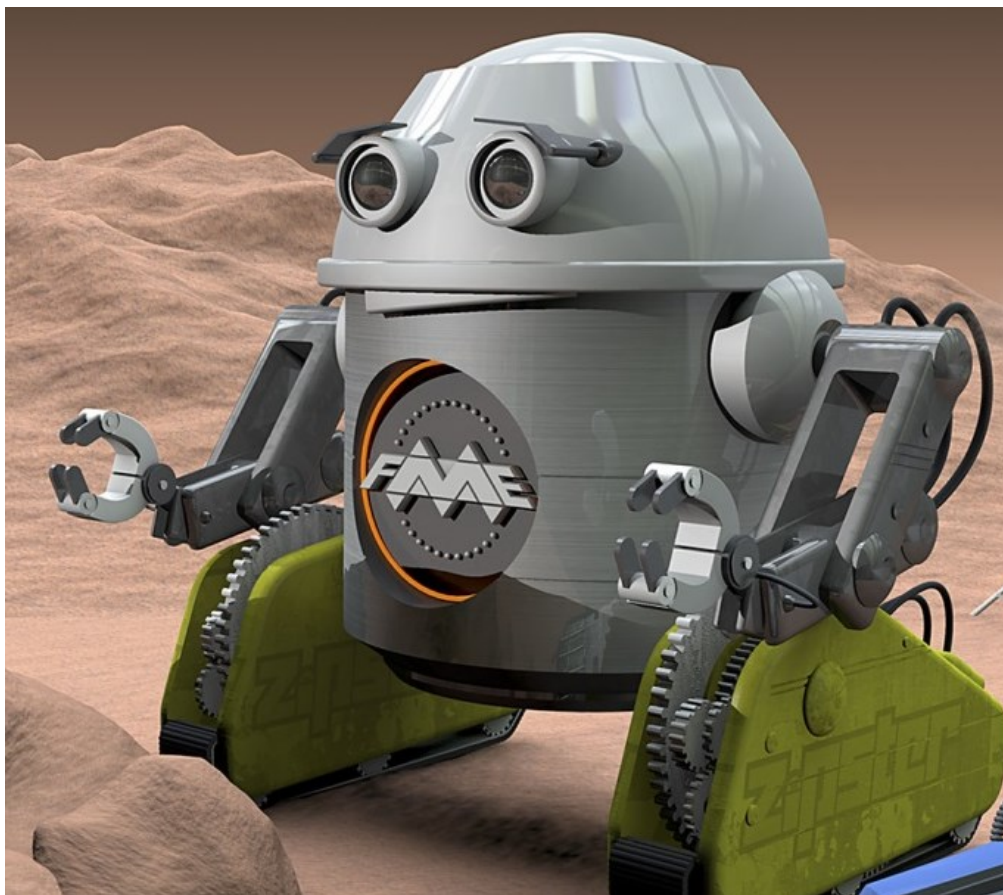


Controletools FME

Rik van Berkum
Geodesie
Hogeschool Utrecht

Afstudeerscriptie



*Versie 4
2 juni 2015*

1 Gegevens afstudeerscriptie

Student

Naam: Rik van Berkum
Adres: J.W. Frisostraat 16, Utrecht
E-mail: rik.vanberkum@student.hu.nl
Telefoon: 0633788996
Studentnummer: 1609428

Studie: Geodesie
Afstudeerrichting: GEO-ICT
Cursuscode: TGEO-AFO8-07
Titel: FME Controletools

Externe begeleider

Naam: Huud Ouwehand
Bedrijf: ProRail
Functie: ProductBeheerder BIS BA–GIS, ICT Services
E-mail: Huud.Ouwehand@ProRail.nl
Adres: Moreelsepark 3, Utrecht

Eerste docentbegeleider: Erik Meerburg
E-mail: Erik.meerbuurg@hu.nl
Tweede docentbegeleider: Elger Heere
E-mail: Elger.Heere@hu.nl

1.1 Changelog

Versie	Datum	Wijzigingen
Conceptversie 1	24 april 2015	---
Conceptversie 2	19 mei 2015	<ul style="list-style-type: none">• Naamgeving conv3rsietools vervangen door controletools• Hoofdstukindeling gewijzigd• Technische details controletools naar bijlage verplaatst• Probleemanalyse toegevoegd• Hoofdstuk <i>Implementatie</i> uitgebreid• Deelvragen duidelijk beantwoord• Opmaak verbeterd• Overige verbeteringen en toevoegingen
Conceptversie 3	26 mei 2015	<ul style="list-style-type: none">• Uitbreiden hoofdstuk <i>Implementatie</i>• Uitbreiden begrippenlijst• Engelse samenvatting• Verbeteren leesbaarheid en taal.• Overige verbeteringen en toevoegingen Bijlage: <ul style="list-style-type: none">• Nieuw hoofdstuk <i>Implementatie</i>• Nieuw hoofdstuk <i>Log</i>
Versie 4	2 juni 2015	<ul style="list-style-type: none">• Voetnoot• <i>Gerefereerde bestanden</i>• Deelvragen uitgebreider beantwoordt• Meer uitleg ter verbetering leesbaarheid Bijlage: <ul style="list-style-type: none">• <i>Implementatie spoortak</i>• <i>Implementatie deelperceel</i>

1.2 Gerefereerde bestanden

[1] Bijlage_controletools.pdf

2 Samenvatting

In de periode van februari tot juni 2015 is de afstudeeropdracht “Controletools FME” uitgevoerd bij ProRail. In deze afstudeerscriptie staat beschreven wat de inhoud en het resultaat is van deze afstudeeropdracht.

ProRail heeft een eigen geo-omgeving. FME is de standaardtool voor het uitvoeren van conversies tussen datasets. Momenteel heeft ProRail ongeveer 500 FME-workbenches in gebruik. Dit grote aantal workbenches is onoverzichtelijk omdat ze niet allemaal eenduidig zijn gebouwd. Workbenches moeten aangepast worden als de datastructuur wijzigt of als blijkt dat ze niet aan de eisen voldoen. Het maken en toevoegen van workbenches om nieuwe gegevens te ontsluiten is een doorlopend proces. Er zijn niet voldoende datacontroles in workbenches. Het bouwen van controles in workbenches gebeurt niet eenduidig en effectief. Dit leidt tot kwaliteitsverlies en tijdsverlies.

Er worden een aantal standaard controletools ontwikkeld die binnen de workbenches hergebruikt kunnen worden. De controletools worden ontwikkeld in FME als custom transformers. Een custom transformer is een reeks aan standaard transformers, samengevat in één enkele transformer (Safe Software, 2014).

De volgende controletools zijn ontwikkeld:

CT_Geocode

De controletool controleert attributen Geocode, Subcode en Geosubcode op juistheid.

CT_GEOM_IN

Valideert de geometrie van originele/ruwe/input datasets.

CT_GEOM_OUT

Valideert de geometrie van bewerkte/output datasets.

CT_GeomType

Controleert features op hun geometrietype.

CT_Measures

Controleert measures met de door FME berekende lengte van het feature.

CT_MissingAttributes

Detecteert null, lege en missende waarden in de opgegeven attributen.

De controletools voldoen aan de eisen van ProRail. Ze hebben een eenduidige naamgeving, documentatie en werking. De controletools zijn getest.

De implementatie bij ProRail is nog niet voltooid. De controletools zijn geïmplementeerd in 5 workbenches in de ontwikkelomgeving. FME gebruikers kunnen de controletools nu gebruiken bij het ontwikkelen van nieuwe workbenches in FME 2015.

2.1 Abstract

This thesis is written over the period from February until June 2015. The thesis is called “Validation tools FME” and is executed in collaboration with ProRail (Dutch rail company).

ProRail sustains her own geo-information. FME is used as the standard tool for performing conversion and data translation between datasets. In the current situation ProRail has about 500 FME-Workbenches in use. This causes a unclear situation because not all workbenches are constructed in the same way and workbenches also need some adjustments and maintenance.

There are no sufficient data validations in workbenches. Data validations are implemented different from workbench to workbench which isn't effective. This leads to degradation of dataquality and loss of time.

Standard validation tools are being developed which can be used inside multiple workbenches. These validation tools will be developed inside FME as custom transformers. A custom transformer is a sequence of standard transformers condensed into a single transformer (Safe Software, 2014).

The next list contains all validation tools which have been build:

CT_Geocode

Validates attributes Geocode, Subcode en Geosubcode.

CT_GEOM_IN

Validates geometry of the original/unedited/input datasets.

CT_GEOM_OUT

Validates geometry of the revised/output datasets.

CT_GeomType

Checks features based on geometry type.

CT_Measures

Checks measures using the by FME calculated length of the features.

CT_MissingAttributes

Detects null, empty and missing attribute values in selected attributes.

The validation tools match the requirements of ProRail. Validation tools have a uniform naming, documentation and operation. The validation tools have been tested using exploratory testing and datatests.

The implementation of the validation tools by ProRail isn't finished yet.

FME users can use the validation tools while developing new workbenches in FME 2015.

When workbenches are converted from FME 2012 to FME 2015 the validation tools can be implemented.

3 Voorwoord

Deze afstudeerscriptie heb ik geschreven als afronding van mijn studie Geodesie/Geo-informatica aan de Hogeschool Utrecht. De afstudeeropdracht heb ik uitgevoerd in samenwerking met ProRail te Utrecht.

Tijdens het tweede jaar van mijn opleiding heb ik als onderdeel van het vak GEO-ICT een practicum FME gedaan. In het derde jaar van mijn studie heb ik twee verschillende stages gedaan. Mijn eerste stage was toegespitst op de landmeetkunde. Mijn tweede stage toegespitst op GIS en kartografie. Tijdens mijn tweede stage bij waterschap Aa en Maas heb ik mijn eerste praktijkervaring opgedaan met FME. Ik ben op de FME World tour 2014 geweest en heb een tweedaagse FME desktop basic training gevolgd bij Vicrea.

Voor mijn afstudeeropdracht heb ik uiteindelijk een bewuste keuze gemaakt voor een onderwerp met FME. FME spreekt mij aan omdat het een sterk automatiseringsaspect heeft en krachtig is in ETL en geo-informatie.

Het uitvoeren van mijn afstudeeropdracht heb ik als leerzaam ervaren. Mijn kennis van FME is natuurlijk enorm gegroeid. Ook zijn mijn onderzoeks- en rapportagevaardigheden verbeterd. Wat ik niet had verwacht is dat ik ook kennis opgedaan heb opgedaan van Python en RegEx.

Ik wil de betrokken collega's bij ProRail bedanken voor het ondersteunen en begeleiden van mijn afstudeeropdracht.

