



Datenqualität als Schlüssel: Bereinigung von Input-Daten für ESRI Utility Network mit FME

Zusammenspiel von ArcGIS und FME in der Praxis



Übersicht

- Warum ist Datenqualität wichtig für ein Utility Network?
 - ArcGIS UN erfordert topologisch saubere, konsistente Daten. Selbst kleine Fehler in den Ausgangsdaten können zu Netzfehlern führen.
- Daher:
 - Der Einsatz von FME als perfektes Tool zur Datenbereinigung und als Integrationsmotor. Durch die nahtlose Einbindung in bestehende Prozesse können auch weitere UN Vorbereitungen in FME getroffen werden.



Anforderungen des ArcGIS Utility Network

- Strukturelle Anforderungen
 - Feature Classes, Subtypes, Domains & Relationship Rules
- Topologische Anforderungen
 - Connectivity Rules, vermeiden von Überlappungen, Gaps & falschen Snaps
- Attributanforderungen
 - Pflichtfelder, eindeutige IDs und korrekte Wertebereiche



Herausforderungen bei Eingangsdaten

- Unterschiedliche Koordinatensysteme
- Fehlende oder doppelte Geometrien
- Uneinheitliche oder leere Attribute
- Leitungen nicht verbunden oder falsch bzw. nicht gesnapped
- Falsche Objektklassenzuweisungen



Vorteile einer Bereinigung in FME

- Einfache Anwendung von komplexen geometrischen Funktionen
- Dateneinblick zu jedem Berechnungsschritt innerhalb eines FME-Workflows möglich – Fehler können direkt erkannt und behoben werden
- Zeitersparnis durch optimierte Algorithmen
- Daten können direkt innerhalb von FME validiert werden
- Reproduzierbar und Automatisierbar



FME-Transformer für Datenbereinigung (und mehr)

- **PointOnLineoverlayer**
- **NeighborFinder**
- **LineOnLineOverlayer**
- **CoordinateExtractor**
- **VertexCreator**
- **LineCombiner**
- **LengthCalculator**
- **Aggregator**
- **Matcher**
- **DuplicateFilter**
- **FeatureJoiner**
- **DuplicateRemover**
- **AreaBuilder**
- **GeometryFixer**
- **LineCleaner**
- **Snapper**
- **ShortestPathFinder**
- **CenterPointReplacer**
- **TopologyBuilder**
- **GeometryValidator**
- **AttributeManager**
- **Tester**



