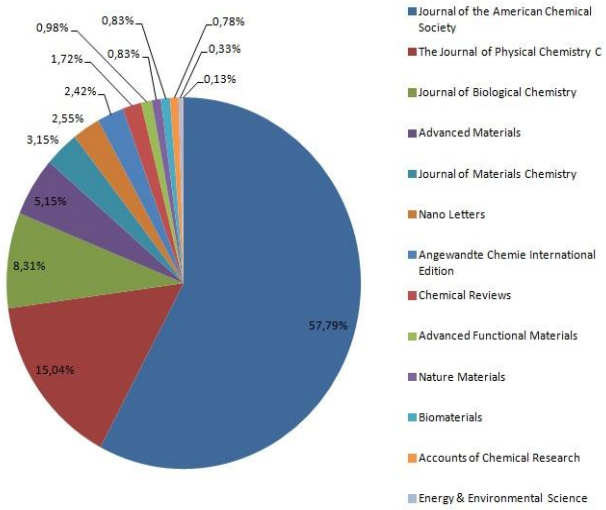

Figure 4. Number of results *synchrotron* in journals [%] received in Google Scholar.

Results in *Figure 4.* show interests of this journals in topics related to synchrotron and can help in the selection of the journal to which the future publication will be sent.

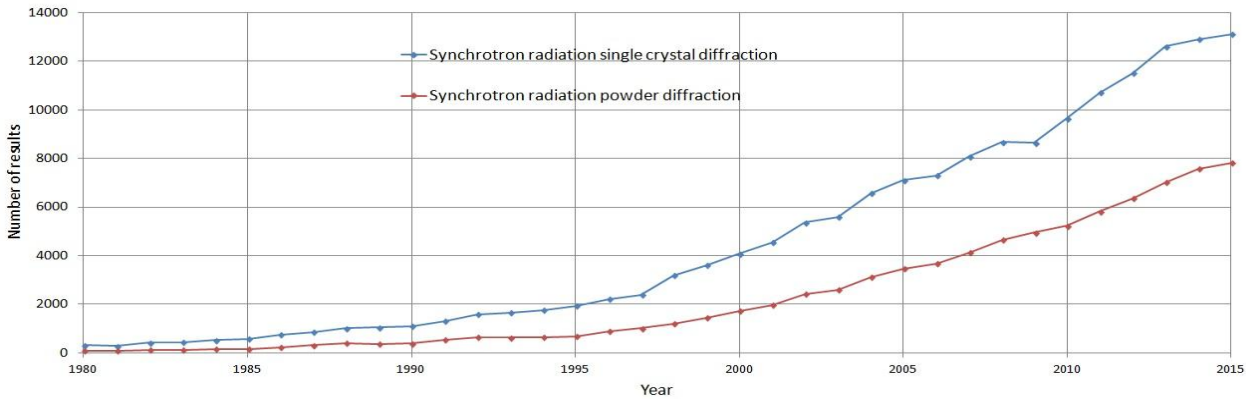

Figure 3. Numbers of results SRSCD/SRPD in the period 1980 – 2015 received in Google Scholar.

This trend can indicate that amount of materials which can be measured using *synchrotron radiation powder diffraction* (SRPD) is more limited than those which are measured using *synchrotron radiation single crystal diffraction* (SRSCD) or, what is also possible, SRSCD gives better rusults than SRPD. More advanced data searching will help in choosing the most suitable technique for specific material.

[1] A-W. Harzing, S. Alakangas, *Scientometrics* (2013) 1-18.
[2] Jankowski W., Hoffmann M., *J Med Internet Res* **18** (2) (2016) e38.
[3] H. R. Jamali, M. Nabavi, *Scientometrics* **105** (2015) 1635.
[4] A. Martín-Martín, J. M. Ayllón, E. Orduña-Malea, E. Delgado López-Cózar, EC3 Working Papers, 17 (2014).