Ook dank aan Erik Meerburg, mijn afstudeerbegeleider van de Hogeschool Utrecht, voor de goede feedback over mijn afstudeerscriptie.

Utrecht, juni 2014.

Rik van Berkum

4 Inhoud

1	Gegevens afstudeerscriptie	2
1.1	Changelog	3
1.2	Gereferende bestanden	3
2	Samenvatting	4
2.1	Abstract	5
3	Voorwoord	6
4	Inhoud	7
5	Inleiding	9
5.1	Aanleiding	9
5.2	Probleemanalyse	9
5.3	Probleemstelling	10
5.4	Belang van controletools	10
5.5	Centrale vraagstelling	11
5.6	Deelvragen	11
6	Theoretische achtergrond	12
6.1	ETL	12
6.2	FME Desktop	12
6.3	FME Server	12
6.4	Huidig gebruik FME en FME Server	12
6.5	Reader en Writer	12
6.6	Transformer	13
6.7	Custom Transformers	13
6.8	Python/TCL en FME	14
6.9	RegEx en FME	15
6.10	FME World Tour	15
7	Eisen controletools	16
7.1	Eenduidige Controletools	16
7.2	Naamgeving	16
7.3	Documentatie	17
7.4	Opbouw	18
7.5	Tests	20
7.6	Conclusie	22

8	Controletools	23
8.1	CT_Geocode	24
8.2	CT_GEOM	25
8.3	CT_GeomType	26
8.4	CT_Measures	27
8.5	CT_MissingAttributes	28
8.6	Hoe zijn de controletools aan de eisen afgestemd?	29
9	Implementatie	30
9.1	FME Desktop	30
9.2	FME Server	30
9.3	Versiebeheer	31
9.4	Implementatie	32
9.5	Resultaten implementatie	33
9.6	Datakwaliteit en tijdsbesparing.....	35
10	Conclusie	37
11	Aanbevelingen.....	38
12	Begrippen en afkeringen	39
13	Bronnen afstudeeronderzoek	42

5 Inleiding

Voor de periode van februari tot juni 2015 is deze afstudeeropdracht uitgevoerd bij ProRail. In deze afstudeerscriptie staat beschreven wat de inhoud en het resultaat is van deze afstudeeropdracht.

5.1 Aanleiding

ProRail heeft een eigen geo-omgeving: een beheeromgeving en een presentatieomgeving. Deze omgevingen bestaan uit verschillende databases. Er vindt dataverkeer plaats tussen de verschillende databases en er gaat data naar viewers en externen.

Om deze data en processen goed te beheren heeft ProRail ongeveer 500 FME-workbenches in gebruik. In een workbench staat de datatransformatie beschreven. Een workbench kan geopend worden in FME en geeft een grafische weergave van datatransformatie. Het grote aantal workbenches is onoverzichtelijk en een aantal workbenches moet nog verbeterd worden. Het toevoegen van workbenches om nieuwe gegevens te ontsluiten is een doorlopend proces. Er is behoefte aan standaardfunctionaliteit die binnen de workbenches hergebruikt kan worden.

5.2 Probleemanalyse

Voordat er een workbench wordt gebouwd, maakt de informatiespecialist een ontwerp. In het ontwerp staan de op te leveren tabellen en de structuur van de dataconversie beschreven. Bij het bouwen van een workbench wordt echter ter plekke bedacht hoe de workbench ingericht moet worden om aan het ontwerp te voldoen.

De datakwaliteit wordt in workbenches niet of alleen ad hoc gecontroleerd. Fouten worden daardoor vaak bij toeval geconstateerd.

Als er controles plaatsvinden in workbenches gebeurt dat niet op gelijke wijze. Hierdoor is hergebruik van de functionaliteit moeizaam.

Soms is dezelfde functionaliteit ingebouwd in meerdere workbenches. Als blijkt dat de functionaliteit niet aan de eisen voldoet, moet het aangepast worden in elke workbench waarin deze functionaliteit voorkomt.

Bovenstaande problemen zorgen voor kwaliteitsverlies in de geo-informatie van ProRail en tijdsverlies bij het maken en onderhouden van workbenches.

Als onderdeel van de probleemanalyse is onderzocht wie de stakeholders zijn bij mijn afstudeeropdracht. Betrokken stakeholders hebben de volgende functies binnen ProRail:

- FME-Consultants: Zij bouwen actief aan huidige en nieuwe workbenches.
- Informatieanalist: De informatieanalist analyseert de klantwens. Op basis hiervan wordt een ontwerp gemaakt. In het ontwerp staat de mapping van de brondata, de op te leveren tabellen en de structuur van de dataconversie beschreven. Het systeemconcept geoprocessing, dat door de informatieanalist wordt onderhouden, is hierbij leidend.
- Applicatiebeheerder: De applicatiebeheerder is verantwoordelijk voor de techniek. Hij moet er bijvoorbeeld voor zorgen dat FME Server blijft draaien.
- Functioneel beheerder: De schakel tussen de applicatie en gebruiker. Hij is ervoor verantwoordelijk dat de gebruiker de applicatie kan blijven gebruiken.
- Gebruikers:
 - Ongeveer 150 GIS gebruikers (ProRail-medewerkers).
 - Systemen die data afnemen en mapservices gebruiken: Railmaps(GIS viewer voor alle ProRail-medewerkers), BBMS (monitoring van onderhoudsconditie), Sigma(spoorgeometrie), SpoorWeb. Van deze systemen maken ook externen gebruik zoals aannemers.

5.3 Probleemstelling

Er zijn niet voldoende datacontroles in workbenches. Het bouwen van controles in workbenches gebeurt niet eenduidig en effectief. Dit leidt tot kwaliteitsverlies van data.

5.4 Belang van controletools

Het volgende artikel van Kenneth Willems beschrijft het belang van geo-informatie bij ProRail:

ProRail is verantwoordelijk voor het spoorwegennetwerk van Nederland. ProRail zorgt derhalve voor het aanleggen, het onderhouden en het beheren van dit spoorwegennetwerk. Dit netwerk beslaat ruim 6.500 kilometer spoor en wordt bereden door ruim 6.500 treinen. De veiligheid hierbij is van zeer groot belang: elke dag verplaatsen zich ongeveer 1.200.000 reizigers en zo'n 100.000 ton goederen over dit spoor. Om dit alles in goede banen te leiden moet er heel wat worden geregeld. ProRail verdeelt de beschikbare ruimte op het spoor, regelt het treinverkeer en voorziet alle betrokkenen (zowel reizigers als vervoerders) van relevante informatie. Daarnaast worden telkens nieuwe ontwikkelingen doorgevoerd, zoals de aanleg en onderhoud van nieuw spoor, het bouwen en beheren van stations en het onderhouden van allerlei bestaand materieel zoals wissels, seinen en overwegen.

Om alles goed te kunnen beheren, om het spoor in stand te kunnen houden en om het plannen van toekomstige infrastructuur te kunnen faciliteren maakt ProRail volop gebruik van kaartmateriaal. Naast infrastructuurobjecten wordt dit kaartmateriaal gebruikt voor het kunnen voldoen aan wet- en regelgeving op het gebied van milieu en veiligheid op en rond het spoor. Daarnaast zijn er externe partijen/klanten die gebruik maken van beschikbaar kaartmateriaal van ProRail. ProRail is belast met het beheer en onderhoud van dit kaartmateriaal. (Willems, K. 2011)

Controletools zijn van belang voor een organisatie als ProRail. Controletools controleren de geo-informatie die ProRail beheert met FME. Door de controles die de controletools

uitvoeren ontstaat kwaliteitswinst in de data. Hier hebben betrokken stakeholders profijt van.

5.5 Centrale vraagstelling

Hoe worden FME en FME Server momenteel gebruikt en kan dit effectiever door het maken van controletools?

5.6 Deelvragen

In deze afstudeerscriptie wordt antwoord gegeven op de volgende deelvragen:

- Wat zijn de mogelijkheden en beperkingen van custom transformers?
- Aan welke eisen moeten de controletools voldoen?
- Hoe zijn de controletools aan de eisen afgestemd?
- Hoe kunnen de controletools geïmplementeerd worden?
- Hoe kunnen de controletools zorgen voor kwaliteitswinst en tijdswinst?



6 Theoretische achtergrond

6.1 ETL

ETL staat voor Extract, Transform en Load. Extract is het ophalen van data. Transform staat voor het converteren van de data. Load staat voor het wegschrijven van de data.

6.2 FME Desktop

FME Desktop is een product van Safe software Inc. FME Desktop is een ETL applicatie en wordt gebruikt voor geautomatiseerde datatranslaties, dataconversies en datavalidatie van geo-informatie. FME is marktleider ETL tools. FME heeft een goede grafische weergave en uitgebreide documentatie.

Met “FME” wordt vaak “FME Desktop” bedoeld, afhankelijk van de context van de zin.

6.3 FME Server

FME Server is een ruimtelijke ETL-applicatie voor translatie en transformatie van geo-informatie over een netwerk. Hierdoor kunnen organisaties hun data bekijken, herstructureren en distribueren. (Safe Software, 2014)

6.4 Huidig gebruik FME en FME Server

FME is de standaardtool voor het uitvoeren van conversies tussen datasets. Alle conversies worden met FME gebouwd en ingeregeld. Voor het uitvoeren van de conversies is FME-server in gebruik. Met FME-server worden de conversies ingeregeld: ze worden op vaste tijdstippen met een vaste frequentie uitgevoerd.

Bestaande conversies die gebruik maken van andere programmatuur (voornamelijk Oracle) zullen naar FME gemigreerd worden. Alleen in gevallen waar het gebruik van FME te ingewikkeld is of te veel tijd kost om te bouwen, kan hiervan afgeweken worden.

(Afwijkingen worden met redenen omkleed in het ontwerp vastgelegd en besproken met applicatiebeheer en functioneel beheer)

ProRail (2014). Systeemconcept Geoprocessing.

Hoe worden FME en FME Server momenteel gebruikt?

FME wordt gebruikt om conversies (van geo-informatie) te bouwen.

FME Server wordt gebruikt om de conversies op vaste tijdstippen uit te voeren.

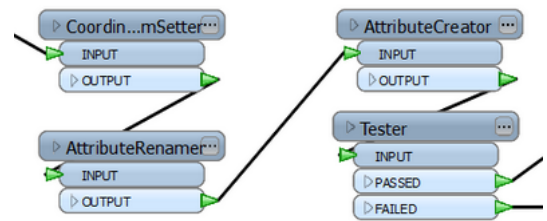
6.5 Reader en Writer

De Reader in FME leest de brondata, het formaat en de locatie.

