

Vraag 1

Voor archieven wordt het allemaal een stuk simpeler, vooral qua migraties, als deze standaard verplicht zou zijn. Nu is een groot deel van ons werk het maken en/of begeleiden van conversies van proprietary formaten naar duurzamere en beter presenteerbare alternatieven.

Kortom, deze verplichting gaat geld schelen.

Vraag 2

Op zich eens dat dit een goede reikwijdte voor MDTO is.

informatie vindbaar, beschikbaar, leesbaar, betrouwbaar, interpreteerbaar en toekomstbestendig [p. 1]

Als dit je identiteit is, of, in ieder geval, hoe je je soms aan de buitenwereld wilt presenteren, moet je je denk ik afvragen of ieder woord hier zijn plek verdient.

Wat is bijvoorbeeld het verschil tussen leesbaar en interpreteerbaar? Welke specifieke gegevens of maatregelen maken MDTO leesbaar, en welke zorgen juist voor interpreteerbaarheid? Kun je een voorbeeld geven van gegevens die *niet* leesbaar zijn, maar wel interpreteerbaar? En leesbaar voor wie? Mensen, machines, of alle twee?

Ik vermoed dat een focus op interpreteerbaar de mens/machine vraag uit de weg gaat, en het concept ‘leesbaar’ geheel omvat.

Tenslotte:

informatie vindbaar, beschikbaar, leesbaar, betrouwbaar, interpreteerbaar en toekomstbestendig [p. 1]

Wat betekent “toekomstbestendig” in deze zin? Is dit een aparte eigenschap bovenop de eigenschappen “vindbaar, beschikbaar, leesbaar, betrouwbaar, interpreteerbaar”, of is toekomstbestendigheid het gevolg van die meer kernachtige eigenschappen? In het laatste geval vind ik “toekomstbestendig” een beetje een ‘weasel word’. Klinkt mooi, maar wat wordt er nou precies mee bedoeld? Bij de andere eigenschappen is de bedoeling veel duidelijker.

Lijkt mij dat je misschien wilt zeggen: data is *toekomstbestendig* wanneer ‘ie leesbaar, vindbaar, etc. is.

De SIP

Ook de SIP wordt hier benoemd. Is de SIP “specificatie” echt een integraal onderdeel van MDTO? In theorie zijn deze compleet los te gebruiken.

Bovendien vind ik veel documentatie rondom de SIP matig geschreven, en lang niet zo duidelijk als de rest van MDTO. Als je de SIP pagina ziet als onderdeel

van een grotere standaard, dan zou ik daar zeker wat redigeren.

Vraag 3

MDTO is hiermee nadrukkelijk niet bedoeld als vervanger voor de andere metagegevensstandaarden [p. 2]

Dit is en blijft goed om te benadrukken. Ik merk bij veel mensen dat ze denken dat *alle* metagegevens in MDTO horen.

Jullie hebben het in hoofdstuk 2 meermaals over data uitwisselen. Tegelijk gaat het meermaals over XML. De afgelopen 20 jaar is er veel kritiek op XML geleverd, vooral in het gebruik als uitwissel formaat, die door de meeste techneuten nog steeds als terecht wordt gezien (zie dit thread op hackernews, een vooraanstaand forum: <https://news.ycombinator.com/item?id=21960863>).

Ik betwijfel sterk of het in de toekomst een goed idee blijft om je zo op XML te blijven richten. JSON en YAML zijn wat de wereld nu gebruikt om data uit te wisselen (i.e. meetbaar populairder), en bovendien beide *veel* leesbaarder voor mensen (en computers ook trouwens, omdat de syntax compatible is met bijna alle populaire programmeertalen. Dit is nooit het geval geweest voor XML). En leesbaarheid is toch iets waar jullie om geven.

In dit artikel wordt denk ik correct beargumenteerd dat XML niet eens bedoeld is als data formaat (ja, echt). Het is natuurlijk wél gebruikt als algemeen data formaat, maar dat is stiekem foutief.

<https://www.devever.net/~hl/xml>

Een andere goede referentie voor problemen met XML:

<https://www.infoq.com/presentations/Heretical-Open-Source/>

(vooral de eerste 15 minuten van deze presentatie)

Aansluiten op ORI (de standaard voor open raadsinformatie) wordt ook makkelijker als de focus minder naar XML gaat.

