



Multi-Layer Content Modelling to the Rescue

XATAPULT

CONTENT ENGINEERING

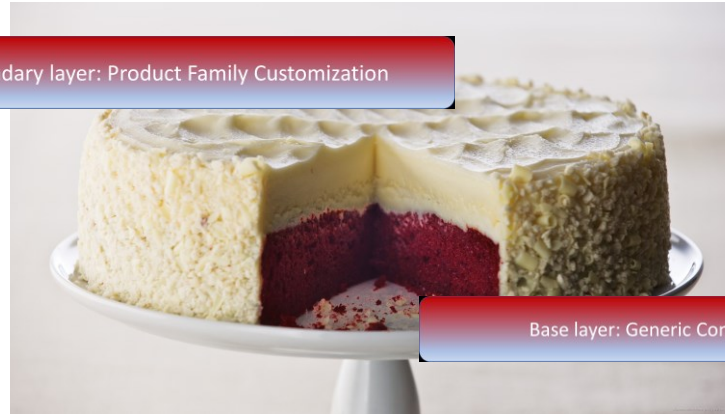
Erik Siegel
erik@xatapult.nl

Multi-Layer Content Modelling?

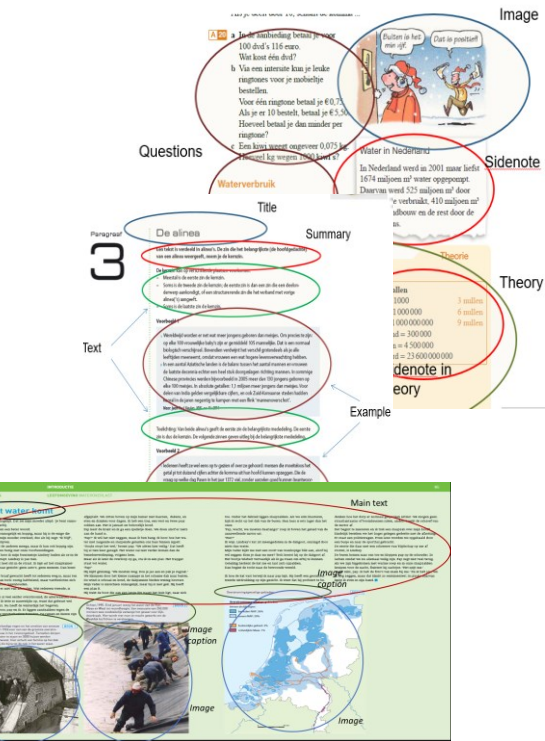


Implementation

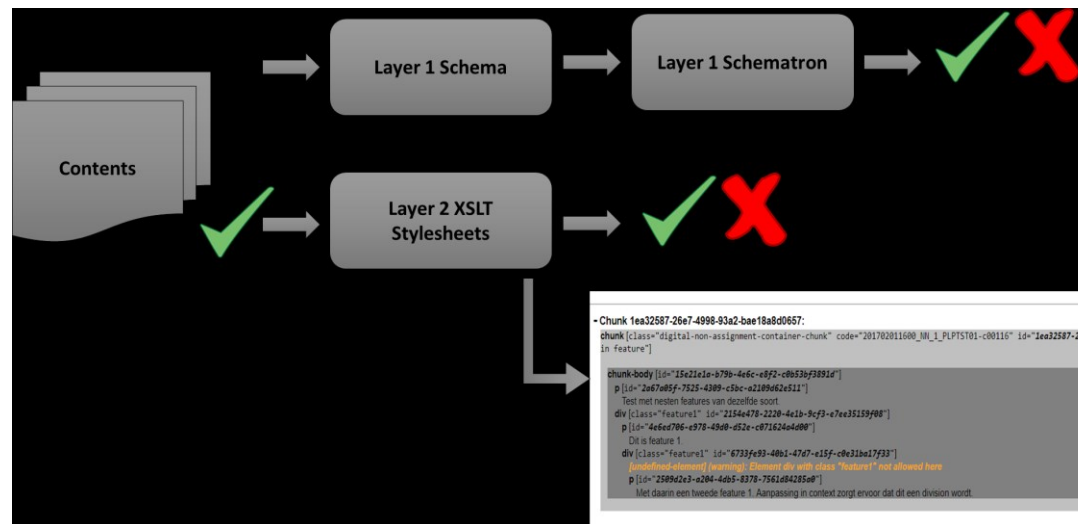
Secondary layer: Product Family Customization



Base layer: Generic Content Model



Use case



Take-aways

The use-case: Educational Publishing


Schooltype	# years
VMBO	4
HAVO	5
VWO	6
Total:	15

What?	#
Subjects	10
Books per subject per year	3
Pages per book	150
Website components per subject per year	500

- **450 books**
- **67.500 pages**
- **75.000 website components**

Example 1: Secondary Education book 1

Image



Questions

A 20 a In de aanbieding betaal je voor 100 dvd's 116 euro. Wat kost één dvd?
b Via een intersite kun je leuke ringtones voor je mobieltje bestellen. Voor één ringtone betaal je €0,75. Als je er 10 bestelt, betaal je €5,50. Hoeveel betaal je dan minder per ringtone?
c Een kiwi weegt ongeveer 0,075 kg. Hoeveel kg wegen 1000 kiwi's?

