

Review PDOK-formaat voor mutatieleveringen

Henri Korver (KING)

Inleiding

Deze review richt zich voornamelijk op het analyseren van de verschillen tussen het nieuwe PDOK- en het bestaande StUF-formaat. De conclusie is dat het nieuwe formaat weinig voordelen biedt. Het lijkt alsof de interne representatie van de database voor het bijhouden van historische voorkomens (versies van objecten) doorschemert in de leveringsberichten van de PDOK. Dat is misschien handig voor de softwareontwikkelaars van de PDOK maar niet voor de afnemers zoals gemeenten.

Daarnaast worden gemeenten geconfronteerd met twee verschillende formaten:

- StUF voor het vullen van de basisregistraties BAG en BGT
- PDOK-levering voor het afnemen van gegevens uit de BAG, BGT en BRK.

Het spreekt voor zich dat dit voor gemeenten geen wenselijke situatie is.

Analyse

De belangrijkste verschillen tussen het mutatieformaat PDOK en StUF zijn:

- Het gebruik van een extra technische identificatie voor historische voorkomens voor een object naast een technische identificatie van het object zelf. De ontvanger van het mutatieformaat zal bij de verwerking deze extra identificatie moeten opslaan in zijn database om mutaties en correcties te kunnen verwerken.
- Het gebruik van een extra voorkomen om eindRegistratie te zetten

Een en ander is aan het werk te zien in het bestand `mutatie_d_bgt_new-change-fix.xml` in de tweede mutatiegroep (regel 36 t/m 114), waarbij een wijziging wordt doorgevoerd. Er wordt begonnen met een was/wordt om eindRegistratie te zetten. Het voorkomen met eindRegistratie gezet krijgt een nieuw id '385e9dbd-1a2b-4f32-bae2-1e5e15c52453'. Daarna volgt een wordt met de nieuwe situatie. De LV bevat drie records:

1. Het record zoals het object gecreëerd is (id=08276e16-6a0b-4647-99af-d643c735bb22)
2. Het eerste record, maar nu met eindRegistratie ook een waarde (id=385e9dbd-1a2b-4f32-bae2-1e5e15c52453)
3. Het record met de nieuwe gegevens (id=94c49817-633e-4e82-9abd-32f1b2f4de2e)

Het tweede record is afleidbaar uit het eerste en het derde record door aan het eerste record eindRegistratie toe te voegen met als waarde tijdstipRegistratie in het derde record.

In StUF wordt volstaan met een was/wordt, waarbij het object wordt geïdentificeerd met het attribute `StUF:sleutelVerzendend`, vergelijkbaar met attribute `objectId` binnen het `<mutatie>` element. Deze `sleutelVerzendend` en `objectId` hebben dan ook dezelfde waarde. Je kunt je overigens afvragen of je in geval van de BGT niet kunt volstaan met de functionele identificatie in het element `imgeo:identificatie` of is het mogelijk dat de inhoud ervan in de levensloop van een object wijzigt?

Het element objecttype binnen het <mutatie> element is vervangen door het attribute imgeo:entiteittype.

```
<cl:was imgeo:entiteittype="pand" StUF:sleutelVerzendend="G0855.44cae3deb10200e6e0530a01fa86e02a">
  <creationDate xmlns="http://www.opengis.net/citygml/2.0">2016-12-29</creationDate>
  <imgeo:LV-publicatiedatum>2017-02-22T11:17:02.000</imgeo:LV-publicatiedatum>
  <imgeo:relatieveHoogteligging>0</imgeo:relatieveHoogteligging>
  <imgeo:inOnderzoek>>false</imgeo:inOnderzoek>
  <imgeo:tijdstipRegistratie>2017-01-26T03:32:09.000</imgeo:tijdstipRegistratie>
  <imgeo:identificatie>
    <imgeo:NEN3610ID>
      <imgeo:namespace>NL.IMGeo</imgeo:namespace>
      <imgeo:lokaalID>G0855.44cae3deb10200e6e0530a01fa86e02a</imgeo:lokaalID>
    </imgeo:NEN3610ID>
  </imgeo:identificatie>
  <imgeo:bronhouder>G0855</imgeo:bronhouder>
  <imgeo:bgt-status codeSpace="http://www.geostandaarden.nl/imgeo/def/2.1#Status">bestaand</imgeo:bgt-status>
  <imgeo:plus-status codeSpace="http://www.geostandaarden.nl/imgeo/def/2.1#VoidReasonValue">geenWaarde</imgeo:plus-status>
  <imgeo:identificatieBAGPND>0855100000540914</imgeo:identificatieBAGPND>
  <imgeo:geometrie2dGrondvlak>...</imgeo:geometrie2dGrondvlak>
  <imgeo:nummeraanduidingreeks>...</imgeo:nummeraanduidingreeks>
</cl:was>
<cl:wordt imgeo:entiteittype="pand" StUF:sleutelVerzendend="G0855.44cae3deb10200e6e0530a01fa86e02a">
  <creationDate xmlns="http://www.opengis.net/citygml/2.0">2016-12-29</creationDate>
  <imgeo:LV-publicatiedatum>2017-05-18T12:53:56.000</imgeo:LV-publicatiedatum>
  <imgeo:relatieveHoogteligging>0</imgeo:relatieveHoogteligging>
  <imgeo:inOnderzoek>>false</imgeo:inOnderzoek>
  <imgeo:tijdstipRegistratie>2017-05-18T10:35:14.000</imgeo:tijdstipRegistratie>
  <imgeo:identificatie>
    <imgeo:NEN3610ID>
      <imgeo:namespace>NL.IMGeo</imgeo:namespace>
      <imgeo:lokaalID>G0855.44cae3deb10200e6e0530a01fa86e02a</imgeo:lokaalID>
    </imgeo:NEN3610ID>
  </imgeo:identificatie>
  <imgeo:bronhouder>G0855</imgeo:bronhouder>
  <imgeo:bgt-status codeSpace="http://www.geostandaarden.nl/imgeo/def/2.1#Status">bestaand</imgeo:bgt-status>
  <imgeo:plus-status codeSpace="http://www.geostandaarden.nl/imgeo/def/2.1#VoidReasonValue">geenWaarde</imgeo:plus-status>
  <imgeo:identificatieBAGPND>0855100000540914</imgeo:identificatieBAGPND>
  <imgeo:geometrie2dGrondvlak>...</imgeo:geometrie2dGrondvlak>
  <imgeo:nummeraanduidingreeks>...</imgeo:nummeraanduidingreeks>
</cl:wordt|
```

