News in PxEdit

Veli-Matti Jantunen
  Statistics Finland
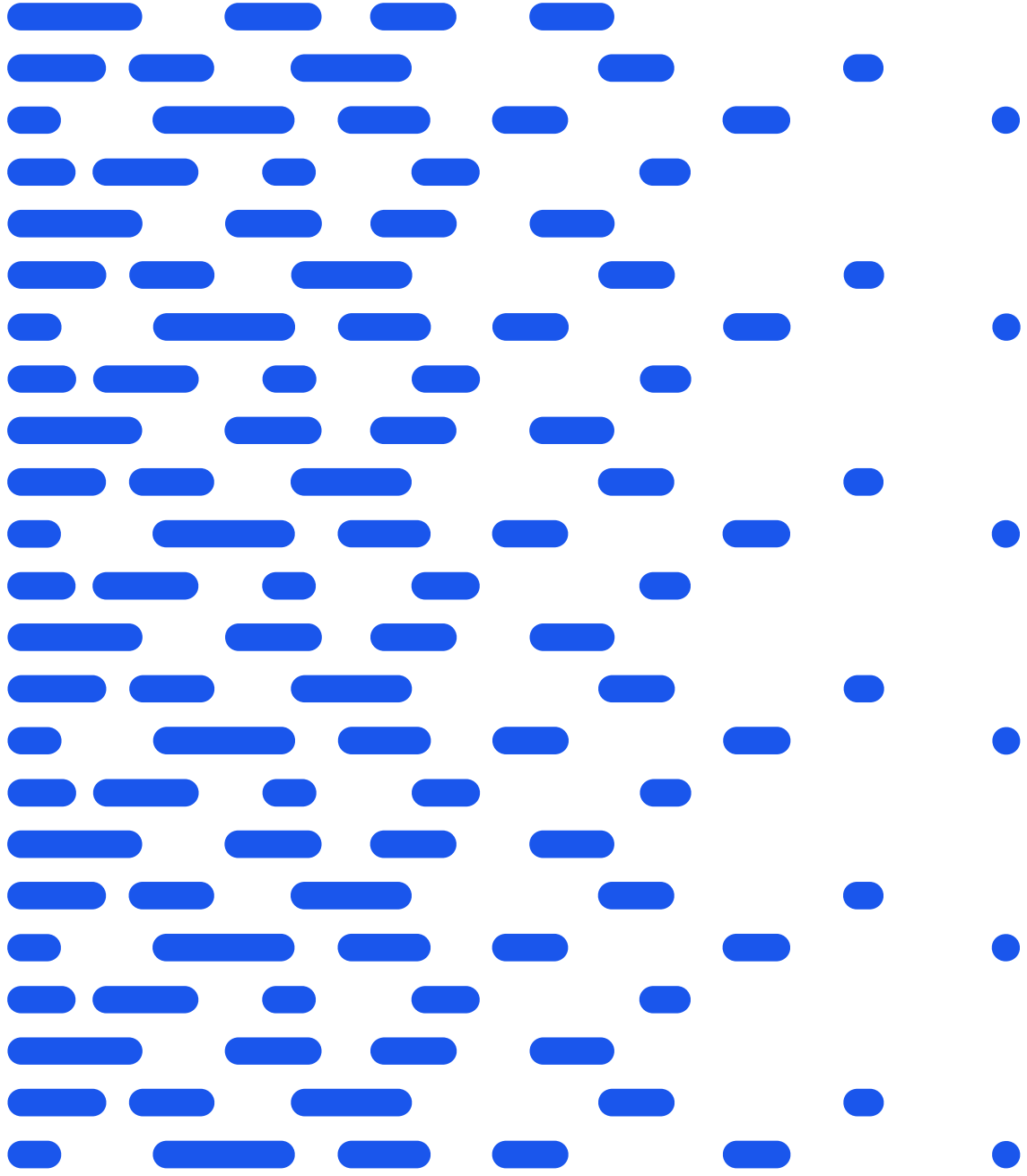 Development and Digitalisation
  Technology Services
  Interface Technologies

