

QPati data integration from XML

Document author: anna.hammais@tyks.fi

Data source

ktpanalytics: /mnt/raw/qpati_xml_autosync directory contains files coming from QPati via “Matin allas” (data pool).

Timestamps in the XML files

Even though the timestamps end with “Z” denoting “Zulu time” /UTC/GMT, they are really just the local time of the system, i.e. EET (eastern European time) or EEST (eastern European summer time).

Statement tables

The following tables contain more than two columns:

```
SYDNEY-LUOKITUS SUPPEA.tbl  
SYDNEY-LUOKITUS.tbl  
MUNUAINEN IMMUNOFLORESENSSI.tbl  
ENNUSTETEKIJÄT.tbl  
RINTASYÖVÄN ENNUSTETEKIJÄT.tbl  
IHO IMMUNOFLORESENSSI.tbl  
Zavaus.tbl  
SOLULASKENTA BAL.tbl  
VIRTAUSSYTOTOMETRIA.tbl  
BETHESDA 2001.tbl
```

In the ETL (*update_patologia.kjb*), these tables are handled by the script

```
ETL/R/qpatitable/qpati_table_<version>.R
```

The following tables contain only one column:

```
OBDUKTIO MIKROSKOOPPISET NÄYTTEET.tbl  
AIVO-OBDUKTIO MIKROSKOOPPISET NÄYTTEET.tbl
```

The values in this column are entered with the variable/suure name ‘Suure:’ in the table *main.patologia_taulukkoarvo*.