De Writer in FME specificeert het formaat en locatie van de output-data.

6.6 Transformer

Een transformer is een object dat een bewerking uitvoert op features. Een feature is een object, bijvoorbeeld een record in een database inclusief attributen en geometrie. Een transformer wordt meestal ingezet tussen de reader en de writer in. FME bevat meer dan 500 verschillende transformers. Transformers zijn gecategoriseerd in de transformer Gallery in FME Desktop.



Figuur 1 Transformers

6.7 Custom Transformers

De controletools worden ontwikkeld in FME als custom transformers. Een custom transformer is een reeks aan standaard transformers, samengevat in één enkele transformer (Safe Software, 2014). Ze kunnen opgeslagen worden in een apart .fmx bestand. In een fmx bestand staat de definitie van de custom transformer opgeslagen. Dit bestand kan vervolgens gebruikt worden om de custom transformer te kunnen delen met andere FME gebruikers.

Wat zijn de mogelijkheden en beperkingen van custom transformers?

De mogelijkheden van custom transformers zijn bijna gelijk aan de mogelijkheden die een reeks van transformers hebben. Custom transformers kunnen geen Reader of Writer gebruiken en de user-parameter mogelijkheden zijn beperkt.



6.8 Python/TCL en FME

Tijdens het ontwikkelen van de controletools kan Python of TCL gebruikt worden in FME.

Python of TCL biedt uitkomst als de grafische interface van FME te beperkt is.

TCL staat voor Tool Command Language. Dit is een programmeertaal die uitsluitend in FME is te gebruiken.

Python is een populaire open-source programmeertaal.

Beide programmeertalen ondersteunen OOP(Object Oriented Programming).

Python is ook in andere programma's te gebruiken. Python-kennis die is opgedaan door het werken met bijvoorbeeld een GIS en Python, kan vervolgens helpen bij het gebruik van Python en FME. Daarom kan het beste Python gebruikt worden als programmeertaal binnen FME.

PythonCaller

De PythonCaller kan Pythonscripts uitvoeren en de FME objects module voor Python gebruiken. Een PythonCaller kan eigenlijk een volledige workbench vervangen. En Python kan zorgen voor extra functionaliteit in FME.

Het nadeel van een PythonCaller is dat het niet inzichtelijk en grafisch wordt weergegeven wat er gebeurt in de workbench. De workbench kan alleen worden onderhouden door iemand met kennis van Python **en** FME. Bij het scripten is het minder inzichtelijk of alle parameters goed gebruikt worden en of het script uitvoert wat de bedoeling is. FME transformers hebben een controle dat alle parameters zijn ingevuld en gedefinieerd. Deze automatische controle is niet mogelijk in een PythonCaller.

FME transformers	PythonCaller
Grafische interface	Script
Interface beperkt mogelijkheden en functionaliteit	Veel mogelijkheden en functionaliteit
Inzicht in dataconversie en datastromen.	Slecht inzichtelijk. Kans op fouten aanzienlijk groter.
Makkelijk te onderhouden	Lastig te onderhouden
Snel workbenches bouwen	Tijdrovend (Tenzij FME interface te beperkt is, dan is het sneller een pythonscript te schrijven dan een workaround te maken met FME transformers.
Alleen kennis van FME vereist	Kennis van Python en FME vereist

De voordelen van de PythonCaller wegen niet op tegen het voordeel van FME transformers. Het gebruik maken van de PythonCaller en Pythonscript kan daarom het best zoveel mogelijk vermeden worden. Behalve als de FME interface te beperkt is voor het doel wat bereikt moet worden en een workaround veel tijd gaat kosten. In dat geval biedt Python een uitkomst.

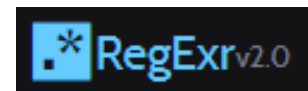
6.9 RegEx en FME

Sommige controletools gebruiken RegEx om attribuutwaardes te controleren.

RegEx staat voor Regular Expression. Het is een handig hulpmiddel voor het vinden en valideren van tekst. Regular expression is een tekst patroon dat gebruikt kan worden in veel modern applicaties en programmeertalen. Dit kan gebruikt worden om input te toetsen of om tekstpatronen te vinden. (Goyvaerts en Levithan, 2012).

RegEx wordt in veel programmeertalen ondersteund. Voor het valideren van data is RegEx een handig hulpmiddel in FME. De StringSearcher-transformer kan een regular expression uitvoeren op attributen.

Om te oefenen met het schrijven en de werking van RegEx, is de simulator op www.regexr.com een handig hulpmiddel om te gebruiken.



Figuur 2 RegExr

6.10 FME World Tour

In kader van deze afstudeeropdracht is de FME World Tour bezocht op 18 mei 2015 in Amersfoort. De FME World Tour is een bijeenkomst van FME gebruikers waarbij uitgewisseld wordt hoe FME ingezet kan worden. Ken Bragg van Safe Software was ook aanwezig en vertelde wat de nieuwe ontwikkelingen zijn voor FME. Op deze dag zijn handige tips gegeven om sneller en effectiever met FME te werken. Die tips zijn gebruikt tijdens het ontwikkelen van de controletools.

Bijvoorbeeld een nieuwe parameter waarbij aangegeven kan worden dat de eerste 100 features uit de database moeten worden gelezen. Deze parameter scheelt tijd bij het testen, omdat dan niet de volledige database wordt ingelezen in FME.

7 Eisen controletools

De volgende eisen en eigenschappen van de controletools zijn opgesteld door gesprekken met FME-gebruikers en door de controletools zelf veelvuldig te gebruiken in FME.

7.1 Eenduidige Controletools

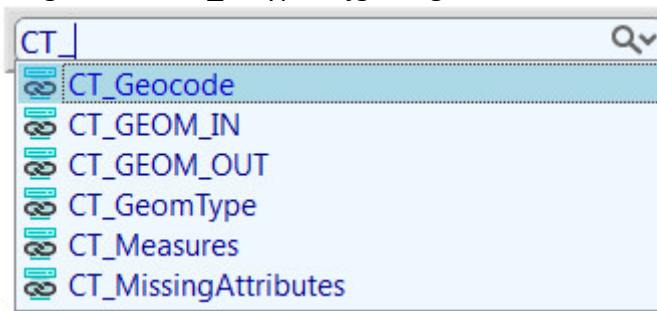
Het is belangrijk dat de controletools een eenduidige naamgeving, werking en documentatie hebben. Dit verbetert het gebruik en het overzicht van de controletools.

7.2 Naamgeving

De naamgeving die binnen ProRail gebruikt zal worden is conv3rsietool. Voor deze scriptie zal ik echter de naamgeving controletool gebruiken, omdat die naam beter overeenkomt met de werking van de controletools.

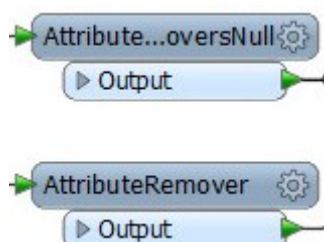
De naamgeving van de controletools begint telkens met “CT_”.

Het toevoegen van een transformer gebeurt vaak door de transformernaam in te typen. Als een gebruiker CT_ intypt krijgt de gebruiker alle controletools in beeld:



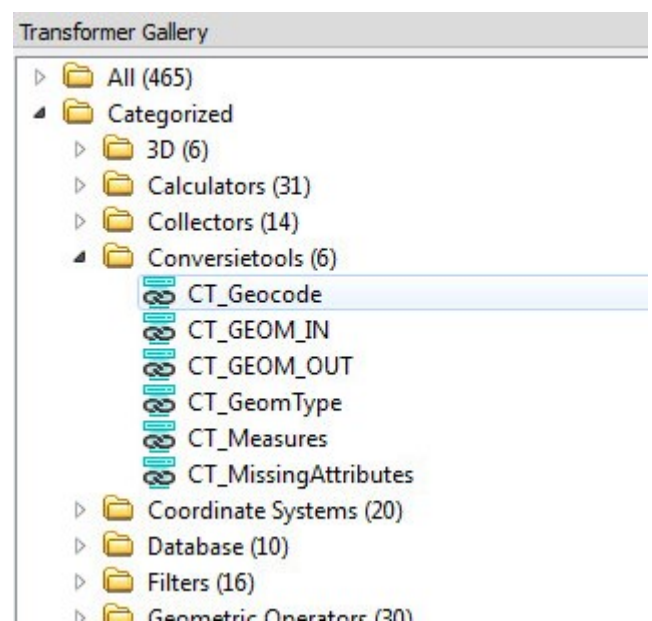
Figuur 3 Controletools zoeken

De namen van de controletools zijn zo gekozen dat ze niet te lang zijn en de functie van de controletool beschrijven. De naam moet in de transformer passen, anders wordt de transformernaam onleesbaar:



Figuur 4 Transformernaam

De Transformer Gallery is een apart venster in FME Desktop met een overzicht van transformers. Controletools worden beschikbaar gemaakt in de Transformer Gallery onder de categorie Controletools. Hierdoor zijn ze ook te vinden door naar de controletool-categorie te bladeren:



