(Geo)Kettle – Ein unvollständiger, subjektiver Erlebnisbericht

Stefan Ziegler

Amt für Geoinformation
Rötistrasse 4
4500 Solothurn

xx. Januar 2014

Inhalt

- Was ist Kettle? Was ist PDI? Was ist Geokettle?
-
- Geo ohne GeoKettle?
- Demo
- Inspiration

Kettle/PDI

- ETL-Tool (Extract, Transform, Load)
- Kettle = PDI (Pentaho Data Integration)
- Dual-licensing

Funktionsweise

- Hübsche Oberfläche
- Relativ intutiv…
- Prozesse gespeichert in einer XML-Datei (in DB-Repo oder filebasiert)
- Kettle *interpretiert* XML-Datei
- Prozess steuerbar mit Parameter und Variablen.
- Prozesse werden gestartet:
  - aus der Oberfläche *Spoon*
  - von der Kommandozeile *Pan* resp. *Kitchen*
  - via Webserver *Carte* (auf entferntem Rechner)
  - oder ganz abgefahren in Cluster.

Weitere ETL-Tools

- Talend Open Studio (TOS):
  - komplizierter?
  - Codegenerator: Erzeugt Java-Code.
- FME:
  - Kenne ich nicht.
  - Arbeitsplatzwechsel resp. Arbeitsmittelwechsel notwendig ☹
  - Interlis ☺

GeoKettle

- Kettle um Geoprozesse erweitert.
- «Geo» *nicht* als Plugin integriert.
- GeoKettle basiert auf Kettle 3 (aktuell Kettle 5).
- aber…

Geo ohne GeoKettle

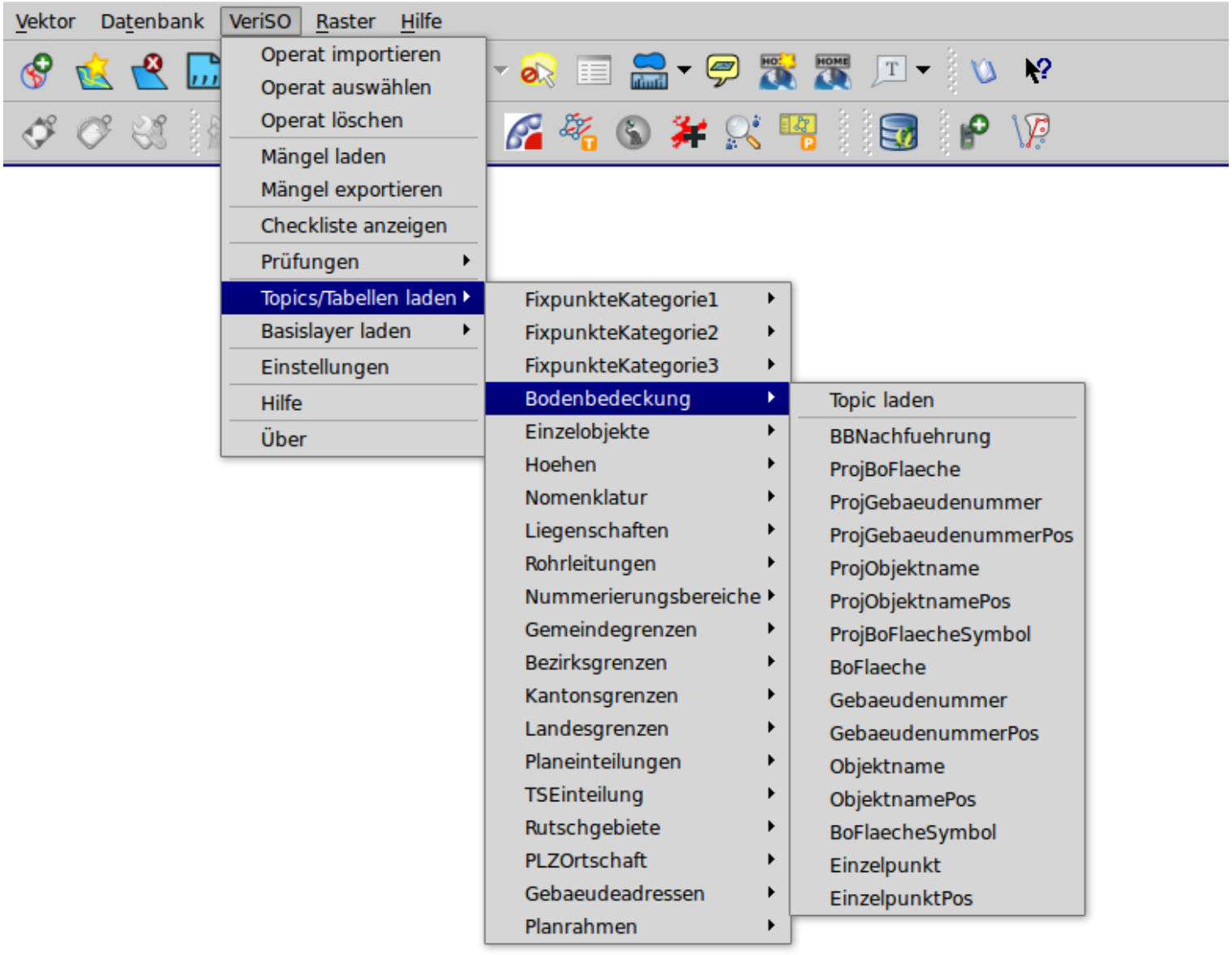
- Sämtliche Geoprozesse werden in der Datenbank gemacht.
- Funktioniert tadellos aber einschränkend, da z. B. kein Import/Export von Shape-Dateien etc. möglich ist.
- Beispiele:
    - Bodenbedeckungsarten pro Liegenschaft (im öffentlichen Eigentum)
    - AV-Fileverifikation