Waterverbruik

O 21 a Lees het krantenartikel hiernaast
b Hoeveel miljoen m³ water werd in de huishoudens verbruikt?

Sidenote

Water in Nederland

In Nederland werd in 2001 maar liefst 1674 miljoen m³ water opgepompt. Daarvan werd 525 miljoen m³ door de industrie verbruikt, 410 miljoen m³ door de landbouw en de rest door de huishoudens.

Theorie

Grote getallen

Voluit geschreven is 1674 miljoen 1 674 000 000. In de krant staat het woord miljoen, omdat je met al die nullen gemakkelijk de tel kwijt raakt. Bij het berekenen van $15 \times 12,5$ miljard tik je op je rekenmachine in $15 \times 12,5 =$. De rekenmachine geeft 187,5. Het antwoord is dus 187,5 miljard. Je hoeft niet alle nullen van miljard in te tikken.

Sidenote in theory

Grote getallen	
duizend = 1000	3 nullen
miljoen = 1 000 000	6 nullen
miljard = 1 000 000 000	9 nullen
300 duizend = 300 000	
4,5 miljoen = 4 500 000	
23,6 miljard = 23 600 000 000	

Theory

Example 2: Secondary Education book 2

84 INTRODUCTIE 85

HOOFDSTUK 4 LEEFOMGEVING WATEROVERLAST

Title
Als het water komt

Main text

Mijn vader is onmogelijk. Dat zei mijn moeder altijd: 'Je bent onmogelijk.' Ze lachte erbij.
Koppig is misschien een beter woord.
Pap is misschien onmogelijk en koppig, maar hij is de enige die ik nog heb sinds mijn moeder overleed, dus als hij zegt: 'Ik blijf', besluit ik ook te blijven.
Hij wil dat ik met de anderen meega, maar ik kan ook koppig zijn. De hele dag zijn we bezig met onze voorbereidingen.
Tegen de middag hoor ik mijn buurmeisje Lindsey huilen als ze in de volgeladen auto stapt. Lindsey is pas tien.
Daarna wordt het heel stil in de straat. Ik kijk uit het slaapkamer-raam. Het is een raar gezicht: geen auto's, geen mensen. Dan komt het water.
Het lijkt alsof het braaf gewacht heeft tot iedereen weg is, maar ten slotte komt het dan toch: rustig kabbelend, maar vastbesloten zich nergens door te laten tegenhouden.
Ik kan mijn ogen er niet van afhouden. Wat iedereen vreesde, is gebeurd.
De zomerdijk was al veel eerder overstroomd; de uiterwaarden stonden onder water. Ik lette er nauwelijks op, want dat gebeurt wel vaker. Dit is anders. Nu heeft de winterdijk het begeven.
We zijn er klaar voor, pap en ik. Er liggen zandzakken tegen de deurdeur en op de vensterbanken beneden. De ramen en kieren zijn

afgeplakt. We zitten boven op mijn kamer met kaarsen, dekens, en eten en drinken voor dagen. Ik heb een trui, een vest en twee paar sokken aan. Het is januari en behoorlijk koud.
Pap leest de krant en ik ga een spelletje doen. We doen alsof er niets aan de hand is.
'Pap?' Ik wil het niet zeggen, maar ik ben bang. Ik hoor hoe het water met zuigende en slurpende geluiden ons huis binnen sijpelt.
'Straks stopt het wel,' bromt pap. 'We zitten hier veilig.' Dat heeft hij al tien keer gezegd. Het water zal niet verder komen dan de benedenverdieping, volgens hem.
Maar als ik later de overloop op ga, sta ik in een plas. Het trapgat staat vol water.
'Pap!'
Hij kijkt grimmig. 'We moeten weg. Doe je jas aan en pak je rugzak.' We klimmen door het kleine raampje in het schuine dak naar buiten. De wind is schraal en koud, de dakpannen bieden weinig houvast. Mijn vader is misschien onmogelijk, maar hij is niet gek. Hij heeft een plan B.
Hij trekt de boot die aan een lange lijn naast het huis ligt, naar zich toe. Onder het dekzeil liggen slaapzakken. Als we erin klauteren, kijk ik recht op het dak van de burens. Hun huis is iets lager dan het onze.
'Pap, wacht, we moeten daarlangs!' roep ik boven het geluid van de aanzwellende motor uit.
'Wat?'
Ik wijs. Lindsey's kat zit ineengedoken in de dakgoot, omringd door niets dan water.
Mijn vader kijkt me met een soort van wanhopige blik aan, alsof hij wil zeggen: Kom je daar nu mee? Toch koerst hij op de dakgoot af. Het bootje wiebelt verwaarloosd als ik ga staan om erbij te kunnen. Gelukkig herkent de kat me en laat zich oppakken.
Dan begint de tocht naar de bewoonde wereld.