Figuur 5 Transformer Gallery

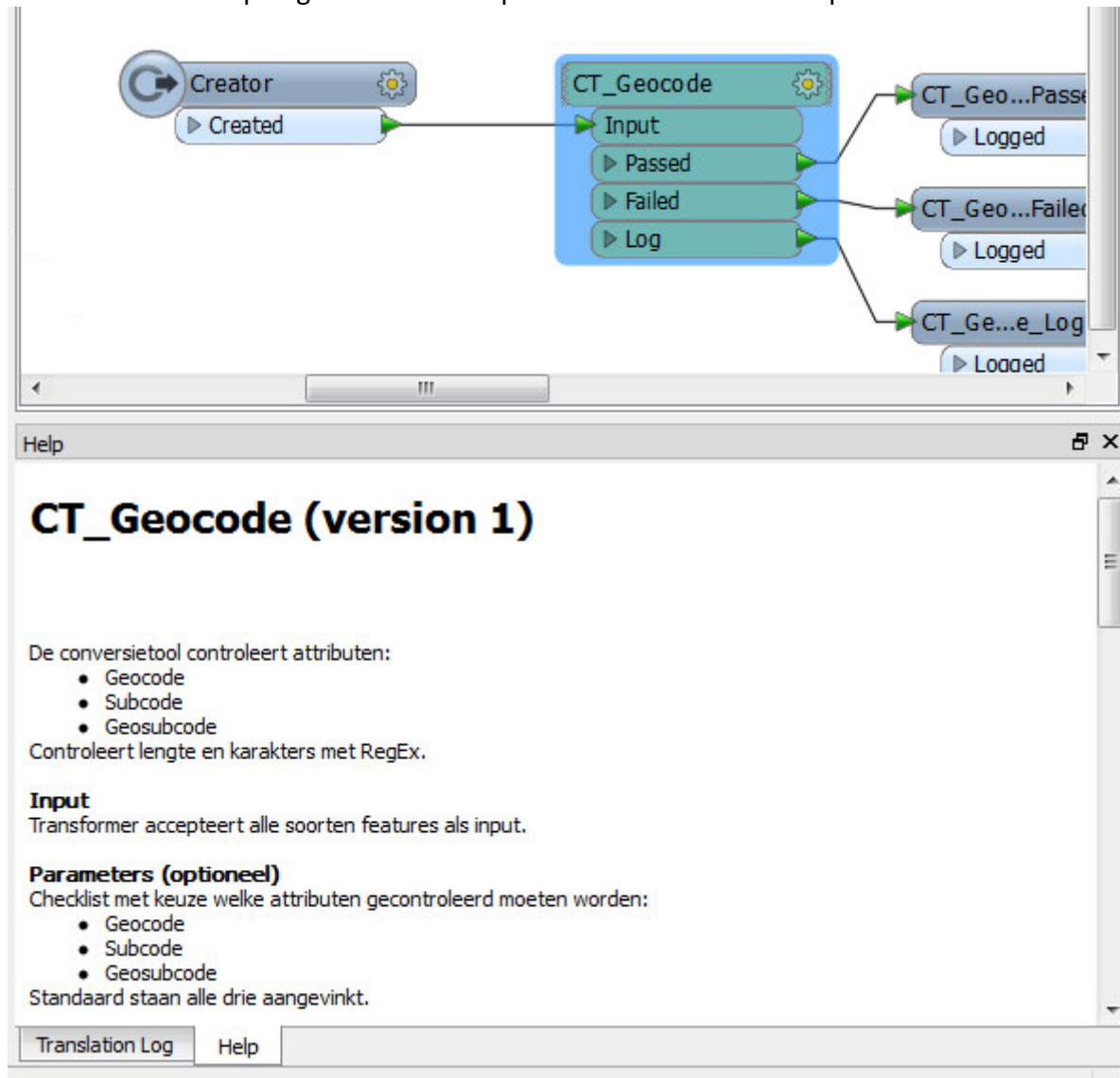
7.3 Documentatie

De controletools zijn eenduidig gedocumenteerd.

De documentatie heeft telkens de volgende opbouw:

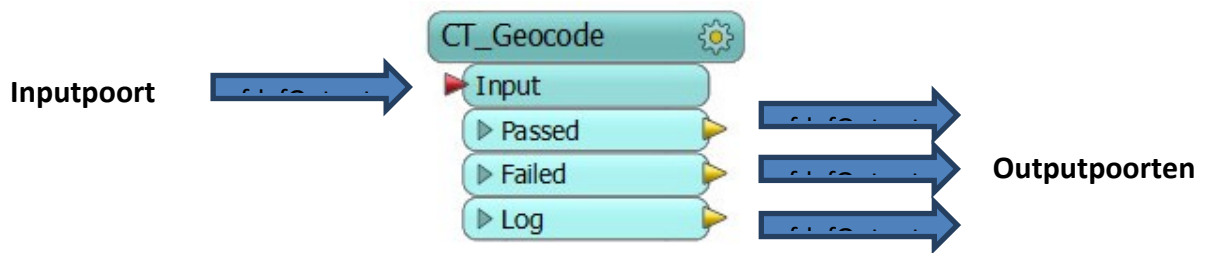
- Korte beschrijving
- Input
- Parameters
- Output

De documentatie is in de description van de controletool vastgelegd. Daardoor is de documentatie raadpleegbaar via het help-window van FME Desktop:



Figuur 6 Help-venster van FME Desktop

7.4 Opbouw



Controletools hebben altijd als outputpoort: Passed, Failed en Log.

Deze outputpoorten hebben de volgende gemeenschappelijke eigenschappen:

Passed: Features die voldoen aan de test van de controletool. Features die uit de passed-poort komen hebben altijd hetzelfde formaat als de input van de controletool.

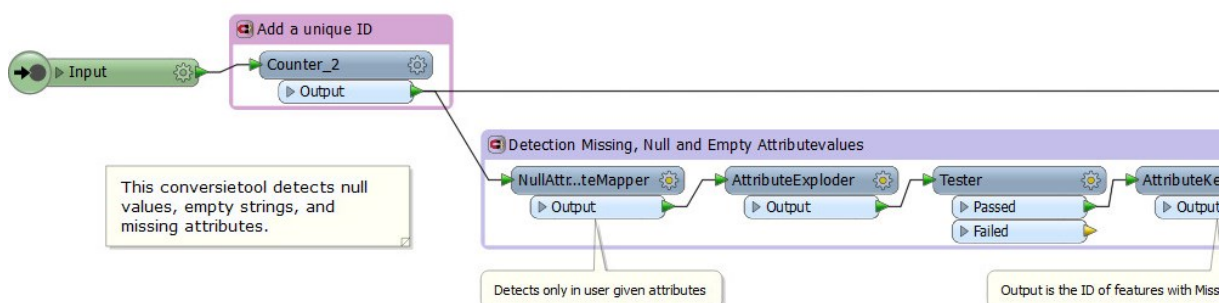
Failed: Features die niet voldoen aan de test van de controletool. Features krijgen een nieuw attribuut 'Fouten' met een beschrijving van de oorzaak van de fout. Deze poort kan gebruikt worden om fouten te inspecteren en weg te schrijven naar een shapefile.

Log: Bevat slechts 1 attribuut met een beschrijving van de oorzaak van de fout. Deze poort kan aangesloten worden op een Logger of Tekst-writer. (Zie bijlage 6)

Als er door de controletool aangemaakte attributen komen uit de output van de controletool, wordt dit altijd gedocumenteerd. De Passed-poort van de controletool heeft altijd hetzelfde formaat input als output. Dat is belangrijk omdat data uiteindelijk naar de Oracle database geschreven wordt. In de Oracle database staat data in een vooraf bepaalde en gedocumenteerde structuur. Een controletool voert alleen controles uit op data en geen conversies (tenzij anders aangegeven). Bij de implementatie van de controletool is het een eis dat de controletool niks aan datastructuur verandert.

De grootte van de workbench van de controletools is een afweging tussen functionaliteit, complexiteit en hanteerbaarheid van de controletool. Tijdens het ontwikkelen van de controletools is het uitgangspunt om zoveel mogelijk functionaliteit te bouwen, met zo min mogelijk transformers.

De workbench van de controletool is altijd overzichtelijk ingericht. Hierdoor is de controletool goed te begrijpen en te onderhouden. Voor extra overzicht en verduidelijking zijn bij iedere workbench Bookmarks en Annotations toegevoegd:



De Bookmarks en Annotations zijn allemaal in het Engels geschreven, omdat FME ook volledig in het Engels geschreven is. Tijdens het gebruik van FME is het duidelijker om alleen Engels te gebruiken. Als Engels en Nederlands door elkaar worden gebruikt kan onduidelijkheid ontstaan en moet er vaak twee kanten op vertaald worden.

Parameters

De parameters van de controletools zijn optioneel of verplicht

- **Verplicht:** Bij het verplicht maken van een parameter, wordt de gebruiker verplicht om de parameter in te stellen. Als de gebruiker de workbench draait zonder de parameter te hebben ingesteld krijgt de gebruiker een waarschuwing “Invalid Transformer Parameters”.
- **Optioneel:** Het is niet verplicht om de parameter in te stellen. De parameter heeft een default(standaard)-waarde. Het is dan wel verstandig dat de gebruiker kijkt of de deze standaardwaarde geschikt is voor zijn of haar doeleinden.

7.5 Tests

Een controletool moet altijd daadwerkelijk uitvoeren wat beschreven is in de documentatie. In de documentatie staat welke fouten de controletool detecteert. Een FME-gebruiker moet ervan uit kunnen gaan dat de controletools daadwerkelijk alle beschreven fouten uit data haalt. Om dit te controleren zullen de controletools getest worden. Deze tests worden gedocumenteerd om te bewijzen dat de controletool getest is en werkt.

Hier volgen twee testmethodes die gebruikt worden om de controletools te testen:

Datatests

Datatests is een methode waarbij testdata wordt gegenereerd. In de testdata worden fouten aangebracht die de controletool zou moeten kunnen detecteren. Deze testdata wordt ingelezen in FME en na de translatie moeten alle fouten in de testdata gedetecteerd worden. Het kost veel tijd om alle mogelijke fouten in testdata te genereren. Daarom moeten de fouten die getest worden de randen van het domein omsluiten. Als een veld bijvoorbeeld uit 3 cijfers moet bestaan, moet een veld aangemaakt worden met 4 cijfers. Als deze wordt gedetecteerd hoeft niet getest te worden wat gebeurt bij velden met 5, 6 of meer cijfers.

Exploratory Testing

Exploratory testing of ad hoc testing is het tegelijkertijd leren, tests ontwikkelen en tests uitvoeren. (Bach, 2003)

Deze methode wordt veelvuldig toegepast bij het ontwikkelen van de controletools. Voordat begonnen wordt met het ontwikkelen van een controletool, wordt eerst een voorbeelddataset opgezocht of gegenereerd. Deze dient als input van de controletool. Tijdens het ontwikkelen worden de tests uitgevoerd op de controletool. Uit deze tests blijkt dat de controletool nog niet aan de eisen voldoet en wordt de controletool aangepast. Door het veelvuldig testen kunnen ook tekortkomingen in de tests ontdekt worden. Tijdens het ontwikkelen breiden de controletool en de tests dus steeds verder uit.

Wanneer de controletool helemaal uitontwikkeld is en ad hoc getest, worden er datatests uitgevoerd op de controletools. Deze datatests staan in aparte workbenches en samen met de gedocumenteerde testresultaten dient dit als bewijs dat de controletool goed werkt.

Unit testing

In de ICT wordt Unit Testing als betrouwbare test gezien. Dit zijn echter tests die bedoeld zijn voor code. Met FME wordt gebruik gemaakt van een interface en Safe Software voert de code uit. Hierdoor kan geen gebruik gemaakt worden van de methode unit testing.

