## Architecture of separated survey and crowdsourcing datasets processing with data validation and quality analysis

General diagram – desired target configuration



## Prioritized list of improvements to realize new fingerprint processing architecture

1. Develop new validators for validation of IVL data
2. Add route and dataset identification to each grid line in grid file
3. Develop a tool to analyze validation logs to determine survey routes reliability.

General idea: analyze validator log for detecting of routes with worse than usual rejection percentage

* 1. …

Result: list of unreliable routes

1. Develop a tool to detect and reject invalid (inconsistent) datasets

General idea: Calculate RMS of difference between mag measurements of each dataset and fingerprint (grid) cells. Analyze a distribution of calculated RMS metrices and reject datasets with the worst metrices (or as an alternative with the metrices which does not fit to normal distribution).

* 1. …

Result: Validated magnetic, *wifi and ble* grids after invalid data rejection

1. Develop a tool to detect unreliable routes and reject their data from the grids

General idea: Select all datasets of specified rout, calculate survey quality metrices for all cells covered by the rout and calculate percentage of good and bad cells for the rout. Reject all datasets of the route if this one contains too much wrong cells.

* 1. …

Results: Validated magnetic, *wifi and ble* grids after invalid data rejection and list of routes recommended to redrawing

1. Add interface to utilize several grids by FPBL Console
   1. …
2. Improve magnetic fingerprint distribution estimation. Calculate distribution parameters depends on histogram form:
   1. Robust estimation for narrow normal-like histograms
   2. Uniform based estimation for wide uniform-like histograms
   3. Use kurtosis for histogram form determination
3. Mag quality map
   1. Mag coverage map -> Mag survey quality map:
      1. Display all cells covered by survey routes as grey, uncovered as blue
      2. If cell has at least one measurement in grid display it as red-yellow-green in depends on survey quality
      3. If cell has one to 100 measurements in grid display it as red ()
      4. If cell has more than 5000 measurements in grid display it as green ()
      5. Calculate survey quality metric for each cell in depending on histogram span

* 1. Total coverage map -> Total survey quality map:
     1. Calculate survey quality metric for each cell as follows:
  2. Improve magnetic uncertainty chart – calculate uncertainty depends on histogram form as described in 8.a

1. Develop IVL routes and dataset validation tool for magnetic grid
   1. TBD
2. Develop a tool for assessment of quality of data surveyed by each surveyor

General idea: analyze validator log for calculating of rejection data percentage for each surveyor and ranking of surveyor’s baed on comparison of individual rejection percentage with average rejection percentage.

* 1. …

Result: list of surveyors with survey quality ranking

1. WiFI uncertainty map
2. Transfer magnetic grid in for database format
3. Transfer wifi, ble and portal grid into database format - optional

## MVP, February sprint

General diagram



Compact diagram