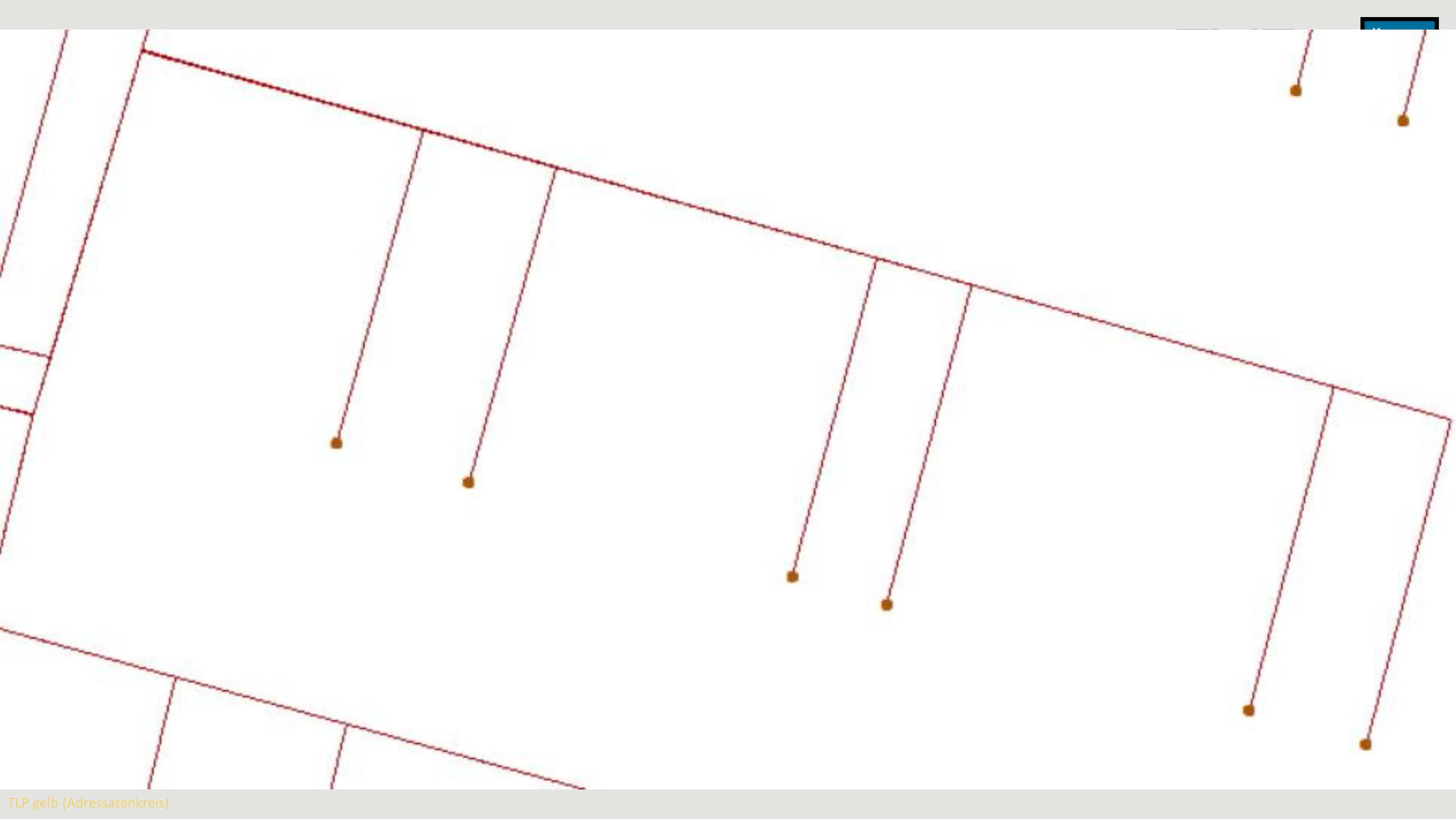
Pipeline-Logik in FME

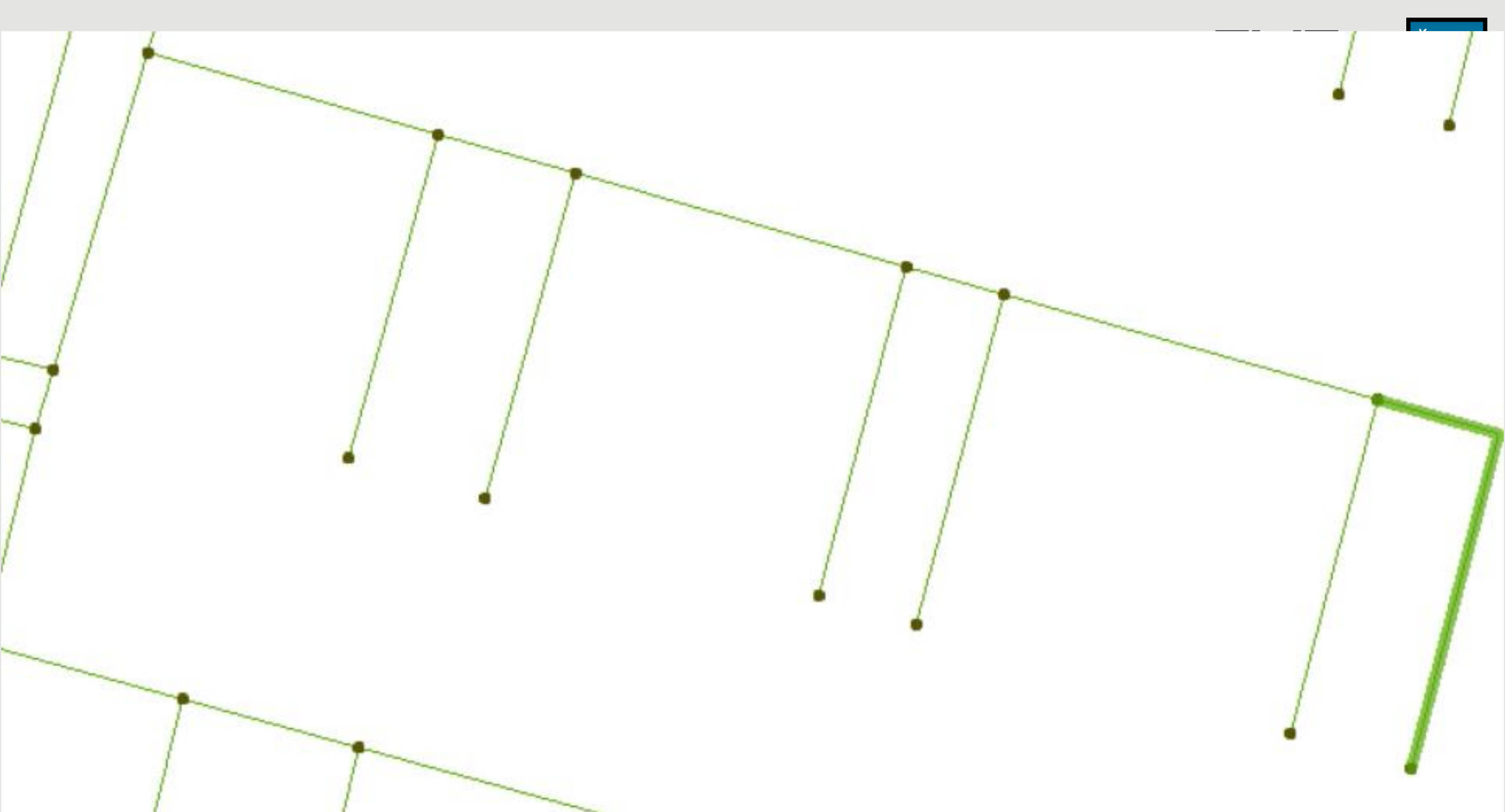
- Input: Einlesen von verschiedenen Datenquellen
- Vorbereitung: Feldnamen anpassen, CRS anpassen
- Geometrische Bereinigung
- Attributbereinigung
- Validierung
- Output: Daten werden geschrieben



Geometrische Bereinigung

- Entfernen von Duplikaten
- Nodes Snappen / Reconnecten
- Topologie wiederherstellen – Clean Lines
- Setzen von Verbindungspunkten