Persönliches Fazit

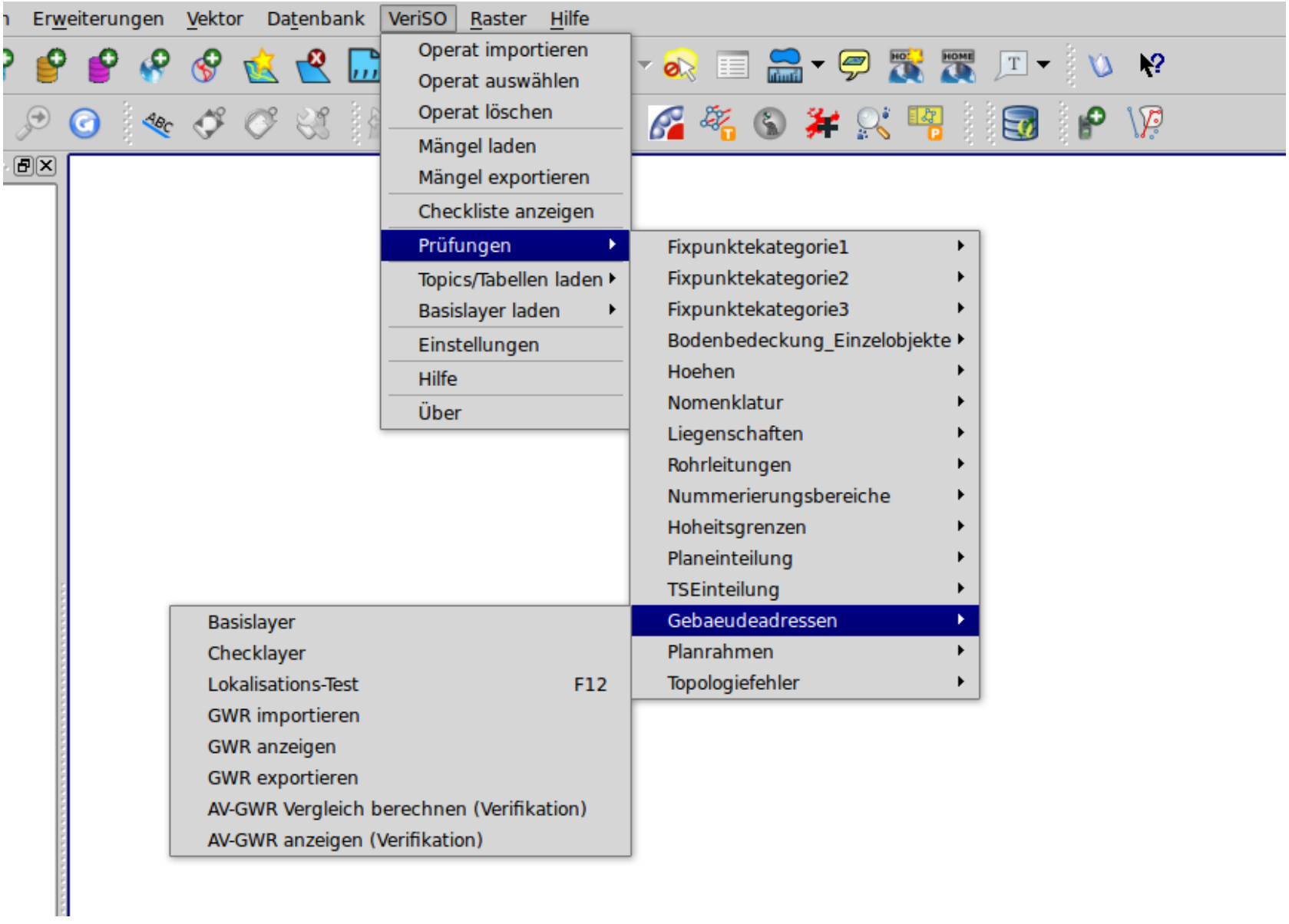
- Keine eierlegende Wollmilchsau.
- Oftmals reicht wahrscheinlich auch ein Skript.
- GeoKettle, quo vadis?

- Fexibel: läuft auch als Cronjob.
- Übersichtliches GUI: einzelne Prozessschritte sind klar sichtbar.
- Tiefe Hemmschwelle: man programmiert nicht.
- Grosse Auswahl an Werkzeugen: E-Mail-Versand, Logging, Access/Excel In-/Output, SAP-Input (?) etc.

QGIS-GUI


Vektor
Datenbank
VeriSO
Raster
Hilfe
Operat importieren
Operat auswählen
Operat löschen
Mängel laden
Mängel exportieren
Checkliste anzeigen
Prüfungen
Topics/Tabellen laden
Basislayer laden
Einstellungen
Hilfe
Über
FixpunkteKategorie1
FixpunkteKategorie2
FixpunkteKategorie3
Bodenbedeckung
Einzelobjekte
Hoehen
Nomenklatur
Liegenschaften
Rohrleitungen
Nummerierungsbereiche
Gemeindegrenzen
Bezirksgrenzen
Kantonsgrenzen
Landesgrenzen
Planeinteilungen
TSEinteilung
Rutschgebiete
PLZOrtschaft
Gebaeudeadressen
Planrahmen
Topic laden
BBNachfuehrung
ProjBoFlaeche
ProjGebaeudenummer
ProjGebaeudenummerPos
ProjObjektname
ProjObjektnamePos
ProjBoFlaecheSymbol
BoFlaeche
Gebaeudenummer
GebaeudenummerPos
Objektname
ObjektnamePos
BoFlaecheSymbol
Einzelpunkt
EinzelpunktPos

