OPEN SCIENCE OVERVIEW

According to prominent open science advocate Michael Nielsen, open science is based on the idea that “scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process”1. Sharing the research process has never been easier, as most data are born digital and stored electronically. Using a cyberinfrastructure of integrated science tools scientists are now able to conduct the research process openly from beginning to end.

There are many benefits to doing research openly, including accelerating the discovery process by making use of collective intelligence and citizen science, reducing the duplication of work thereby saving time and resources, as well as increasing the visibility and impact of research2. Although open science has not been fully embraced by everyone in the research community, there have been many successful open science projects. By working with researchers and policy makers, as librarians, we can help the open science movement to continue to grow.

...PAST

For much of history, science has been based on observation and description. The development of the scientific method resulted in a shift from description to experiment and measurement. Below is a summary3 of the evolution of the scientific process:

- **Empirical science**
  - Thousands of years ago - describing natural phenomena

- **Theoretical science**
  - Past few hundred years - using models and generalizations to explain phenomena

- **Computational science**
  - Past few decades - using digital technology to simulate phenomena

THE FOURTH PARADIGM

The past 10 years or so have seen the most recent evolution in the scientific process, which was coined ‘the fourth paradigm’ by computer scientist Jim Gray4. The fourth paradigm refers to data-intensive scientific discovery, the focus of which is reusing existing data to find new meaning.

THE HGP AND THE BERMUDA AGREEMENT

The Human Genome Project (HGP) officially began in 1990. Not long afterward, labs all over the world began to churn out large volumes of genomic data. The following graph illustrates the exponential increase in data that was deposited in the data repository GenBank between 1982-20085.

Despite the positive implications of the HGP on medical advancements, some scientists were reluctant to deposit sequences openly, so in 1996 a number of researchers from various sequencing centers around the world met in Bermuda and set principles for the release of data. Now known as the Bermuda Agreement6, the principles were:

1. Primary genomic sequences should be in the public domain
2. Primary genomic sequences should be rapidly released

The HGP and the Bermuda Agreement set a precedent for openness in the field of human genetics and successfully changed scientific culture.

OPEN SCIENCE: PAST, PRESENT AND FUTURE

...PRESENT...

Thousands of open science tools are used by researchers every day. Different tools support different functions at various stages of the research process. The following are some examples of the different types of open science tools currently used:

- **Project planning**
  - Tools that track the development of ideas and experimental methods

- **Data collection and analysis**
  - Data repositories and tools that make use of collective intelligence and citizen science

- **Information retrieval**
  - Integrated search tools that link data with literature

- **Demonstrating impact**
  - Networking tools and tools that illustrate impact of all forms of contributions

VARIOUS TOOLS

THE Role of the librarian

As librarians, we have an important role in participating in open science. Our various skills can help ensure that research is maintained with the goal of long-term preservation and future use. The figure below illustrates the data lifecycle and the points at which librarians can be involved.

FUTURE...

The following solutions proposed by the Royal Society, Science Policy Centre7 address the challenges mentioned.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislation and policy</td>
<td>• standards for sharing information in order to ensure usability and interoperability</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>• continual development of new software tools able to hold large data sets</td>
</tr>
<tr>
<td>Training and support</td>
<td>• involvement of experts (i.e. librarians) able to manage and support digital data</td>
</tr>
<tr>
<td>Scientific culture</td>
<td>• increased openness among scientists themselves, as well as with the public and media</td>
</tr>
<tr>
<td></td>
<td>• consistent recognition for gathering and sharing data</td>
</tr>
</tbody>
</table>

SELECTED SOURCES