Abstract

The report introduces basic definitions on the topic, provides information on the history and current state of the institutional repository of Ural Federal University, comments on some systemic issues, and analyzes statistics. An attempt has been made to show the connection between the work done and the position of the repository in the world

PART 1

Today we shall discuss the Ural Federal University, which is in Yekaterinburg (Russia), at the Department of information and analytical support. At URFU, we are responsible for the operation of the institutional repository and are not part of the library structure.

The report will be formally divided into several parts, beginning with an introduction of a few terms:

Institutional repository - an electronic archive for long-term storage, accumulation and provision of long-term and reliable open access to the results of scientific research carried out at the institution.

This definition is as general as possible and can be found in the Wikipedia article on institutional repositories. Still, it is important to reflect the "spirit" of the entity, not the "letter", and if a service so defined is called a Digital Library, a Digital Archive or otherwise, but is nevertheless used for long-term storage, accumulation, as well as providing access to the results of scientific and research activities of an organization, we will consider it an institutional repository.

As noted above, the institutional repository accomplishes the following tasks:

- Ensuring free access to the results of scientific research conducted at the university through self-archiving;
- Access to scientific research of the university for the world community;
- Concentration of materials in one place;
- Preservation of other electronic materials, including unpublished (so-called gray literature), such as dissertations and technical reports.

When talking about the quality, ranking, and evaluation of repositories, we usually turn to TRANSPARENT RANKING: Institutional Repositories by Google Scholar by CyberMetrics Lab.

The purpose of this ranking is to support open access initiatives and thus free access to scholarly publications in electronic form and other academic materials. Web indicators are used here to measure the global visibility and impact of scientific repositories.

The methodology used in the evaluation within this ranking is not uncontroversial, hardly repeatable, but nevertheless understandable. We should not treat this rating as a competition, but it is the best way to assess our own and the world's dynamics.

PART 2

When we are done with general information, we can move on to the essence. What is the Institutional Repository of the Ural Federal University today? In a nutshell:

- The project started in 2002
- The use of DSpace started in 2004
- Batch loading has been practiced since 2013
- The project is registered and indexed in the main industry aggregators (more details on the slide in the bottom right part)
- More than 20 years of "high availability"
- More than 100 thousand documents in the public domain

For a slightly fuller picture, as well as to answer the question about the scalability and portability of good practices, let's look at a couple more repositories from Yekaterinburg:

Electronic Archive of the Russian State Vocational Pedagogical University:

- The project started in 2014
- Batch loading has been practiced since 2014
- The project is registered and indexed in the main industry aggregators
- More than 40 thousand documents in the public domain

Electronic Archive of the Ural State Forest Engineering University:

- The project started in 2012
- Batch loading has been practiced since 2015
- The project is registered and indexed in the main industry aggregators
- More than 10 thousand documents in the public domain

It should be noted that colleagues with our support have adopted many of the practices used at URFU, which allowed to achieve certain results regardless of the size of the university, the material base, etc. I thought about how to compare the three universities brought above, and decided to compare them by year of foundation, to show the history and base of results of scientific activity, to compare the budget enrollment in the first year, to show the difference in the size of universities and compare repositories.

The smallest university with a budget enrollment of 600 places and a total volume of students of all forms and levels of study shows not a bad result. The other two universities are also great fellows.

PART 3

Speaking of dynamics, systematic work and other manifestations of work in the format, from 2004 to 2021 – these are the actual years of the DSpace repository. It shows the publication years of the materials. The materials published in the period from 1920 to 1989 are combined in one line.

The number of materials in the archive for a specific year of publication, the dynamics of uploading materials to the repository, and the number of materials published in the archive in the year of publication are also depicted. One table can give information about the speed of downloading actual data, the annual growth, and the dynamics and depth of reprovisioning, etc.

Many of you here who follow Webometrics are probably aware that the Lab publishes at least two rankings - Universities and Repositories. The latter ranking is divided into several more - institutional and aggregating, journal portals, CRIS systems. One of the results of our years of systematic work is the graph where the horizontal axis are edits of the rating from 4 to 14, and on the vertical axis is the position in the rating. It is to be noted that we have always been in the TOP50 of the world, and in the latest edition we are in the TOP25. In terms of dynamics, we can try to look for a connection between the rate of fund growth and place, to draw conclusions about the growth of the world and our growth, but these would be empirical conclusions. It will be an idea for discussion, not a technology that guarantees results when applied.

The second result is the demand for repository content. In the first three weeks of September, we see 19.7 million impressions of our content in Google searches and 336,000 user clicks from Google searches to the repository. Here we can confidently talk about the absence of robots. Overall, September ended with the numbers reflected on the right side of the slide. In our opinion, it is quite worthy.

CONCLUSIONS

Finally, I have rather freely formulated conclusions:

The content is in demand. Not our content, but basically content typical of an institutional university repository.

The description and presentation are accurate. This follows from a good ranking in the search indexes of general and specific search engines.

The world is growing. This is probably the only indisputable statement.

We are growing at about the same rate. This conclusion can be made based on the data of analytical and webometric tools, as well as ratings. But this does not mean that an increase in the fund of 1000+ titles guarantees growth for everyone.