QGIS-GUI


Erweiterungen
Vektor
Datenbank
VeriSO
Raster
Hilfe
Operat importieren
Operat auswählen
Operat löschen
Mängel laden
Mängel exportieren
Checkliste anzeigen
Prüfungen
Topics/Tabellen laden
Basislayer laden
Einstellungen
Hilfe
Über
Fixpunktekategorie1
Fixpunktekategorie2
Fixpunktekategorie3
Bodenbedeckung_Einzelobjekte
Hoehen
Nomenklatur
Liegenschaften
Rohrleitungen
Nummerierungsbereiche
Hoheitsgrenzen
Planeinteilung
TSEinteilung
Gebaeudeadressen
Planrahmen
Topologiefehler
Basislayer
Checklayer
Lokalisations-Test F12
GWR importieren
GWR anzeigen
GWR exportieren
AV-GWR Vergleich berechnen (Verifikation)
AV-GWR anzeigen (Verifikation)

QGIS-GUI: ILI → XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<topics>
  <topic id="FixpunkteKategorie1" title="FixpunkteKategorie1" group="FixpunkteKategorie1">
    <table>
      <group>FixpunkteKategorie1</group>
      <title>LFP1Nachfuehrung</title>
      <schema>av_dm01avso24</schema>
      <table>fixpunktekategorie1_lfp1nachfuehrung</table>
      <geom>perimeter</geom>
      <style></style>
    </table>
    <table>
      <group>FixpunkteKategorie1</group>
      <title>LFP1</title>
      <schema>av_dm01avso24</schema>
      <table>fixpunktekategorie1_lfp1</table>
      <geom>geometrie</geom>
      <style></style>
    </table>