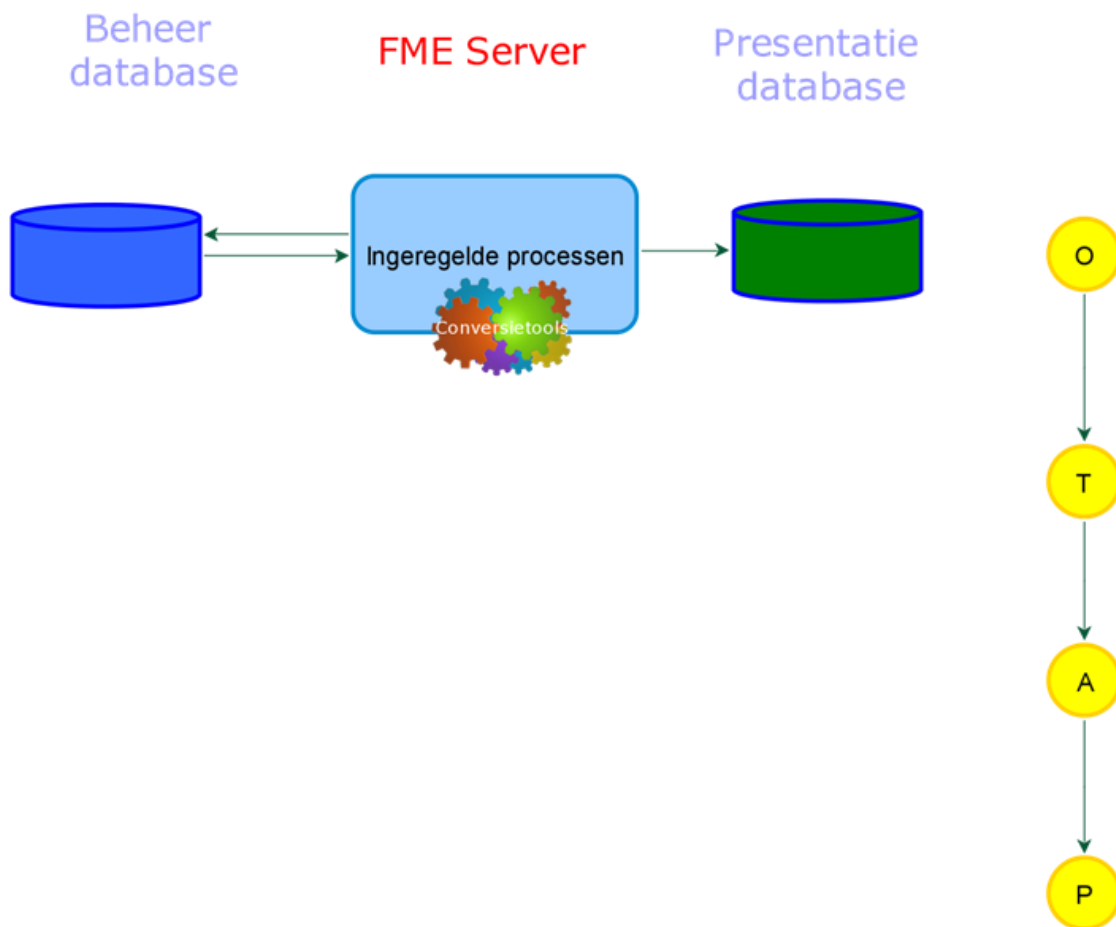
OTAP

ProRail maakt gebruik van OTAP-systematiek. OTAP is een methode die veel wordt gebruikt in de ICT en staat voor: Ontwikkeling, Test, Acceptatie en Productie.

Een controletool wordt als eerste gebruikt tijdens het ontwikkelen in de ontwikkelomgeving. Vervolgens zal de controletool op FME Server gezet worden in de testomgeving. In de testomgeving kan de controletool getest worden met data van ProRail.

In de acceptatieomgeving wordt bekeken of de functionaliteit van de workbench met de controletool naar wens is. Als dat zo is zal de controletool in productie genomen worden in de productieomgeving.

De volgende figuur illustreert waar de controletools in de OTAP-omgevingen staan:



7.6 Conclusie

Aan welke eisen moeten de controletools voldoen?

Eenduidige naamgeving

Door een eenduidige naamgeving zijn de controletools snel vindbaar en goed te herkennen in FME.

Eenduidige documentatie

De documentatie is snel raadpleegbaar. Er staat duidelijkheid beschreven wat de werking van de controletools is en waarvoor de parameters zijn.

Eenduidige opbouw

Doordat de controletools overzichtelijk zijn opgebouwd is de werking van de controletool duidelijk.

Gebruik verplichte of optionele parameters

Door gebruik van verplichte of optionele parameters wordt de gebruiker geholpen met het juist instellen van de controletools.

Gedocumenteerde tests

Door het testen en het documenteren daarvan wordt aangetoond dat de controletools werken zoals beschreven in de documentatie van de controletool.

OTAP implementatie

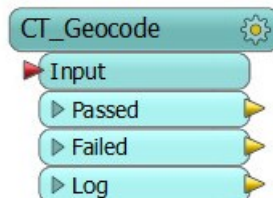
Doordat de controletools geïmplementeerd worden volgens de OTAP-methode zullen de controletools in productie goed werken.

8 Controletools

In dit hoofdstuk staan de functionele beschrijvingen van de zes controletools die zijn ontwikkeld. Alle controletools zijn ontwikkeld als custom transformer.

Zie Bijlage_controlertools.pdf voor:

- Technische beschrijving controletool
- Testrapport controletool
- Workbench controletool
- Test-workbench controletool



8.1 CT_Geocode

8.1.1 Aanleiding controle

Geocode is een belangrijk attribuut dat in veel ProRail datasets voorkomt. Het wordt gebruikt voor navigatie en het opdelen van gebieden.

Geocode

Een geocode is een gebied waarbinnen spoor- en spoor gerelateerde objecten zijn gelegen, aangeduid met een nummer (bijvoorbeeld 009), naam en een eigen kilometrering. Met de combinatie geocode en kilometrering is ieder infra-object met een unieke plaatsaanduiding te beschrijven.

Geosubcode

*Sommige geocodegebieden bevatten meer kilometerbereiken. Het geocodegebied is dan opgesplitst in subcodes. De geocode krijgt er dan een letter bij (bijvoorbeeld 009_a). Elk van deze subcodes heeft zijn eigen bereik.
(ProRail, 2014)*

Er kunnen fouten in geocode ontstaan tijdens het aanmaken van nieuwe workbenches die geocode genereren of bewerkingen toepassen op geocode-velden.

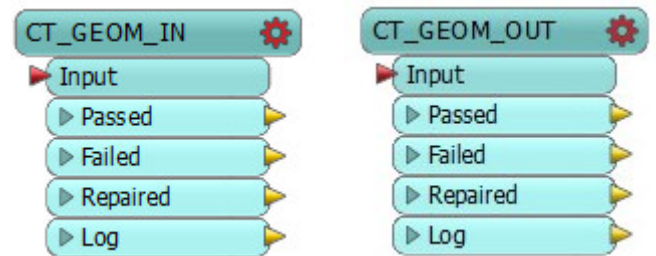
8.1.2 Beschrijving

De controletool controleert attributen:

- Geocode
- Subcode
- Geosubcode

Controleert lengte en karakters met RegEx.

8.2 CT_GEOM



8.2.1 Aanleiding

Er kunnen fouten in geometrie ontstaan tijdens het ontwikkelen van workbenches die geometrie genereren of bewerkingen toepassen op geometrie.

Oracle en ArcGIS geven foutmeldingen als ze foute geometrie moeten tekenen of importeren. In ArcGIS wordt het laden van de dataset zelfs onderbroken als de dataset een of meerdere foute geometrieën bevat. Dit is zeer hinderlijk bij gebruik van een viewer op basis van ArcGIS.

8.2.2 Beschrijving

De controletool voert een geometriecontrole uit. De controletool controleert een dataset op ruimtelijke fouten, bijvoorbeeld zelfsnijdende polygoenen.

De geometriecontrole berust voornamelijk op de GeometryValidator-transformer in FME.

Voor het valideren van geometrie zijn twee controletools ontwikkeld:

CT_GEOM_IN: Geometrie van de originele/ruwe/input dataset valideren.

Eigenschappen controletool:

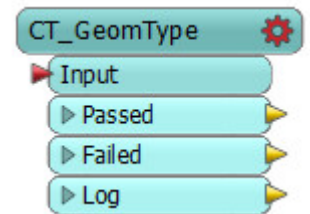
- Valideert geometrie
- Repareert geometrie(optioneel)
- Detecteert dubbele geometrie(optioneel)
- Detecteert Null-geometrie
- Detecteert multi-geometrie

CT_GEOM_OUT: Geometrie van de bewerkte/output dataset valideren.

Eigenschappen controletool:

- Valideert geometrie
- Repareert geometrie(optioneel)
- Detecteert dubbele geometrie(optioneel)
- Detecteert Null-geometrie
- Verwijdert dubbele punten in een geometrie

8.3 CT_GeomType



8.3.1 Aanleiding

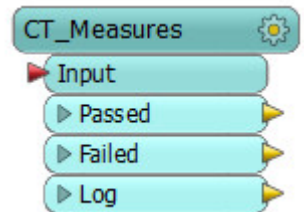
Bij ProRail wordt ArcGIS gebruikt door GIS-gebruikers. De Railmaps-viewer van ProRail is ontwikkeld op basis van een ArcGIS-viewer.

ArcGIS ondersteunt geen meerdere geometrietypen in één tabel. Als er bijvoorbeeld een punt en een lijn in één tabel voorkomen kan de gehele tabel niet geladen worden. Deze controle zorgt ervoor dat slechts één geometrietype in de data komt.

Oracle ondersteunt wel meerdere geometrietypen. De controle is nodig om latere problemen met ArcGIS af te vangen.

8.3.2 Beschrijving

De controletool controleert features op hun geometrietype (punt, lijn of vlak).



8.4 CT_Measures

8.4.1 Aanleiding

Een measure is een kilometrering en is te vergelijken met de kilometerbordjes langs de snelweg. Juiste measures zijn van belang voor ProRail. Omdat ProRail geo-informatie over het spoor beheert en spoor altijd measures bevat. Fouten in measures kunnen ontstaan als measures gegenereerd of geprojecteerd worden.

De kilometrering van het spoor wordt beschreven met measures. Kilometrering wordt gebruikt als locatieaanduiding voor het plegen van onderhoud en het coördineren van calamiteiten. Bijvoorbeeld: Reparatie aan sein A bij Geosubcode 103_b op kilometrering 24.4.

Een groot deel van de geo-informatie van ProRail bevat kilometrering ook als attribuut. Dit wordt ook voor locatieaanduiding gebruikt.

De controletool kan ook gebruikt worden om kilometerlint te controleren.

Een kilometerlint is een lijn die een specifiek gedeelte van het spoorwegnetwerk beschrijft.