Vraag 4

Mooi dat jullie TOOI en MDTO als complementair zien. Ik denk dat meer kruisbestuiving goed is voor beide standaarden. In dat opzicht laat de huidige MDTO documentatie nog wel wat te wensen over. Er komen bijvoorbeeld nul verwijzingen naar TOOI waardenlijsten voor in de vernieuwde voorbeeld bestanden.

Ik ga er vanuit dat de precieze uitwerkingen van de “symbiose” later helderder wordt.

Vraag 5

Ik denk dat de expertgroep een goede doorsnede is.

Vraag 6

MDTO is vrij minimalistisch. Maar dat is ook de kracht: het focust zich echt op gegevens die *noodzakelijk* zijn voor duurzame toegankelijkheid van informatie. Hiermee is het inderdaad toepasbaar binnen alle overheden.

Als ontwerper van een metadata standaard vond ik het makkelijk om te bedenken en te implementeren hoe men deze standaard met MDTO zou kunnen combineren. Hiermee sluit ik me aan bij sectie 5.1.3

Over 5.1.4.5: in MDTO kun je opnemen welke actoren betrokken zijn geweest bij een informatieobject. Middels **beperkingGebruik** kun je volgens de huidige definitie van het element restricties op een informatieobject leggen, maar *niet* op de metadata rondom het informatieobject (bijv. op de MDTO XML zelf). Hiermee kan er bijvoorbeeld onbedoeld bekend worden hoe bepaalde personen zich verhielden tot het informatieobject, wat volgens mij als gevoelige informatie kan gelden. (misschien is er altijd impliciet bedoeld: “alle restricties van toepassing op het informatieobject zijn ook van toepassing op bijbehorende metadata”, maar dit word nergens duidelijk gestipuleerd)

Met **beperkingGebruik** kun je in theorie veel regelen qua afscherming, maar is het gebruik hiervan momenteel gestandaardiseerd genoeg om bijvoorbeeld te kunnen zeggen “op de metadata zelf liggen geen restricties” of “alleen bepaalde mensen kunnen dit document in zien”? Ik denk van niet. De bestaande begrippenlijst geeft je eigenlijk alleen de mogelijkheid om te zeggen “nee” (oftewel, ‘geen beperking’), “misschien” (oftewel, ‘nader te bepalen’), “niks precies over te zeggen” (oftewel, ‘overig’).

Hier kun je geen implementatie om heen bouwen. Daarvoor moet er echt meer gestandaardiseerd worden (nu zal in de praktijk iedereen maar wat invullen bij de beperkingen). En omdat die standaardisatie nu ontbreekt, sluit ik me niet helemaal aan bij de adviezen in 5.1.4.4 en 5.1.4.5. Jullie benoemen weliswaar een bestaande classificatie “Departementaal Vertrouwelijk (DepV)”, maar is die breed toepasbaar genoeg? Welke begrippenlijsten zijn mogelijk nog meer van toepassing?

Met goede implementatie zijn de risico’s beheersbaar [p. 12]

Zeker, maar het is jullie taak om implementaties in goede banen te leiden door ambiguïteiten zoals bovenstaande te voorkomen.

Hoe zouden jullie bijvoorbeeld aanraden om met behulp van MDTO _meta_ gegevens (ipv reguliere gegevens) af te screenen? Of is dat jullie inziens niet nodig?

Vraag 7

Over 5.2.2.2: TOOI hanteert git voor het bijhouden van versie wijzigingen. Wat mij betreft is dat *good engineering practice* (het paper Best practices for Technical standard creation onderschrijft het belang van *good enineering*), en is bovendien transparanter dan de huidige pagina. Het is met zekerheid voorgekomen dat veranderingen aan de XML voorbeelden/XSD niet zijn genoemd op de wijzigingen pagina. Meerdere mensen hebben bijvoorbeeld gemerkt dat de voorbeeldbestanden (de eerste variant) ooit invalide waren. Dat is gecorrigeerd, maar is nergens terug te zien Ik heb in mail contact met Wout een ander voorbeeld van een verborgen wijziging beschreven (een verbetering in een XML comment).

Ik heb geverifieerd dat dit allemaal echt heeft plaatsgevonden door in jullie webarchief te zoeken (<https://nationaalarchief.sitearchief.nl/#archive>).