    ....
  <topic id="Bodenbedeckung" title="Bodenbedeckung" group="Bodenbedeckung">
    <table>
      <group>Bodenbedeckung</group>
      <title>BoFlaeche</title>
      <schema>dm01avso24</schema>
      <table>bodenbedeckung_boflaeche</table>
      <geom>geometrie</geom>
      <style></style>
    </table>
    ....
 </topic>
</topics>
```

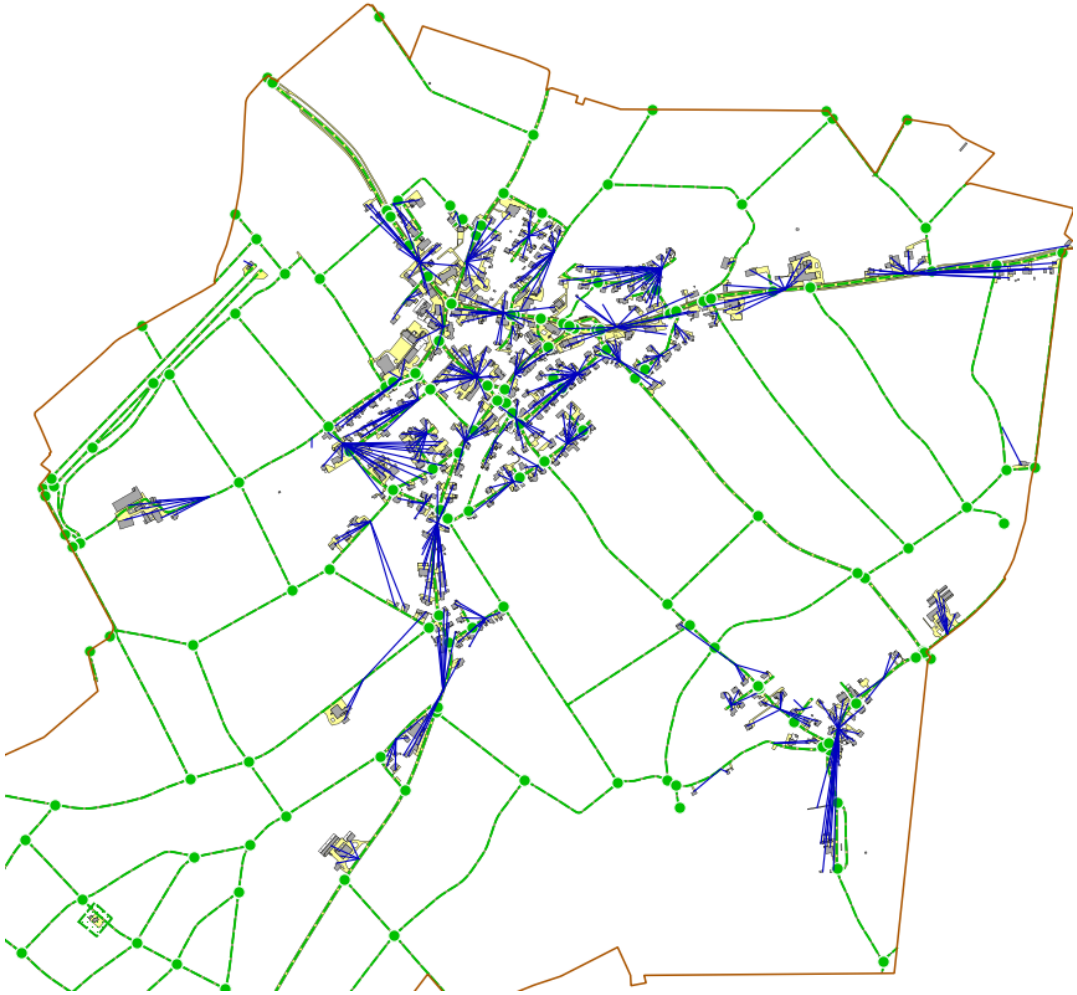
QGIS-GUI: XML für Checks

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<checks>
  <check id="fixpunktekategorie1" title="Übersicht" group="Fixpunktekategorie1"
         topic="Fixpunktekategorie1" type="complex">
    <file>complexchecks.fp1</file>
  </check>
  ...
  <check id="gebaeudeadressen_basis" title="Basislayer" group="Gebaeudeadressen - Basislayer"
         topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_basislayer</file>
  </check>
  <check id="gebaeudeadressen_check" title="Checklayer" group="Gebaeudeadressen - Checklayer"
         topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_checklayer</file>
  </check>
  <check id="gebaeudeadressen_lokalisations" title="Lokalisations-Test"
         group="Gebaeudeadressen - Lokalisations-Test" topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_lokalisation</file>
    <shortcut>F12</shortcut>
  </check>
  ...
</checks>
```