Het geeft een route van de ene locatie naar de andere locatie met een eigen kilometrering.

Met de combinatie kilometerlintnaam en kilometrering is ieder infra-object met een plaatsaanduiding te beschrijven.

(ProRail, 2014)

8.4.2 Beschrijving

De controletool is een controle op de measures uit het LRS(Linear Reference System).

Features hebben een X,Y, (Z) en M(Measure).

De controletool controleert measures met de door FME berekende lengte van het feature.

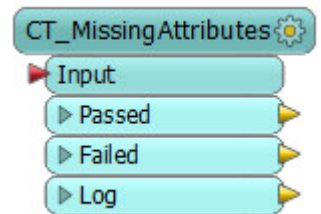
8.5 CT_MissingAttributes

8.5.1 Aanleiding

De geo-informatie van ProRail heeft attributen waarvan het belangrijk is dat ze altijd gevuld zijn met een attribuutwaarde. Bijvoorbeeld ID, projectnummer of status.

8.5.2 Beschrijving

De controletool detecteert null, lege en missende waarden in de opgegeven attributen.



8.6 Hoe zijn de controletools aan de eisen afgestemd?

Alle controletools zijn ontwikkeld conform de eisen die zijn opgesteld in hoofdstuk 7, “Eisen controletools”. Door gebruik te maken van exploratory testing zijn de controletools al grotendeels aan de eisen afgestemd.

Om te kunnen controleren en bewijzen of de controletools aan de eisen voldoen zijn datatests uitgevoerd. Deze datatests zijn gedocumenteerd en te vinden in de bijlage.

Om te testen hoe de controletools in de praktijk bij ProRail ingezet kunnen worden, zijn de controletools geïmplementeerd in 5 workbenches. In het volgende hoofdstuk wordt de implementatie beschreven.

9 Implementatie

In dit hoofdstuk staat hoe de controletools in FME Desktop en FME Server zijn te gebruiken. Daarna wordt uitgelegd waar de controletools in de geo-omgeving bij ProRail worden ingezet. Tenslotte worden de resultaten van de implementatie besproken.

9.1 FME Desktop

Controletools worden beschikbaar gemaakt in de Transformer Gallery onder de categorie Controletools. De controletools worden gedocumenteerd binnen FME.

In FME Desktop kan een controletool gebruikt worden op twee manieren:

Embed betekent dat de definitie van de controletool wordt bepaald in de huidige workbench. De controletool alleen in de huidige workbench worden gebruikt. De controletool is groen gekleurd. Een controletool mag niet embed worden, want dan is het niet meer mogelijk te updaten naar de laatste versie van de controletool.

Linked betekent dat de controletool gedefinieerd is in een apart .fmx bestand. Hierdoor kan de controletool gebruikt worden in meerdere workbenches. De controletool staat dan lokaal bij de huidige gebruiker of op de FME Server. Als de workbench draait wordt gekeken naar die locatie. De controletool is blauw gekleurd.

De controletools moeten altijd als **Linked** gebruikt worden in de workbench. Als er wijzigingen worden doorgevoerd op de controletool worden deze doorgevoerd in alle workbenches die de controletool gebruikt.

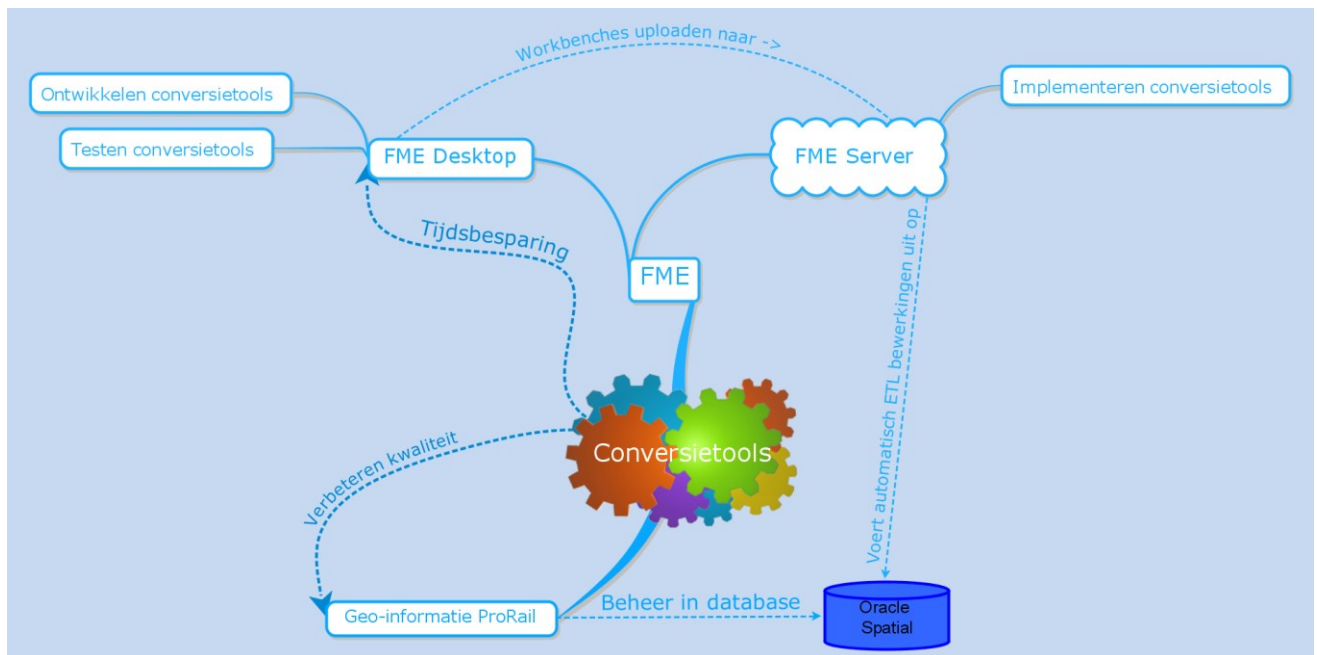
Als een gebruiker een controletool toevoegt in de workbench is deze automatisch een linked controletool. Als de Linked controletool wordt aangepast en opgeslagen geeft FME een waarschuwing dat dit ook andere workbenches beïnvloedt.

Password

Een controletool kan password protected worden gemaakt. Daardoor kan de controletool niet aangepast worden door andere FME gebruikers waardoor de controletool ook in andere workbenches wijzigt. Dit is echter ook een beperking omdat dit de vrijheid van de gebruiker beperkt.

9.2 FME Server

De controletools worden apart naar FME Server geüpload. (Zie bijlage 6)
Workbenches die controletools bevatten kunnen vervolgens op FME Server gezet worden. De workbenches gebruiken dan automatisch de controletool zoals die op FME Server staat gedefinieerd.

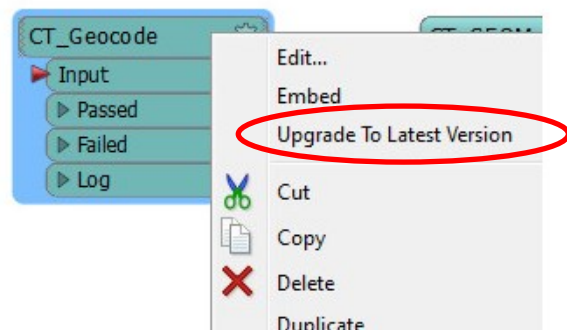


Figuur 7 De relatie tussen de controletools(conversietools), FME en de geo-informatie van ProRail

9.3 Versiebeheer

Als een controletool wordt verbeterd of aangepast, wordt dit doorgevoerd in alle workbenches waar de controletool in wordt gebruikt. Daarom is het soms verstandiger om een nieuwe versie te gebruiken bij het aanpassen van de controletool. FME Desktop ondersteunt versiebeheer. Als een controletool een nieuwere versie heeft, kan snel geüpdatet worden door rechtermuisklik 'Upgrade to latest version'(figuur 8).

FME Server heeft helaas geen versiebeheerfunctionaliteit.



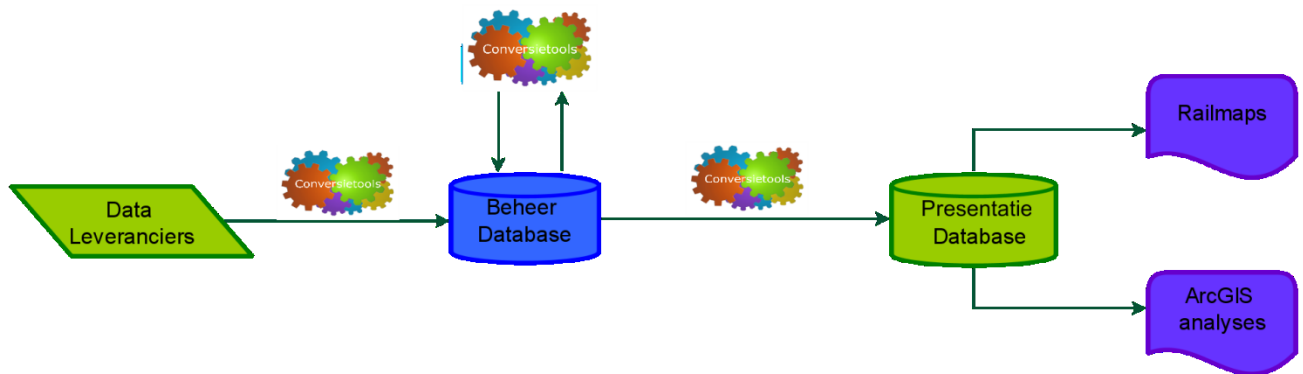
Figuur 8

Source control/Revision control

Omdat veel workbenches gebruik gaan maken van de controletools, moet er een Source Control komen voor de controletools. Source control houdt wijzigingen bij. Als er fouten ontstaan door wijzigingen kan er teruggedaan worden naar een oudere versie. Ook wordt bijgehouden wie wanneer wijzigingen uitvoert. (Sink, 2004)

9.4 Implementatie

Onderstaande figuur illustreert waar de controletools geïmplementeerd worden:



Figuur 9 Implementatie controletools

De controletools zijn geïmplementeerd in de ontwikkel- en testomgeving van ProRail. De controletools zijn daar beschikbaar in FME Desktop en FME Server.

De controletools kunnen momenteel nog niet in de acceptatie- en productieomgeving gebruikt worden. De controletools zijn ontwikkeld in FME 2015. Momenteel gebruikt ProRail (nog) FME 2012 in de acceptatie en productieomgeving. FME 2012 ondersteunt geen FME 2015 workbenches.