15/7/2023
PX-Meeting in **Riga**

Statistics Finland

PxEdit 4.1

New version (and some other) features /1

- Upgraded version
  - new dll's
  - Internet connection package

- Speedups
  - especially **huge** eccentric tables, such as large Excel tables, text files and px tables with lots of metadata
    - up to 40-60%
  - Excel/clipboard reading still have the upper limit which comes from available RAM
    - ..but I have some ideas ☺

- Bug fixes
  - yes, there has been **some** ☹

- Tutorial presentations available (you may add your own `.pdf` files if needed)
  - *Help|Tutorials*

New version (and some other) features /2

- Open database connection
  - There has to be a [**Database**] section in the main settings file

```
databases=StatFin,Paavo;SiStat                                         list of available open databases
StatFin_api=https://statfin.stat.fi/pxweb/api/v1/fi/StatFin/           API path for retrieving metadata
StatFin_path=https://statfin.stat.fi/database/StatFin/                 path for retrieving whole tables
Paavo_api=https://statfin.stat.fi/pxweb/api/v1/fi/Postinumeroalueittainen_avoin_tieto
Paavo_path=https://statfin.stat.fi/database/Postinumeroalueittainen_avoin_tieto
..
```

- Comparing current table to the database version
  - PxEdit: *Edit|Compare to database*
  - PxJob: `set= -h313 out=`

- List the used web links (in the keywords)
  - PxJob: `-h200 out=`

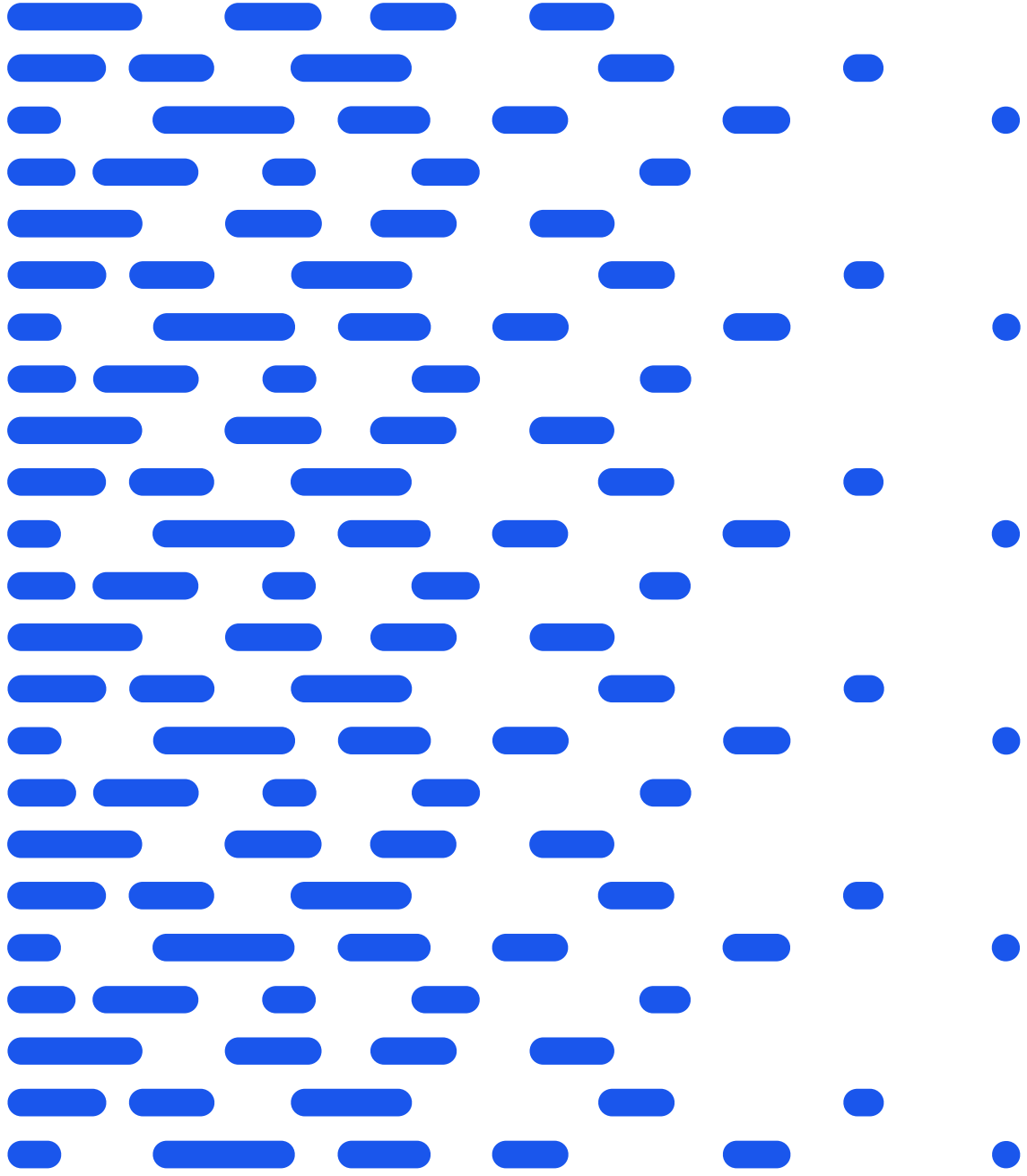- Test the database web links
  - PxJob: `set= -h404 out=`
  - BTW: may be quite slow ☹

