

SYSTEMATISCH WERKEN AAN BETERE GEODATA



Stockphoto

Veldsteekproef datakwaliteit bij de waterkering in Zutphen

Vanouds zijn metingen en locatiebepalingen van cruciaal belang voor de waterschappen. Nu er steeds meer gegevens bijkomen en datagedreven werken de ambitie is, wordt het verwerken en beschikbaar maken van 'waterdata' steeds belangrijker. En al die data moeten kloppen. Waterschap Rijn en IJssel is een voorloper in de zorg voor goede en betrouwbare data.

Kwaliteit van data staat centraal in de 'I-Strategie' van Waterschap Rijn en IJssel (WRIJ) uit 2018. Goede data zijn belangrijk voor interne en externe gebruikers, maar bijvoorbeeld ook vanwege de eisen van de nieuwe Omgevingswet. Bovendien zijn data van slechte kwaliteit een risico. Bij calamiteiten kunnen incomplete gegevens zelfs leiden tot verkeerde beslissingen. In principe moet je op data kunnen vertrouwen. Daarbij speelt altijd de vraag: hoe goed is goed genoeg? Welke eisen stel je aan de data voor welke toepassing? En hoe meet je de kwaliteit van de grote hoeveelheid geodata?

Geo-team

Sinds 2016 heeft Waterschap Rijn en IJssel een 'Geo-team' van 11 mensen. Dit team, sinds 2019 inclusief een data-kwaliteitsanalist, werkt aan kwaliteit in de hele 'dataketen', van inwinning, opslag en mutaties tot ontsluiting en gebruik.

Voor het Geo-team is de relatie met business units als Waterkeringen & Vaarwegbeheer en Watersystemen

cruciaal. Daar ligt het 'eigendom' van de data. In elk van deze units is één persoon verantwoordelijk voor het gegevensbeheer. Deze functioneel gegevensbeheerders bepalen met de medewerkers welke data thuishoren in de kernregistratie en aan welke eisen ze moeten voldoen. Zij bepalen ook de prioriteiten: voor welke data is opname in de kernregistratie urgent?

Een onmisbare leidraad is het WRIJ-gegevenswoordenboek. Naast definities, afspraken over gegevenseigendom en invulafspraken bevat het woordenboek ook eisen aan inwinning, aanlevering en ontsluiting van data. Deze basisafspraken vormen het uitgangspunt voor het aanleveren van data, bijvoorbeeld bij oplevering van projecten of het verlenen van vergunningen.

Grip op datakwaliteit

Het waterschap werkt vanuit het Geo-team structureel aan het monitoren, verbeteren en borgen van de kwaliteit van de geodata. Door continu een vinger aan de pols te houden, krijgt de organisatie als geheel grip op datakwaliteit en



Plan-do-act-cyclus datakwaliteit

kunnen er gerichte, bewuste en transparante keuzes worden gemaakt.

Bij het bepalen van de kwaliteitseisen staan de eisen en wensen van gebruikers centraal, om te beginnen van de interne gebruikers. Inmiddels zijn de kwaliteitseisen voor de belangrijkste datasets van de programma's Waterkeringen en Watersystemen bepaald (ruwweg driekwart van hun data). Op basis daarvan lopen sinds vorig jaar de eerste geautomatiseerde kwaliteitscontroles. Binnenkort zijn de geodata van Zuiveringsbeheer en en Riolerings aan de beurt.

De belangrijkste manier om datakwaliteit te monitoren, bestaat uit maandelijkse geautomatiseerde controles met de Datareviewer van ArcGIS. Voor elke dataset is op een monitoringsdashboard te zien hoe de kwaliteitsscore zich ontwikkelt, welke fouten prioriteit hebben en of een bepaalde controle ineens veel meer of minder meldingen oplevert. Gegevensbeheerders kunnen met de Datareviewer ook snel kleine foutjes herstellen na invoering van nieuwe gegevens.

De meeste metingen aan waterlopen en kunstwerken komen van externe meetbureaus. Nu worden deze data nog handmatig gecontroleerd. We werken aan een automatische controlestap om ontbrekende gegevens en onlogische waarden te onderscheppen voordat de data in de centrale database terecht komen.

Eenmalige controles, steekproeven, terugmeldingen

Soms zijn andere analyses nodig. Zo wordt vanaf 2019 voor het hele werkgebied van het waterschap systematisch gewerkt aan duikers die er wel liggen, maar niet in het systeem zitten en omgekeerd. Soms is daarvoor een check op basis van luchtfoto's voldoende, soms is een veldbezoek nodig.

Voor waterkeringen is het waterschap gestart met steekproeven. Met behulp van een app worden jaarlijks op drie locaties de data gecontroleerd op aan- en afwezigheid van

objecten, de juistheid van beschrijvende gegevens en een correcte ligging of afbakening.

Tot slot is terugmelding door gebruikers een onmisbare bron van informatie. Door het interne Geo Informatie Portaal en via apps, viewers en dashboards komen steeds meer collega's in aanraking met data en dus ook met tekortkomingen daarin. Gebruikers kunnen via het terugmeldingssysteem zaken als verkeerde stuwtypes of ontbrekende bruggen melden, waarna de gegevenseigenaar verantwoordelijk is voor de afhandeling.

Verbeteracties

Bij het verbeteren van de datakwaliteit is het soms moeilijk om te kiezen welke fouten je het eerste aanpakt. WRIJ werkt in zorgvuldige en haalbare stappen, en kiest daarbij voor een risicogestuurde planning van verbeteracties. De kwaliteitsmonitoring stelt het waterschap nu in staat hier bewuste keuzes in te maken.

Datakwaliteit is een veelomvattend vraagstuk, waarbij een organisatie graag goede afwegingen wil maken tussen kosten en baten. Dat begint met inzicht en het gesprek aangaan met gebruikers. Door datakwaliteitsmanagement als continu en deels geautomatiseerd proces in te richten, wordt goede datakwaliteit routine, groeit het vertrouwen en worden de risico's en kosten van slechte data kleiner.

Bram de Vlieger, Djurra Scharff, Emiel Huizinga, Kristan Margry, Judith Janssen (*Waterschap Rijn en IJssel*)

Een uitgebreide versie van dit artikel is te vinden op H₂O-Online. Het is te lezen door gebruik te maken van de QR-code of te kijken op www.h2owaternetwerk.nl (onder H₂O-vakartikelen).



SAMENVATTING

Waterschap Rijn en IJssel werkt sinds 2019 steeds systematischer aan datakwaliteit. De belangrijkste geodata-sets van de programma's Waterkeringen en Watersystemen zijn inmiddels opgenomen in geautomatiseerde maandelijkse kwaliteitscontroles. Door de samenwerking tussen het geo-team en functioneel gegevensbeheerders wordt de organisatie als geheel steeds beter in het verbeteren van datakwaliteit aan de voorkant (fouten voorkomen) en achterkant (fouten oplossen).