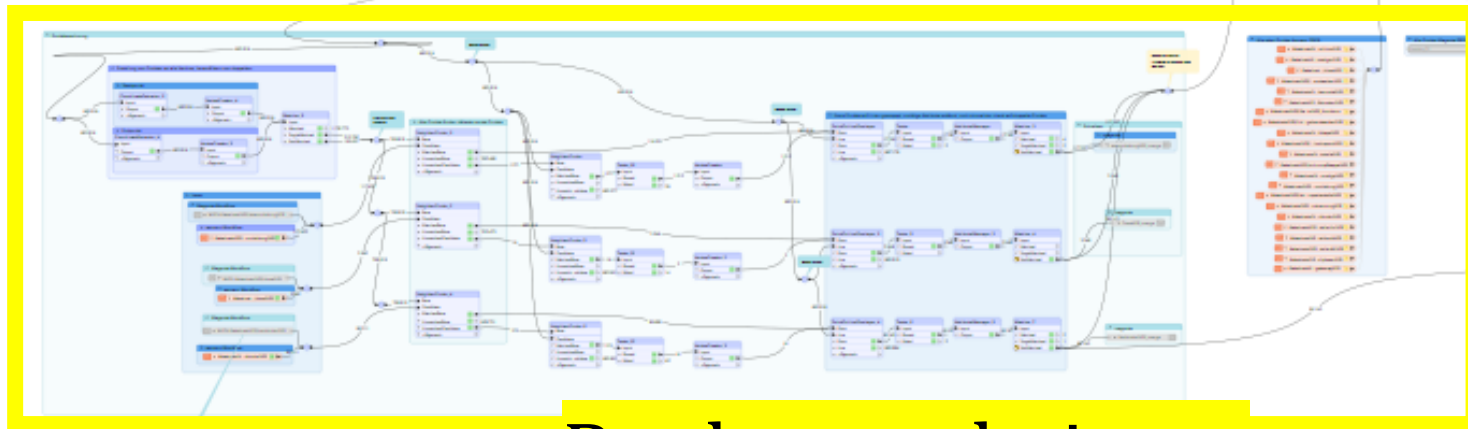
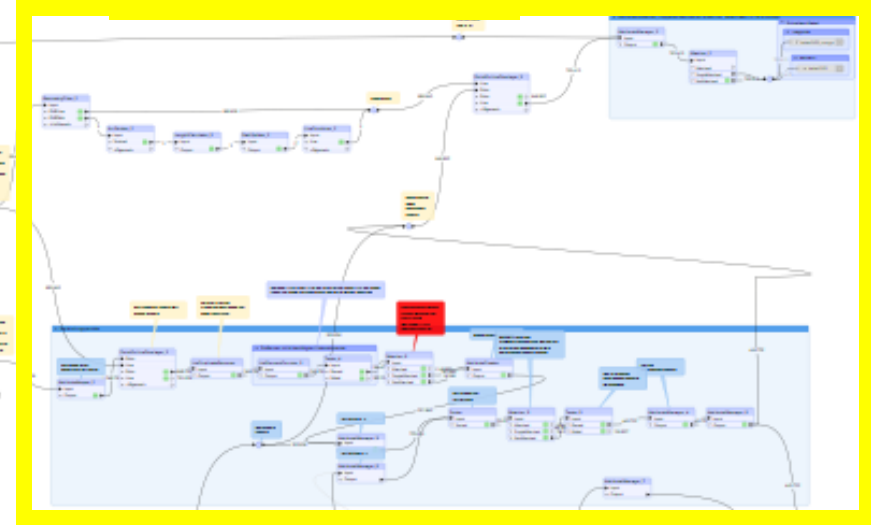
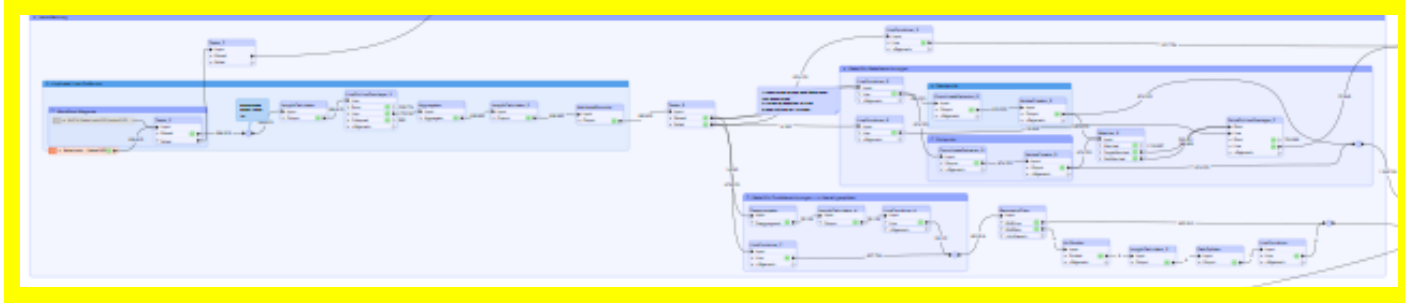