New version (and some other) features /3

- Reading metadata from control file in PxEdit: *Edit|Import from control file*

- Enhanced *TIMEVAL* handling
  - support for new frequencies: **D** (daily) and **T** (three times per year)
  - week frequency handles the week number 53
  - possible to switch the *TIMEVAL* checking off (for non-standard time variables)
    PxEdit: *Settings|File reading|Skip TIMEVAL*
    PxJob: `!`**T**

- Keep the original *DECIMALS* when joining tables
  - PxEdit/Join window: *Keep DECIMALS*
  - PxJob: `-j`**d**

- Analyse the .vs and .agg files used in the database
  - PxJob: `set= -h`**101** `out=`

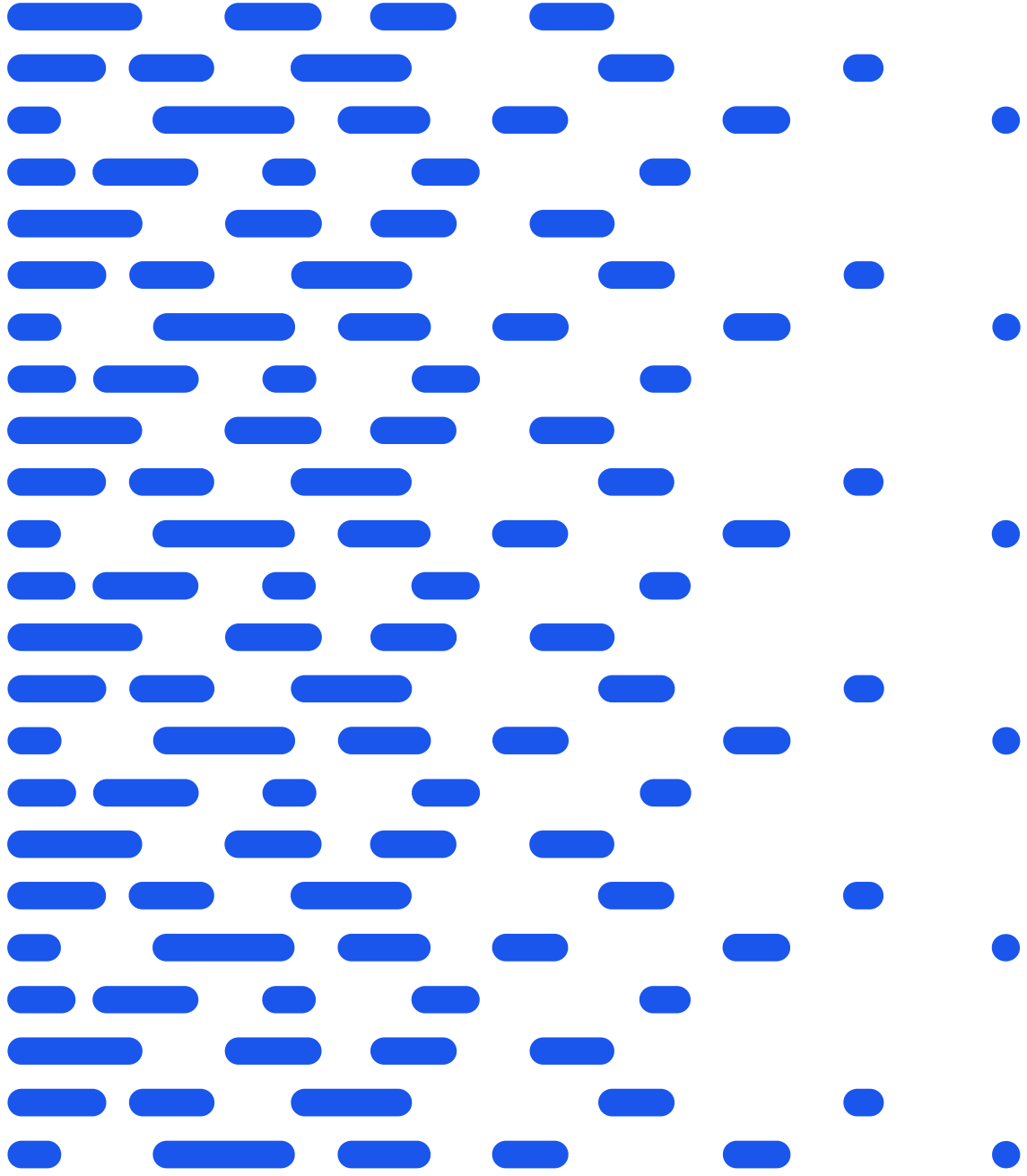- **Ctrl+R** opens the Report window

- Variable font …

Metadata enriching

There are several ways to inject metadata

- Editing **manually** in PxEdit

- Using **structural** tables with keyword blocks

- Table **joining**
  - PxEdit: *Edit|Join tables*
  - PxJob: **in**=..,.. **-j**

- Using **default** values defined in the settings file
  - PxEdit: *Edit|Defaults* or **Ctrl+k**
  - PxJob: **set**=.. **-m**

- **Importing** metadata from another `.px` (or `.pxk`) file
  - PxEdit: *Edit|Import keywords* or **Ctrl+i**
    Edit|Import from   (from **open** table)
  - PxJob: **meta**=`../import.px`   (metadata **file**)
    meta=`../import/`   (**directory** & sub-directories, the source file has similar name as the original one)

- Reading metadata from a **control file**
  - PxEdit: *Edit|Import from control file*
  - PxJob: **meta**=`..control.csv`   (there are extra options, such as deleting and replacing as well)

Future

Statistics Finland

Retirement...

- I *could* have retired last summer - ..but won't do that in the near future ☺

- It is good time to ponder what to do with PxEdit/PxJob so that they will be available for (some) years to come
    - *GitHub* seems to be the answer to that problem

- My aim is to develop the version **5.0** for being open software, which means that
    - there will be **intermediate** versions which will slowly be (internally) changed to the open version
    - not much new features – actually I will remove old stuff as much as possible
    - perhaps only **64-bit** version

- I have an **alpha** version which contains most of the things, such as
    - informative object names (doesn't everybody use one-character variable names, though?)
    - better code structure (from flat architecture to a more namespace-related one)

- There's still a lot of work to be done
    - source code to **text files**
    - comments and documentation in English
    - ancient code remnants should be upgraded
    - even the programming environment can be standardised

These features will be removed in the future (if there is no enough resistance)

- **CoSSI/XML**
  - **not** being used in Statistics Finland any more

- eXplorer output format

- PxSQL output

- *Statistics Finland specific*
  - *eXist* connection
  - output packages
  - some special/outdated PxJob functions

- Others
  - obscure/outdated settings (e.g. `ISO_date`)