Figuur 1 In StUF zijn slechts twee voorkomens nodig om de mutatie te specificeren.

StUF heeft één voorkomen minder nodig en bevat in minder dan 2/3 van het PDOK mutatieformaat alle voor de verwerking benodigde informatie. In het voorbeeld hierboven zijn alle elementen van het object opgenomen in de was/wordt. Dit is functioneel niet nodig. Het is voldoende om naast de identificerende gegevens alleen de wijzigende elementen op te nemen. In het bovenstaande voorbeeld wordt de was/wordt dan:

```
<cl:was imgeo:entiteittype="pand" StUF:sleutelVerzendend="G0855.44cae3deb10200e6e0530a01fa86e02a">
  <imgeo:LV-publicatiedatum>2017-02-22T11:17:02.000</imgeo:LV-publicatiedatum>
  <imgeo:tijdstipRegistratie>2017-01-26T03:32:09.000</imgeo:tijdstipRegistratie>
  <imgeo:identificatie>
    <imgeo:NEN3610ID>
      <imgeo:namespace>NL.IMGeo</imgeo:namespace>
      <imgeo:lokaalID>G0855.44cae3deb10200e6e0530a01fa86e02a</imgeo:lokaalID>
    </imgeo:NEN3610ID>
  </imgeo:identificatie>
</cl:was>
<cl:wordt imgeo:entiteittype="pand" StUF:sleutelVerzendend="G0855.44cae3deb10200e6e0530a01fa86e02a">
  <imgeo:LV-publicatiedatum>2017-05-18T12:53:56.000</imgeo:LV-publicatiedatum>
  <imgeo:tijdstipRegistratie>2017-05-18T10:35:14.000</imgeo:tijdstipRegistratie>
  <imgeo:identificatie>
    <imgeo:NEN3610ID>
      <imgeo:namespace>NL.IMGeo</imgeo:namespace>
      <imgeo:lokaalID>G0855.44cae3deb10200e6e0530a01fa86e02a</imgeo:lokaalID>
    </imgeo:NEN3610ID>
  </imgeo:identificatie>
</cl:wordt>
```

Figuur 2 In StUF zijn de berichten compacter omdat naast de identificerende gegevens alleen de gewijzigde elementen worden opgenomen.

Als we ook de elementen met gelijke waarden in was en wordt weglaten, dan gaan we terug van ruwweg 74 regels voor het PDOK mutatieformaat naar 20 regels voor StUF.

StUF laat daarnaast afnemers vrij in de wijze waarop historie in hun database wordt gerepresenteerd. Om voordeel te hebben van de voorgestelde PDOK mutatie-opbouw moeten afnemers hun database inrichten met technische sleutels voor historische voorkomens gelijk aan de door PDOK geleverde id's. Ik kan me voorstellen dat dit op bezwaren stuit bij partijen die voor een andere systematiek voor het omgaan met historie gekozen hebben. Denk in het bijzonder aan gemeenten waar WOZ-systemen dankzij StUF vrij zijn geweest in het kiezen van een model voor het opslaan van historie.

De werkwijze van StUF heeft één nadeel en dat zie je aan het derde voorbeeld van de correctie van versie 1 in het bovengenoemde bestand (regels 118 t/m 168). Het voorgestelde mutatieformaat kan een correctie in de historie doorgeven als een was/wordt voor dat voorkomen. In StUF kan dat alleen worden doorgegeven met behulp van een synchronisatiebericht, omdat StUF alleen het muteren van actuele gegevens in kennisgeving definieert. Correcties in historische gegevens kunnen alleen worden doorgevoerd met behulp van synchronisatieberichten. In de derde open vraag wordt al geopperd dat een mechanisme à la de StUF synchronisatieberichten wenselijk kan zijn. Het grote voordeel van het mechanisme van StUF voor afnemers is dat zij alleen de verwerking van wijzigingen in actuele gegevens hoeven te implementeren en geen mechanisme hoeven te implementeren voor het corrigeren van historische gegevens. Dit geeft de afnemers meer keuzevrijheid voor het implementeren van historie.

De opmerking in de regels 116 en 117 dat bij een correctie van gegevens de tijdstipRegistratie niet verandert, is merkwaardig. Een correctie wordt toch op een later tijdstip in de database doorgevoerd dan het registreren van de achteraf foutief gebleken gegevens of wordt de betekenis van tijdstipRegistratie niet goed begrepen?

Een ander voordeel van de keuze voor StUF richting afnemers is dat aangesloten kan worden bij de berichten die door de bronhouders aan de LV BGT worden geleverd. Deze kunnen één-op-één doorgezet worden met uitzondering van de stuurgegevens: zender, ontvanger, referentienummer en tijdstipBericht. Er kan natuurlijk ook voor worden gekozen om berichten periodiek uit te leveren door voor een bepaalde periode verzamelde berichten op te nemen in een StUF-berichtenset.