

# [EN] 07. The Statistics of the Use of the Reader Application and of the Digital Objects Made Available through That Application

The reader application of the dLibra system has a built-in statistics mechanism for end users' basic actions, such as searching or displaying the content of objects. Those data are collected solely on the basis of data obtained by the reader application, so end users cannot block that tracking, which is possible when data are tracked by additional external tools, such as Google Analytics. What is more, the statistics collected by the reader application also include calculations of direct references to files with content which are not easy to track with the use of such tools as Google Analytics.

The table below contains a breakdown of the basic statistical possibilities offered by the built-in statistics mechanism of the dLibra system, compared to the use of the Google Analytics tool.

We also recommend reading the [subchapter about the details of the possibilities of analyzing the behavior of users of the reader application of version 6 of the dLibra system with the use of Google Analytics](#).

	The Built-In Statistics of Version 6 of the dLibra System	Google Analytics
The Data Collection Method	On the reader application side, on the basis of the received HTTP requests.	On the user's web browser side, on the basis of the Java Script tracking code.
Data completeness	The data include all requests from users while the browser bot traffic is ignored (see below).	The data do not include:: <ul style="list-style-type: none"><li>• users who block the user tracking in their browsers with the use of dedicated plugins;</li><li>• Users who refer directly to publication files, without the mediation of the digital library website (for example, by going to a PDF directly from Google results); and</li><li>• browser bot traffic (see below).</li></ul>
Ignoring the traffic generated by browser bots	<p>The programs (browsers) which are not included in the statistical calculations are selected based on the content of the <b>"ignored_agents.txt"</b> configuration file. Every entry is treated as a regular expression (there is a sample entry for Googlebot in the file). Character chains which correspond to the "User Agent" field (HTTP heading) of the programs which should not be calculated (the file contains an example for Googlebot) should be placed there. User agent names can be found online, for example, on the <a href="http://en.wikipedia.org/wiki/User_agent">http://en.wikipedia.org/wiki/User_agent</a> site. Once the changes have been introduced in the files, the Tomcat should be restarted so that the new values can be taken into account.</p> <p>Once Tomcat has been started up, all ignored agents from the <a href="http://dlibra.psnc.pl/ignored_agents.txt">"http://dlibra.psnc.pl/ignored_agents.txt"</a> file are saved to the "ignored_agents.txt" file.</p>	Internal Google Analytics mechanisms are responsible for ignoring the bot traffic.

Available statistics	<ul style="list-style-type: none"> <li>Statistics concerning the development of the resources of the digital library; <ul style="list-style-type: none"> <li>the total number of objects available in a period of time (accurate to within a month)</li> <li>the total number of new objects in a month (accurate to within a month)</li> <li>the total number of objects in a month, broken down into object formats (accurate to within a month)</li> </ul> </li> <li>statistics concerning user behavior <ul style="list-style-type: none"> <li>the total number of generated web pages (accurate to within a month)</li> <li>the total number of searches, broken down into simple and advanced searches (accurate to within a month)</li> <li>the total number of visitors (sessions; accurate to within a month)</li> <li>the total number of displayed objects (accurate to within a month, several impressions of the same object during one user session is counted only once)</li> <li>the total number of displayed objects, broken down into object formats (accurate to within a month, several impressions of the same object during one user session is counted only once)</li> </ul> </li> </ul>	Only statistics concerning user behavior, very complex, dependent on the configuration of Google Analytics (for more information, see <a href="#">here</a> ).
Data format	Charts (PNG), tables (HTML, CSV).	Interactive charts and tables, export to the CSV, XLS, and PDF formats.
Access method	By default, publicly for all, at the <main address of the digital library>/stats (for example, <a href="http://sbc.org.pl/stats/">http://sbc.org.pl/stats/</a> ).	Access only for authorized users, through the <a href="#">Google Analytics service</a> .