

## Long-term availability of bioinformatics web services

Sebastian J. Schultheiss<sup>1,\*</sup>, Marc-Christian Münch<sup>2</sup>, Gergana D. Andreeva<sup>2</sup>,  
Gunnar Rätsch<sup>1</sup>

<sup>1</sup> Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany

<sup>2</sup> Wilhelm Schickard Institute for Computer Science, University of Tübingen, Germany

\* E-mail address: [sebastian.schultheiss@tuebingen.mpg.de](mailto:sebastian.schultheiss@tuebingen.mpg.de) or [sebi@umich.edu](mailto:sebi@umich.edu)

URL to the original data: <http://tinyurl.com/fml-webservices> or

<http://www.fml.tuebingen.mpg.de/raetsch/suppl/long-term-availability-of-bioinformatics-web>

This work is licensed under the Creative Commons Attribution 3.0 Unported License

We studied all 927 open-access web services found in the seven NAR Web Server Issues published between 2003 and 2009. We checked their availability, usability and functionality. A web service is defined here as an application usable through a web browser.

An astonishing number of 90% of services are still reachable. However, while the original web page may still be online, up to 62% of services no longer function as published when tested with their example data. Surprisingly, the services published in 2003 clearly stand out from the following years in terms of availability, as this issue contains many of the most accessed and cited services in computational biology developed before 2003. The requirements for publication of services in NAR have risen constantly, with services this year having to provide the following information directly on their web pages: contact information, example data, help texts, and version information. The presence of this information, and service usability, has risen constantly from year to year. Based on these qualities, we created the Long-Term-Score, which reflects the number of criteria fulfilled by a service, thus measuring service quality.

We counted the number of times a service's NAR publication was cited. We can show that bad design choices correlate with a low number of citations: the ratio of citations for services with high usability vs. services with low usability shows that better services are cited 2.2 times more often.

Given these findings, we can provide reviewers, editors, and most notably web service developers with guidelines that make the use of their service easier for everyone and simultaneously allow for time-saving maintenance of the service for the years after publication, hopefully leading to a higher number of citations. The most effective way of ensuring continued access to a service is a persistent web address, offered either by the publishing journal or instituted on the authors' own initiative.