Checks

- «Alles ist möglich!»
- «einfach»:
  - Vordefinierte PostGIS-Views oder Tabellen (falls View zu langsam).
  - Tabellen werden während des Importes abgefüllt.
  - Queries sind in spezieller Tabelle gespeichert.
- «komplex»:
  - beliebig kompliziert
  - Alles was die Python- und PyQGIS-API hergibt, z. B.:
    - Excel-Export
    - PDF erzeugen
    - Daten mit WFS requesten und vergleich.
    - …
- Alles in *mindestens* einer Python-Datei verpackt.

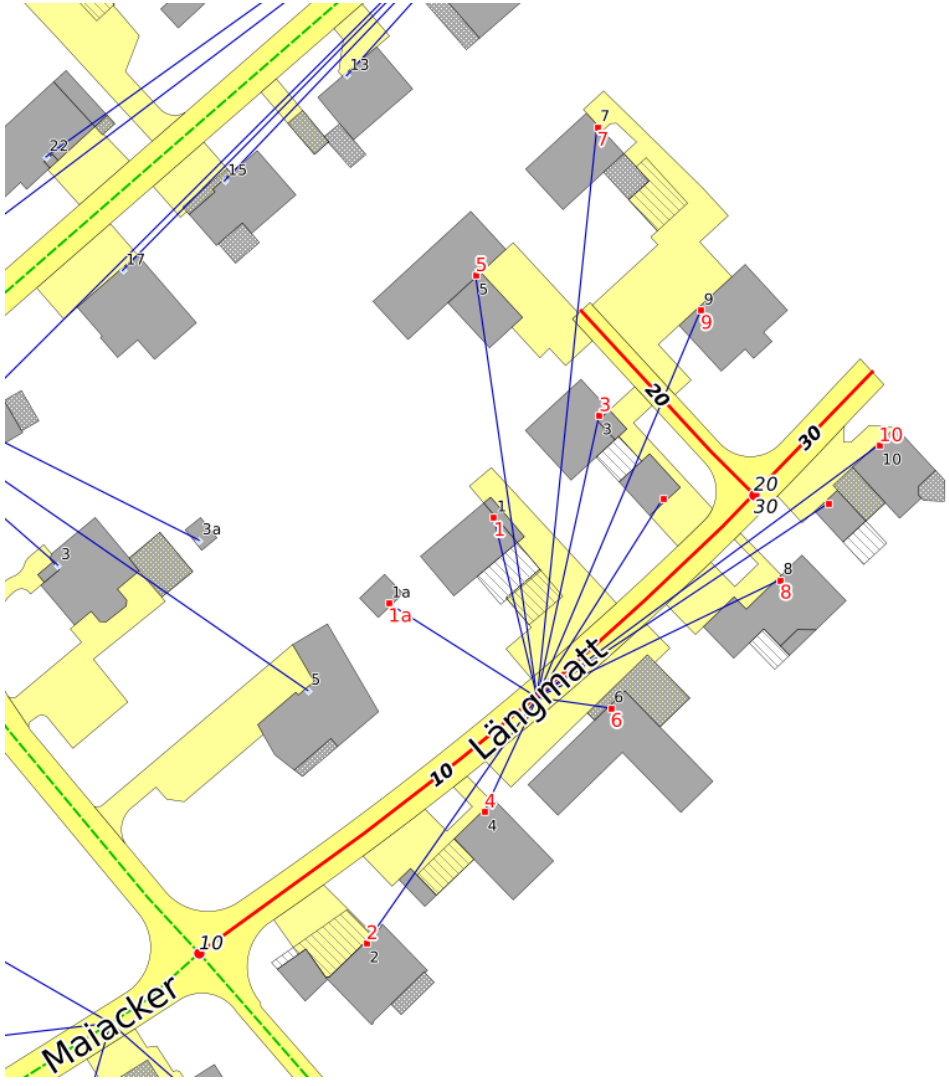
Checks



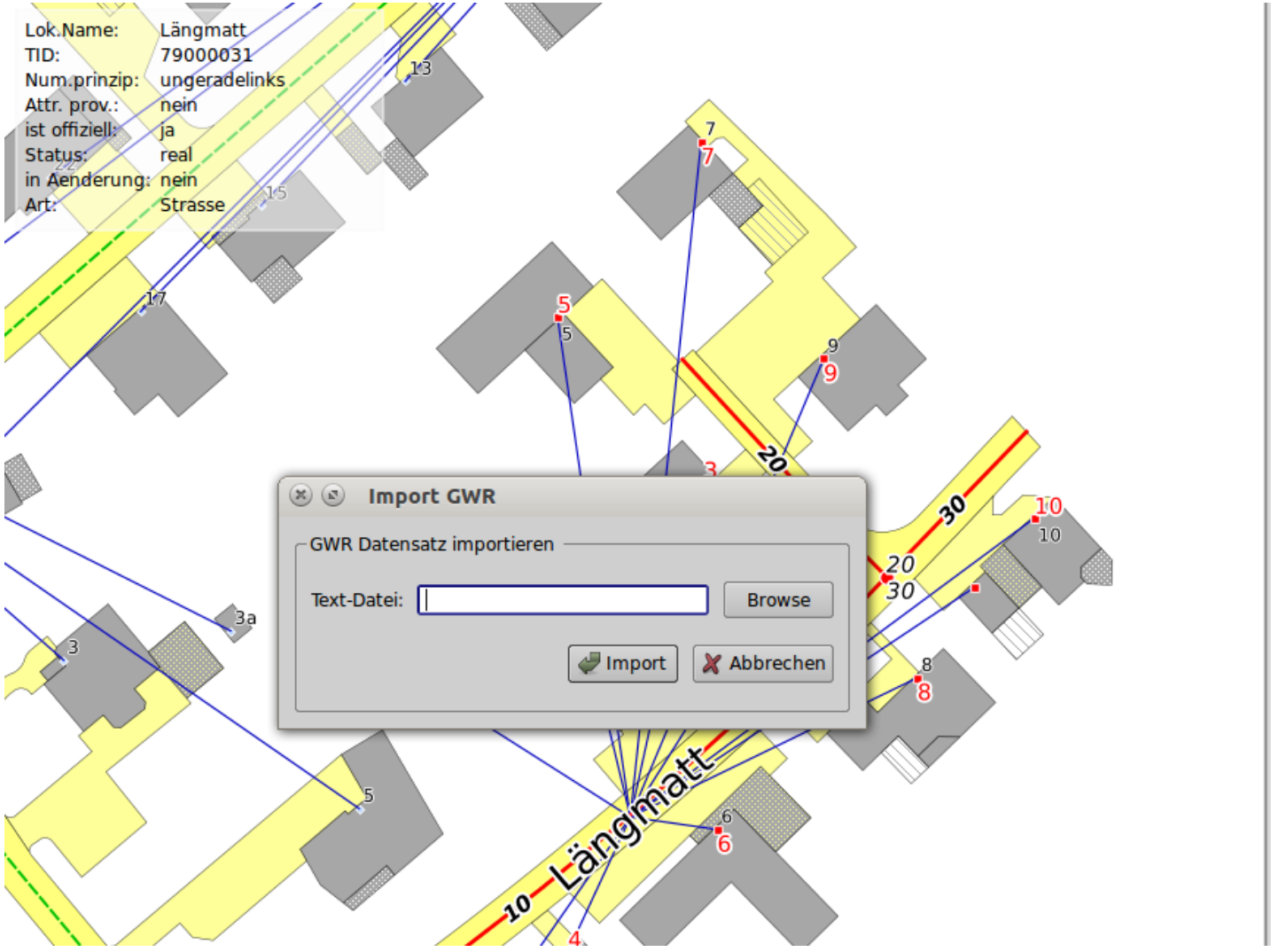
Checks


Maiacker
Längmatt
13
22
15
24
17
7
5
9
19
3
10
3a
1
3
1a
8
5
6
10
4
2
2
1
3

Checks



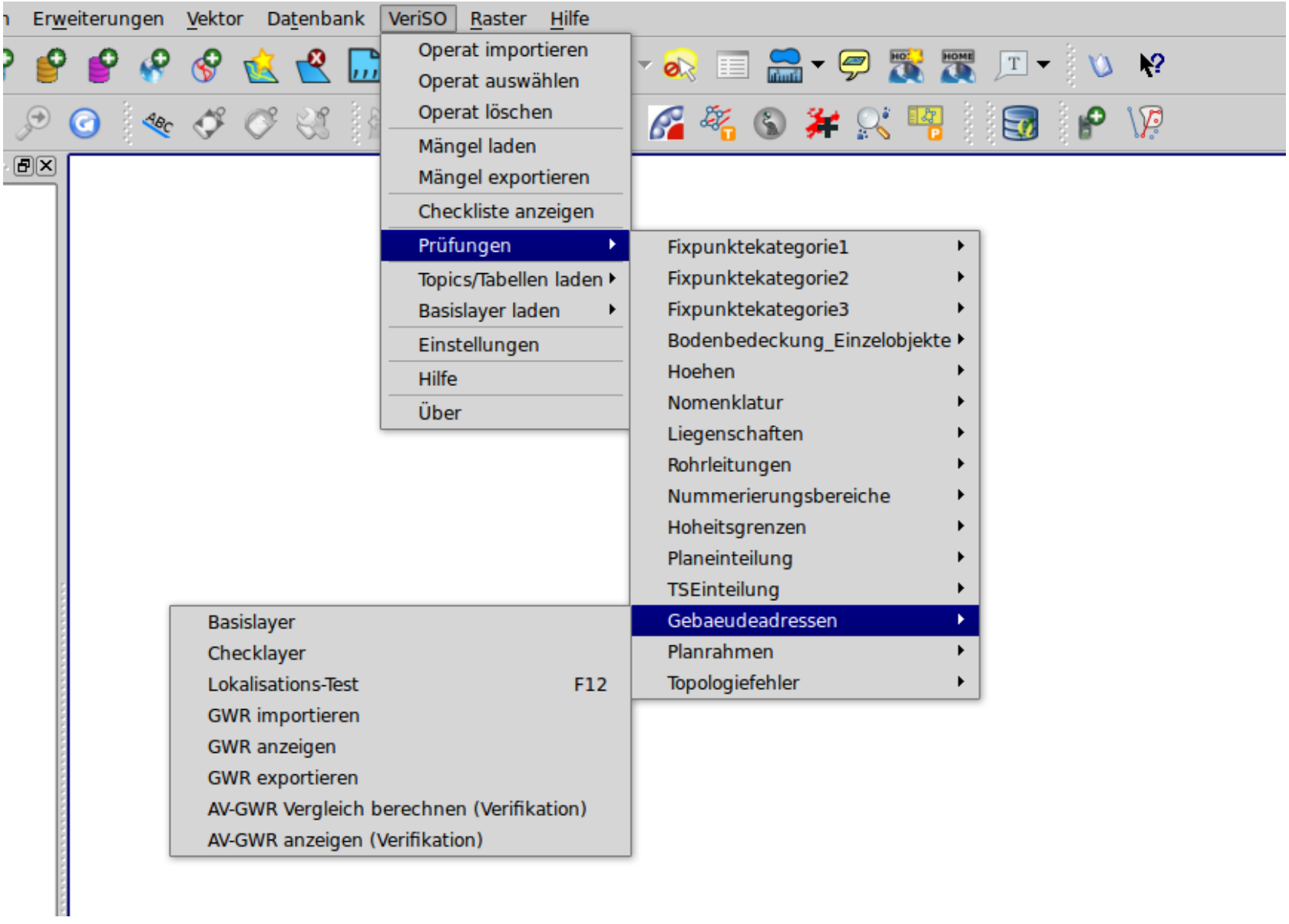
Checks


Lok.Name: Längmatt