Door met Github/Gitlab te werken ga je dit soort haiten in transparantie tegen. Toegegeven, de dingen die niet op de wijzigingen pagina te zien zijn klein, maar ze scheppen wel verwarring en kosten tijd. Een open — in de zin van transparante — standaard zou beter kunnen en moeten doen.

Bovendien maakt een Github/Gitlab bredere samenwerking mogelijk, inclusief een gedeelde plek voor technische feedback.

Over 5.2.4.1:

MDTO werkt volgens de principes van semantic versioning . . . Voor MDTO zijn tot nu toe uitsluitend Z-wijzigingen doorgevoerd [p. 13]

Hier kun je aan twifelen. Semantic versioning zegt zelf:

Patch version Z (x.y.Z | x > 0) MUST be incremented if only backward compatible bug fixes are introduced. A bug fix is defined as an internal change that fixes incorrect behavior.

De vraag is of de MDTO XSD inderdaad volgens dit principe is geversioneerd. Niet alle MDTO XML die valide is volgens de 1.0.0 XSD is valide volgens de 1.0.1 XSD. Dus deze “bugfix” 1.0.1 release — als je ’m zo wilt/kunt noemen — was niet backwards compatible.

Anderzijds vind ik de vraag wat een bug is in de context van een XML-schema moeilijk. Misschien zijn dat gevallen waar het schema niet overeenstemt met de (human-readable) documentatie (zie ook <https://ori-a.nl/downloads#semantisch-versioneren>)

Kort gezegd: ik denk dat hier nog wel wat duidelijkheid geschept kan worden.

(een bug in een programma/API implementatie kan ook crash zijn. Z-wijzigingen zijn deels bedoeld voor dat soort wijzigingen. Maar dit soort wijzigingen vinden niet echt plaats in een XSD, omdat een XSD geen software is)

Vraag 8

De documentatie kan echt overzichtelijker en beter. Hierin sluit ik me dus aan bij “de experts”.

Kijk eens naar andere documentatie sites voor technische/metadata standaarden, zou ik zeggen. En dan niet alleen de vormgeving, maar ook hoe teksten zijn opgebouwd, en wat ze allemaal wel/niet bevatten (ik mis zelf een tutorial/prosa uitleg van de XML, zie bijv. <https://ori-a.nl/hoe-werkt-ori-a>). Concreet heb ik het gevoel dat sommige teksten nauwelijks geredigeerd zijn (de SIP pagina), en dat je teveel moet klikken voordat je vindt wat je zoekt.

Voor het navigeerbaarder maken van de documentatie kun je misschien denken aan een compleet aparta documentatie site, die niet 100% hoeft te voldoen aan de restricties van de huidige NA website (althoewel dat ook weer nadelen heeft). Er zijn heel projecten waarmee je mooie en overzichtelijk documentatie sites mee kan maken. Zoiets maakt technici/ontwikkelaars ook meteen vertrouwt met je site.

Vraag 9

Bieden meerdere leveranciers ondersteuning voor de standaard?

De REE (MAIS) biedt ook al vrij lang ondersteuning.

Vraag 10

Ik mis geen grote spelers.

Vraag 12

Ik ken zelf nog meer archieven die ook wat met MDTO doen! Dus ja.

Vraag 15

Ik sluit me sterk aan bij bullet #2. Ook denk ik inderdaad dat het taalgebruik beter kan — mensen lezen op het internet meestal vrij vluchtig, dus schrijf ook voor dat publiek.

Het toevoegen van meer validatietesten kan misschien wel interessant zijn, maar wat betekent dat bovenop voldoen aan de XSD? Niet alleen valide, maar ook inhoudelijk logisch, of iets in die richting? Hier kan je nog wel wat duidelijker over zijn.

(kan natuurlijk ook zijn dat afnemers niet altijd technologisch in staat iets met een XSD te doen, maar dat is gedeeltelijk een human resources probleem bij

overheden, niet een probleem van het NA/MDTO. Tegelijk is het probleem ook dat XSD al decennia geen populaire technologie meer is, en dat de tooling dus een beetje obscuur is. De beheerder van xmllint, een veel gebruikt programma voor XML validatie middels een XSD, is deze maand bijv. afgestapt. Zie ook mijn antwoord op vraag 3)