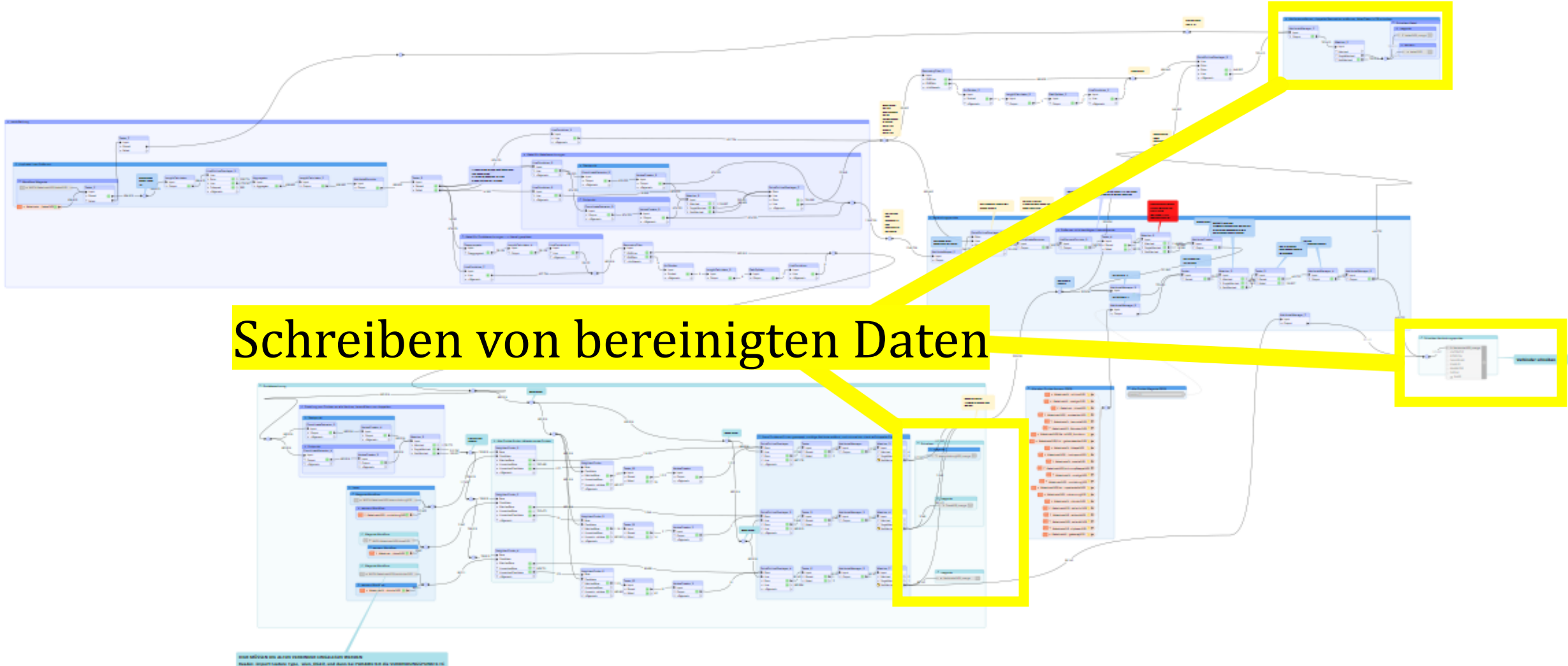
Attribute & Zielschema

Linienverarbeitung

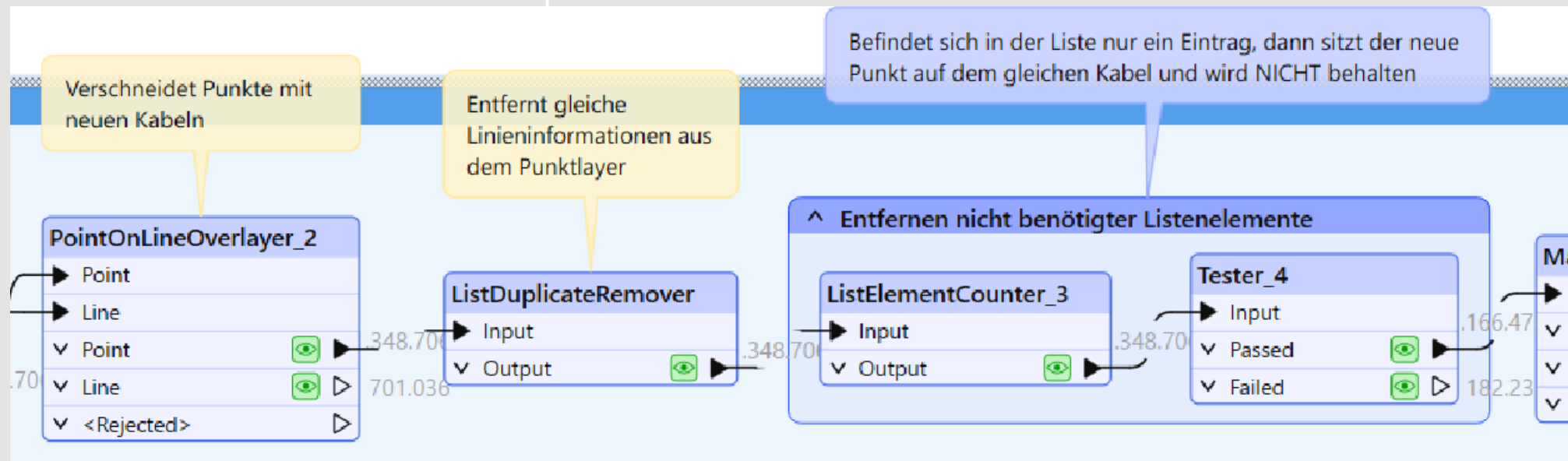
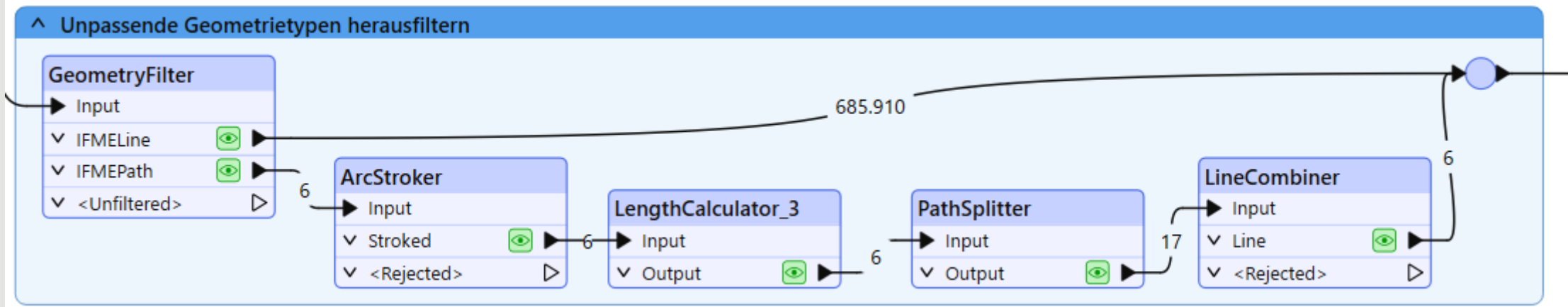


Punktverarbeitung

Das Merkmal des letzten Features (Umschaltzeitpunkt) werden, entsprechend Typen, Wert, Status und anderer Parameter von den Werten des vorherigen Features übernommen.



Schreiben von bereinigten Daten





Laufzeit

Wie lange ein Workspace läuft, hängt von vielen Faktoren ab.

Das gezeigte Beispiel enthält
 ± 700.000 Kabel und ± 700.000 Punkte
und läuft in etwa 32 Minuten komplett durch



Fazit

Datenqualität ist ein Faktor, den wir meist nicht selbst bestimmen können – daher ist zumindest eine Validierung essenziell, auf die eine Bereinigung meist folgt.

FME erlaubt es, individuell angepasste Lösungen für Datenprobleme zu entwickeln, diese performant auszuführen, die Daten direkt zu validieren und ohne weitere Zwischenschritte z.B. in eine SDE Datenbank zurückzuschreiben.