Hoe kunnen de controletools geïmplementeerd worden?

De implementatie moet nog voltooid worden door ProRail. Een goed moment om dit te doen is bij het omzetten van workbenches van FME 2012 naar FME 2015.

9.5 Resultaten implementatie

Het is belangrijk dat de controletools goed geïmplementeerd kunnen worden. Om dit te testen zijn de controletools geïmplementeerd in 5 workbenches. Deze workbenches staan op FME Server 2015 in de testomgeving. De controletools moeten nog in alle overige workbenches geïmplementeerd worden.

De implementatie van de 5 workbenches is gedetailleerd gedocumenteerd. Deze documentatie is te vinden in de bijlage.

Resultaten implementatie controletools (globaal):

- Fouten in workbench gedetecteerd.
- Fouten in data gedetecteerd.
- Controletools zijn verbeterd naar aanleiding van problemen/onvolkomenheden tijdens de implementatie.

Best practices tijdens implementatie:

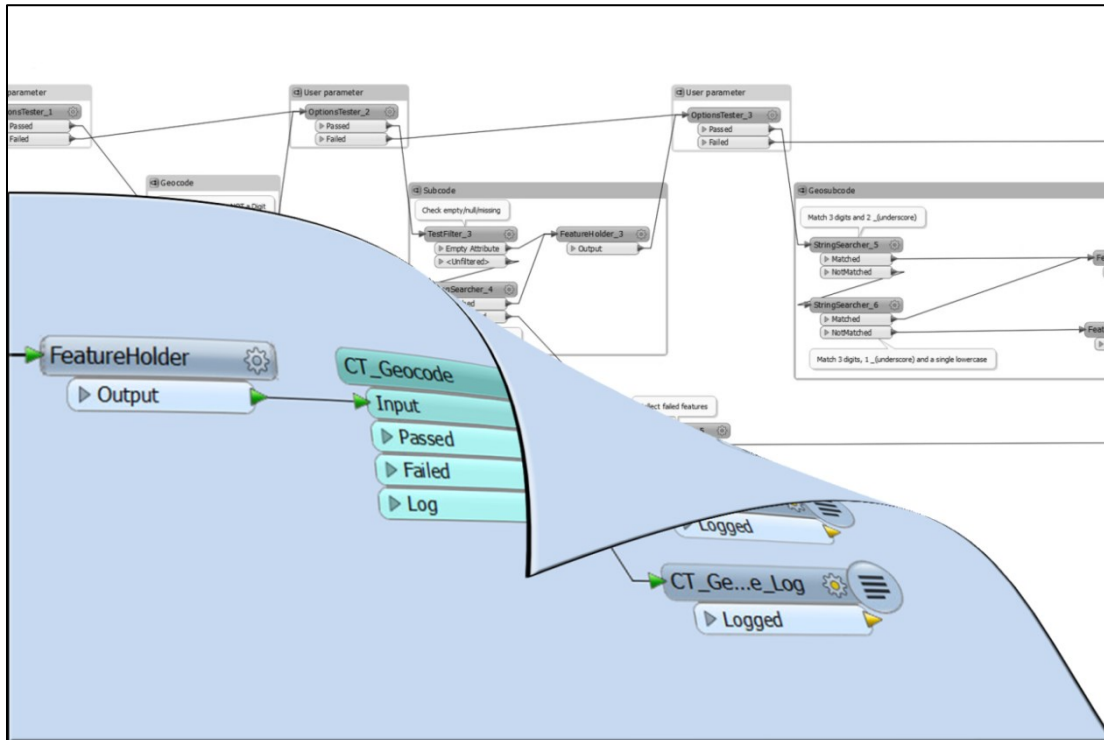
- De output van de workbench moet ongewijzigd blijven.
- Het is belangrijk erop te letten dat de controletools juiste input krijgen.
- Er moet goed overwogen worden waar en wanneer is welke controletool nodig is: Is bij het inlezen van data een controle nodig? Van sommige data kan ervan uit gegaan worden dat het foutloos is.
- De opbouw en werking van een workbench bepaalt welke controletools geïmplementeerd moeten worden. Als er bijvoorbeeld transformers gebruikt zijn die de geometrie aanpassen, moet CT_GEOM_OUT gebruikt worden.
- De log van de controletools helpt bij het vinden van de oorzaak van een fout.

Tijdens de implementatie kan het voorkomen dat een controletool niet geschikt is door de specifieke eisen in een workbench. Dit moet opgelost worden met behulp van de volgende twee vuistregels:

- ❖ Als een controletool niet geschikt is:
 - Features die onjuist worden afgewezen door de controletool in de workbench uitfilteren en om de controletool heen leiden.
- ❖ Als het vaak voorkomt dat er dezelfde aanpassingen gedaan moeten worden voor goede werking van de controletool:
 - Extra parameter en/of extra functionaliteit in de Linked controletool verwerken.

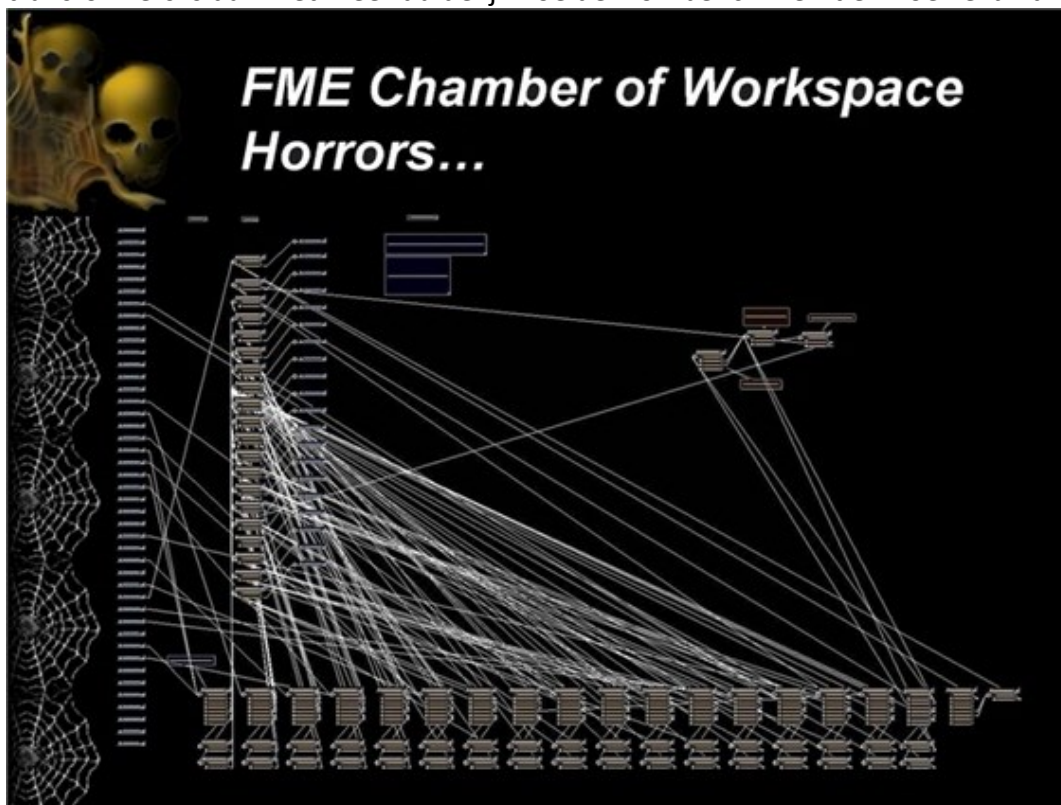
Overzicht in de workbench

Uit de implementatie is gebleken dat de controletools overzicht brengen in de workbench doordat ze complexe functionaliteit verbergen:



Figuur 10 Bij het gebruik van een controletool wordt achter de schermen een hoop transformers toegevoegd.

Door gebruik van de controletools worden workspace-horrors voorkomen. Bij veel transformers is dan niet meer duidelijk hoe de workbench werkt en hoe verbindingen lopen:



Figuur 11 Bron: Safe Software, 2012

9.6 Datakwaliteit en tijdsbesparing

Datakwaliteit en tijdsbesparing zijn de twee grootste voordelen van het gebruik van controletools.

Omdat de controletools op een eenduidige wijze werken en goed gedocumenteerd zijn kunnen ze gemakkelijk gebruikt worden. De controletools zijn betrouwbaar want ze zijn getest.

De functionaliteit van de controletools hoeft niet in iedere workbench opnieuw gebouwd te worden. Dezelfde functionaliteit wordt op meerdere plekken gebruikt, maar hoeft slechts één keer centraal aangepast te worden en dat wordt vervolgens automatisch doorgevoerd in alle workbenches die gebruik maken van de controletool. Hierdoor wordt herhaling in workbenches voorkomen. Dit zorgt voor tijdswinst bij het onderhouden van FME workbenches.

De controletools kunnen gebruikt worden aan het begin van een workbench om de aangeleverde data te controleren op fouten. Hierdoor kan terugkoppeling naar brondatabasebeheerders plaatsvinden waardoor fouten in de brondata hersteld kunnen worden. Door deze controle is ook direct duidelijk of de oorzaak van de fout in de workbench zit of in de aangeleverde data. Door de goede logging van de controletools kan snel bepaald worden wat de fout in de data is en wat de oorzaak van de fout is. De kans om fouten in workbenches te ontwikkelen wordt kleiner als er controletools worden gebruikt vlak voor het wegschrijven van de data. Dit helpt ook het voorkomen van latere problemen door verkeerde data.

“Voorkomen is beter dan genezen” is een toepasselijk gezegde op de controletools. Het gaat tijd kosten om de controletools te implementeren. Na de implementatie zullen er minder fouten in de data zitten wat ProRail als organisatie in zijn geheel profijt zal opleveren.

Gebruikers die tijdswinst ondervinden:

- FME-Consultants: Zij bouwen actief aan huidige en nieuwe workbenches en kunnen daarbij de controletools gebruiken om betere workbenches te ontwikkelen en fouten in de workbench snel en gemakkelijk te detecteren.
- Applicatiebeheerder: De applicatiebeheerder is verantwoordelijk voor de techniek. Hij moet er bijvoorbeeld voor zorgen dat FME Server blijft draaien. Door de controletools zal de ArcGIS viewer beter werken en hoeft de applicatiebeheerder minder problemen op te lossen.
- Functioneel beheerder: Hij is de schakel tussen de applicatie en gebruiker. Hij is verantwoordelijk dat de gebruiker de applicatie kan blijven gebruiken.

