INFM600: Quarry Queries

# Query-Set 1

**User:** Zache

**Query 1:** *Fiwiki abuse filter hits per day*. Retrieved October 16, 2016, from Quarry, <https://quarry.wmflabs.org/query/634>

**Query 2:** *Fiwiki abuse filter hits per day limited by filter\_id.* Retrieved October 16, 2016, from Quarry, <https://quarry.wmflabs.org/query/639>

**Query 3:** *Fiwiki abuse filter hits per day limited by afl\_actions.* Retrieved October 16, 2016, from Quarry, <https://quarry.wmflabs.org/query/640>

**Table used:** *abuse\_filter\_log*

**Description:** The *abuse\_filter\_log* table comes under the AbuseFilter extension of wiki. This extension enable the privileged users to set up triggers when other users’ actions match certain specified criteria (e.g. When the edits made by user contains inappropriate words, he/she would be banned) (“Extension: AbuseFilter,” 2009).

The first query is aimed at calculating the total number of abuse filter hits in Finnish Wikipedia. In the second one, the user drills down to specific filter ids and for the third one, to filter actions. The user might have been trying to figure out how many people violate the terms of using Wikipedia or perform illegal actions each day. In this case, he/she is trying to calculate the number of people that are banned from editing every day.

**Query 4:** *Fiwiki revert analysis.* Retrieved October 16, 2016, from MediaWiki, <https://quarry.wmflabs.org/query/688>

**Tables used:** *page, revision*

**Description:** The *page* table contains the details of each and every page on wiki (“Manual: Page table,” n.d.) whereas *revision* table logs every edit made to any of the pages (“Manual: Revision table,” n.d.).

Through this query, the user might be trying to analyze the Finnish-wiki pages whose revisions had been reverted within a specific time range (*here:* a week) using the comments associated with each edit/revision.

# Query-Set 2

**User:** Tgr

**Query 1:** *Infinte-blocked editors with the most edits on huwiki.* Retrieved October 17, 2016, from Quarry, <https://quarry.wmflabs.org/query/12603>

**Tables used:***ipblocks, user*

**Description:** As the name suggests, *ipblocks* table contains details like IP addresses of the users who have been blocked from editing (“Manual: Ipblocks table,” n.d.).

Through the query the user might be trying to find out the details of indefinitely blocked users having the most number of edits in the Hungarian wiki. The user implemented join on the *user* and *ipblocks* table to find out the number of edits for each user.

**Query 2:** *Most linked stubs on hu.wikipedia*. Retrieved October 17, 2016, from Quarry, <https://quarry.wmflabs.org/query/531>

**Tables used:** *pagelinks, templatelinks, page*

**Description:** The *pagelinks* table stores all the internal links in the wiki (“Manual: Pagelinks table,” n.d.) and *templatelinks* stores the details like id of host page, namespace of the linked page, etc (“Manual: Templatelinks table,” n.d.).

The user might be aiming to get the list of pages in Hungarian wiki with the links to another particular wiki page (here: page whose title begins with ‘Csonk’ substring). I believe that the user has implemented inner join on *page* and *pagelinks* table to eliminate the pages which have been deleted or renamed and consider only those pages which are currently available on the wiki.

# Query-Set 3

**User:** Stefan2

**Query 1:** *File redirects with file uploads.* Retrieved October 17, 2016, from Quarry, <https://quarry.wmflabs.org/query/1281>

**Tables used:** *page, image*

**Description:** The *image* table gives the details about image/file uploads in wiki (“Manual: Image table,” n.d.).

The user might be aiming to find out all the redirected English wiki pages containing file uploads. The inner join between the *page* and *image* tables ensures that the page contains uploaded files or images and the page\_is\_redirect column ensures that the page is indeed a redirected page.

**Query 2:** *Cross-namespace redirects to files*. Retrieved October 17, 2016, from Quarry, <https://quarry.wmflabs.org/query/5825>

**Tables used:** *redirect, page*

**Description:** The *redirect* page holds the details of all pages which are redirects (“Manual: Redirect table,” n.d.).

The user might be trying to find out the page titles of English wiki that are redirected to the file namespace. But the catch here is that the page titles themselves should not belong to the file namespace. Hence, the query is named ‘Cross-Namespace Redirects to Files’. The join on the *page* and *redirect* tables is to figure out the source of the redirects.

**Word Count: 672**

# References

[1] *Extension: AbuseFilter.* (2009, April 19). Retrieved October 16, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Extension:AbuseFilter>

[2] *Manual: Page table.* Retrieved October 16, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Page_table>

[3] *Manual: Revision table.* Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Revision_table>

[4] *Manual: Ipblocks table.* Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Ipblocks_table>

[5] *Manual: Pagelinks table.* Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Pagelinks_table>

[6] *Manual: Templatelinks table*. Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Templatelinks_table>

[7] *Manual: Image table*. Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Image_table>

[8] *Manual: Redirect table*. Retrieved October 17, 2016, from MediaWiki, <https://www.mediawiki.org/wiki/Manual:Redirect_table>