TID: 79000031
Num.prinzip: ungeradelinks
Attr. prov.: nein
ist offiziell: ja
Status: real
in Aenderung: nein
Art: Strasse
Import GWR
GWR Datensatz importieren
Text-Datei:
Browse
Import
Abbrechen
Längmatt

Checks: Umsetzung GUI

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<checks>
  <check id="fixpunktekategorie1" title="Übersicht" group="Fixpunktekategorie1"
         topic="Fixpunktekategorie1" type="complex">
    <file>complexchecks.fp1</file>
  </check>
  ...
  <check id="gebaeudeadressen_basis" title="Basislayer" group="Gebaeudeadressen - Basislayer"
         topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_basislayer</file>
  </check>
  <check id="gebaeudeadressen_check" title="Checklayer" group="Gebaeudeadressen - Checklayer"
         topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_checklayer</file>
  </check>
  <check id="gebaeudeadressen_lokalisations" title="Lokalisations-Test"
         group="Gebaeudeadressen - Lokalisations-Test" topic="Gebaeudeadressen" type="complex">
    <file>complexchecks.gebaeudeadressen_lokalisation</file>
    <shortcut>F12</shortcut>
  </check>
  ...
</checks>
```

Checks: Umsetzung GUI



Checkliste

- Möglichst einfach.
- Keine zusätzliche Software.
- HTML-Formulare
- Persistentes Speichern von Daten mittels *Web Storage*.
- Achtung: Daten «nur» im Browser gespeichert!
- Für jedes Operat wird HTML-Datei von Template kopiert.