Gebruikers die direct belang hebben bij datakwaliteit die ontstaat door controletools:

- Ongeveer 150 GIS gebruikers (ProRail-medewerkers).
- Systemen die data afnemen en mapservices gebruiken: Railmaps, BBMS, Sigma, SpoorWeb. Van deze systemen maken ook externen gebruik zoals bijvoorbeeld aannemers.

Hoe kunnen de controletools zorgen voor kwaliteitswinst en tijdswinst?

- De controletools controleren de geo-informatie van ProRail op fouten. Dit zorgt voor kwaliteitswinst.
- De controletools voorkomen herhaling in workbenches. Dit zorgt voor tijdswinst.

10 Conclusie

De onderzoeksvraag was: Hoe worden FME en FME Server momenteel gebruikt en kan dit effectiever door het maken van controletools?

FME wordt gebruikt om conversies (van geo-informatie) te bouwen.

FME Server wordt gebruikt om de conversies op vaste tijdstippen uit te voeren.

De volgende controletools zijn ontwikkeld:

CT_Geocode

De controletool controleert attributen Geocode, Subcode en Geosubcode op juistheid.

CT_GEOM_IN

Valideert de geometrie van originele/ruwe/input datasets.

CT_GEOM_OUT

Valideert de geometrie van bewerkte/output datasets.

CT_GeomType

Controleert features op hun geometrietype.

CT_Measures

Controleert measures met de door FME berekende lengte van het feature.

CT_MissingAttributes

Detecteert null, lege en missende waarden in de opgegeven attributen.

De controletools voldoen aan de eisen van ProRail. De controletools hebben een eenduidige naamgeving, documentatie en werking. Hierdoor zijn ze gemakkelijk te gebruiken in workbenches. Ze brengen overzicht in de workbench. Door de logging van de controletools kunnen fouten en hun oorzaak snel gedetecteerd worden. Dit maakt het gebruik van FME effectiever.

De controletools zijn getest en deels geïmplementeerd in de ontwikkelomgeving van ProRail. De implementatie bij ProRail is nog niet voltooid. FME gebruikers kunnen de controletools nu gebruiken bij het ontwikkelen van nieuwe workbenches in FME 2015 in de ontwikkelomgeving. Bij het omzetten van de workbenches van FME 2012 naar FME 2015 kunnen de controletools geïmplementeerd worden in bestaande workbenches.

11 Aanbevelingen

FME 2015

De controletools kunnen momenteel nog niet in de acceptatie- en productieomgeving gebruikt worden. De controletools zijn ontwikkeld in FME 2015. Momenteel gebruikt ProRail (nog) FME 2012 in de acceptatie en productieomgeving. FME 2012 ondersteunt geen FME 2015 workbenches. ProRail zal moeten upgraden naar FME Desktop 2015 en FME Server 2015 in de acceptatie en productieomgeving.



Controletools implementeren

De implementatie moet nog voltooid worden door ProRail. Een goed moment om dit te doen is bij het omzetten van workbenches van FME 2012 naar FME 2015.

Source control/Revision control

ProRail moet een source control opzetten om wijzigingen in de controletools bij te houden.

Controletools Log

De log-informatie van de controletools kan worden gebruikt in een soort dashboard van processen. Met dit dashboard kan overzichtelijk inzichtelijk worden gemaakt waar fouten gedetecteerd worden en wat de oorzaak van de fout is.

FME ETL snelheid

Een nadeel van FME is dat het alle features moet inladen in het werkgeheugen en in FFS (FeatureStoreFiles) en vervolgens weer wegschrijft. Dat is veel trager dan conversies uitvoeren op databaseniveau.

Het kan tijd schelen om een where clause in te stellen voor een Reader. Hierdoor wordt bij het inlezen van een database al een query op de database uitgevoerd. Hierdoor leest FME niet de gehele database in, maar slechts de benodigde data.

FME niet altijd de beste oplossing

Soms kan een enorme workbench vervangen worden door een klein stukje Oracle Spatial SQL-script van 10 regels. Bijvoorbeeld bij het projecteren van measures.



In zo'n geval is een stukje SQL-script beter te onderhouden dan een grote onoverzichtelijke workbench. Conversies met SQL-script op databaseniveau hebben een betere ETL-snelheid. ProRail houdt momenteel nogal krampachtig vast aan gebruik van FME, terwijl FME in sommige gevallen niet de beste oplossing is.

12 Begrippen en afkorting

BBK	BasisBeheerKaart. Dit is een openbare kaart met alle spoorobjecten.
Controletool	Een controletool is een custom transformer met functionaliteit die in meerdere workbenches gebruikt kan worden. Een controletool is generiek in te zetten in bestaande en nieuwe workbenches. Een controletool is getest en gedocumenteerd en voldoet aan de eisen van ProRail.
Creator (transformer)	Creëert features op basis van de door de gebruiker gegeven parameters.(Safe Software, 2014)
Custom Transformer	Een custom transformer is een reeks aan standaard transformers, samengevat in één enkele transformer. (Safe Software, 2014)
Database	Een verzameling van gegevens die gestructureerd zijn opgeslagen.
Datatests	Een testmethode waarbij testdata wordt gegenereerd.
ETL	Extract, Transform en Load.
Exploratory testing	Het tegelijkertijd leren, tests ontwikkelen en tests uitvoeren. (Bach, 2003)
Feature	Een feature is een object. Bijvoorbeeld een record in een database inclusief attributen en geometrie.
FME	Feature Manipulation Engine. De naam van een aantal producten van Safe software Inc.
FME Data Inspector	Onderdeel van FME Desktop. Wordt gebruikt voor het visualiseren en inspecteren van geo-informatie.
FME Desktop	FME Desktop is een product van Safe software Inc. FME Desktop is een ETL applicatie en wordt gebruikt voor geautomatiseerde datatranslaties, dataconversies en datavalidatie van geo-informatie. De eerste versie kwam uit in 1996. De versie gebruikt tijdens deze scriptie is FME 2015.
FME Server	FME Server is een ruimtelijke ETL applicatie voor translatie en transformatie van geo-informatie over een netwerk. Hierdoor kunnen organisaties hun data bekijken, herstructureren en distribueren. (Safe Software, 2014) De versie gebruikt tijdens deze scriptie is FME Server 2015.
Geocode	Een geocode is een gebied waarbinnen spoor- en spoor gerelateerde objecten zijn gelegen, aangeduid met een nummer (bijvoorbeeld 009), naam en een eigen kilometrering. Met de combinatie geocode en kilometrering is ieder infra-object met een

	unieke plaatsaanduiding te beschrijven. (ProRail, 2014)
GIS	Geografisch Informatie Systeem
Inspector	Een transformer die de input visualiseert in de FME Data Inspector.
NaN	Not a Number
OGC	Open Geospatial Consortium
OOP	Object Oriented Programming
Oracle Spatial	Oracle Spatial is een database die binnen veel organisaties wordt gebruikt om geo-informatie in op te slaan en te beheren. Met Oracle Spatial kunnen ook datatransformaties op databaseniveau uitgevoerd worden.
OTAP	OTAP is een methode die veel wordt gebruikt in de ICT en staat voor: Ontwikkeling, Test, Acceptatie en Productie.
Parameter	Alle transformers hebben parameters. In parameters kun je de eigenschappen van de transformers definiëren.
Python	Python is een populaire open-source programmeertaal.
Railmaps	GIS viewer voor alle ProRail-medewerkers
Reader	De reader leest de brondata, het formaat en de locatie.
RegEx	Regular expression is een tekst patroon dat gebruikt kan worden in veel modern applicaties en programmeertalen. Dit kan gebruikt worden om input te toetsen of om tekstpatronen te vinden. (Goyvaerts en Levithan, 2012).
Repositorie	Wordt gebruikt om workbenches in FME Server te categoriseren en op te slaan.
Source control/ Revision control	Source control houdt wijzigingen bij. Als er fouten ontstaan door wijzigingen kan er teruggedaan worden naar een oudere versie. Ook wordt bijgehouden wie wanneer wijzigingen uitvoert.
TCL	Tool Command Language. Dit is een programmeertaal die uitsluitend in FME is te gebruiken.
Transformer	Een transformer is een object dat een bewerking uitvoert op features. Een transformer wordt meestal ingezet tussen de reader en de writer in. FME bevat meer dan 500 verschillende transformers. Transformers zijn gecategoriseerd in de transformer Gallery in FME Desktop.
Unit testing	Een testmethode waarbij algoritmen afzonderlijk getest worden.
Workbench/workspace	Dit is een grafische weergave van datatransformatie. Een

	workbench wordt opgeslagen als .fmw bestand en kan geopend, gewijzigd en uitgevoerd worden met FME Desktop.
Writer	De writer specificeert het formaat en locatie van de output-data.

13 Bronnen afstudeeronderzoek

Referenties

Bach, J. (2003). Exploratory Testing Explained.

Conterra (2012). Webinar Python Scripting in FME.

Goyvaerts, J. en Levithan, S. (2012). Regular Expressions Cookbook, 2nd Edition.
ISBN-13: 978-1449319434

Kolawa, A (2009). Unit Testing Best Practices.

ProRail (2014). Productbeschrijving BBK Geleidingsysteem.

ProRail (2014). Productbeschrijving Referentiesysteem.

ProRail (2014). Systeemconcept Geoprocessing.

Shaw, Z. (2013). Learn Python the hard way. *ISBN-13: 978-0321884916*

Sink, E (2004). Source Control HOWTO.

van Beveren, E (2012). De OTAP werkwijze. *ISBN-9789081701518*

Willems, K. (2011). OGH-visie.

Internetreferenties

Python documentatie (februari – juni 2015):

URL: www.docs.python.org/2.7

FME Transformer documentatie (februari – juni 2015):

URL: http://docs.safe.com/fme/html/FME_Transformers/FME_Transformers.htm

Safe Software (februari – juni 2015):

URL: www.safe.com

FMEpedia (februari – juni 2015):

URL: <http://fmepedia.safe.com>

Oracle Spatial User's guide and reference (februari – juni 2015):

URL: http://docs.oracle.com/cd/B19306_01/appdev.102/b14255/toc.htm

URL: http://docs.oracle.com/cd/B19306_01/appdev.102/b14255/sdo_util.htm#BJEICGGA

URL: http://docs.oracle.com/cd/B19306_01/appdev.102/b14255/sdo_objgeom.htm

OGC Open Geospatial Consortium (februari – juni 2015):

URL: <http://www.opengeospatial.org/>

RegEx simulator (februari – juni 2015):

URL: www.regexr.com

IGO afstuderen (februari 2015):

URL: <https://onderwijsteams.sharepoint.hu.nl/fnt/IGO-afstuderen/default.aspx>