Ik hou de kat vast terwijl ik naar pap kijk. Hij heeft een geconcentreerde uitdrukking op zijn gezicht. Ik weet dat hij probeert te bedenken hoe het dorp er normaal gesproken uitziet. We mogen geen straatlantaarns of boomkruinen raken, anders breekt de schroef van de motor af.
Het begint te miezeren en ik trek een slaapzak over mijn hoofd. Eindelijk bereiken we het hoger gelegen gedeelte met de afzetting. Er staat een politiewagen. Even later worden we opgehaald door een busje en naar de sporthal gebracht.
De eerste die daar met een schreeuw van blijdschap op me af stormt, is Lindsey.
De burens komen naar ons toe en kloppen pap op de schouder. Ze lachen omdat we nu allemaal veilig zijn. Pap zegt niet veel terug. Als we zijn bijgekomen met warme soep en in onze slaapzakken kruipen voor de nacht, fluistert hij zachtjes: 'Het spijt me.'
'Wat niks, pap. Ik heb de foto's van mam bij me.' En ik heb jou, wil ik nog zeggen, maar dat klinkt zo sentimenteel. In plaats daarvan knijp ik even in zijn hand. ■

Image caption
Ochten, 1995. Eind januari steeg het water van de Rijn, Maas en Waal tot recordhoogte. Een evacuatie van 250.000 mensen was noodzakelijk vanwege het gevaar voor dijkdoorbraak. Hier wordt met man en macht gewerkt om de Waaldijk bij Ochten te verstevigen. **BRON 2**

Image caption
Mook, 1926. Overvloedige regen en het smelten van sneeuw zorgden in januari 1926 voor een van de grootste overstromingen van de eeuw in het rivierengebied. Tientallen dorpen kwamen onder water te staan en 3000 huizen werden beschadigd of verwoest. Hier schuilt een familie op het dak van hun boerderij die bijna tot de nok in het water staat. **BRON 1**

Image caption
Overstromingsgevoelige gebieden **BRON 3**


Image

Image

Image

Differences matter!

Image



Questions

A 20 a In de aanbieding betaal je voor 100 dvd's 116 euro. Wat kost één dvd?
 b Via een internet site je leuke ringtonen voor je mobiele telefoon bestellen. Voor één ringtone betaal je €0,75. Als je er 10 bestelt, betaal je €5,50. Hoeveel betaal je dan minder per ringtone?
 c Een kwart weegt ongeveer 0,075 kg. Hoeveel kg wegen 1000 kwart's?

Water in Nederland

In Nederland werd in 2001 maar liefst 1674 miljoen m³ water opgepompt. Daarvan werd 525 miljoen m³ door de industrie verbruikt, 410 miljoen m³ door de landbouw en de rest door de huishoudens.

Waterverbruik

a Lees het krantenartikel hiernaast.
 b Hoeveel miljoen m³ water werden in de huishoudens verbruikt?

Side note

Grote getallen

Voluit geschreven is 1674 miljoen 1 674 000 000.
 In de krant staat het woord miljoen, omdat je met al die nullen gemakkelijk de tel kwijt raakt.
 Bij het berekenen van $15 \times 12,5$ miljard tik je op je rekenmachine in 15 \times 12,5 = .
 De rekenmachine geeft 187,5.
 Het antwoord is dus 187,5 miljard.
 Je hoeft niet alle nullen van miljard

Theorie

Grote getallen

duizend = 1000	3 miljoen
miljoen = 1 000 000	6 miljoen
miljard = 1 000 000 000	9 miljoen
300 duizend = 300 000	
4,5 miljoen = 4 500 000	
23 miljard = 23 600 000 000	

Sidenote in Theory



Title

Als het water komt

Main text

Op zaterdag 11 oktober 2008... (Text continues with details about water management in the Netherlands, mentioning the 'Waterwet' and the 'Rijkswaterstaat'.)


Image caption

Image



Image caption

Image




Support for Interactive Website Contents

The screenshot shows a website interface with a left-hand bar containing navigation and information. The main content area features a drag-and-drop interaction where words from a list on the left are being moved to a grid on the right. The words in the list include 'proef', 'trainen', 'zichtbaarheid', 'primaten (al. 1)', 'portret (al. 1)', 'experiment (al. 3)', 'onderscheid (al. 4)', 'irrelevant (al. 5)', and 'voorheen (al. 6)'. The grid contains words like 'vrienden', 'vroeger', 'verschil', 'afbeelding van een gezicht', 'apen', and 'onbelangrijk'. A blue circle highlights the left-hand bar, and another blue circle highlights the drag-and-drop interaction area.

Drag and drop interactions

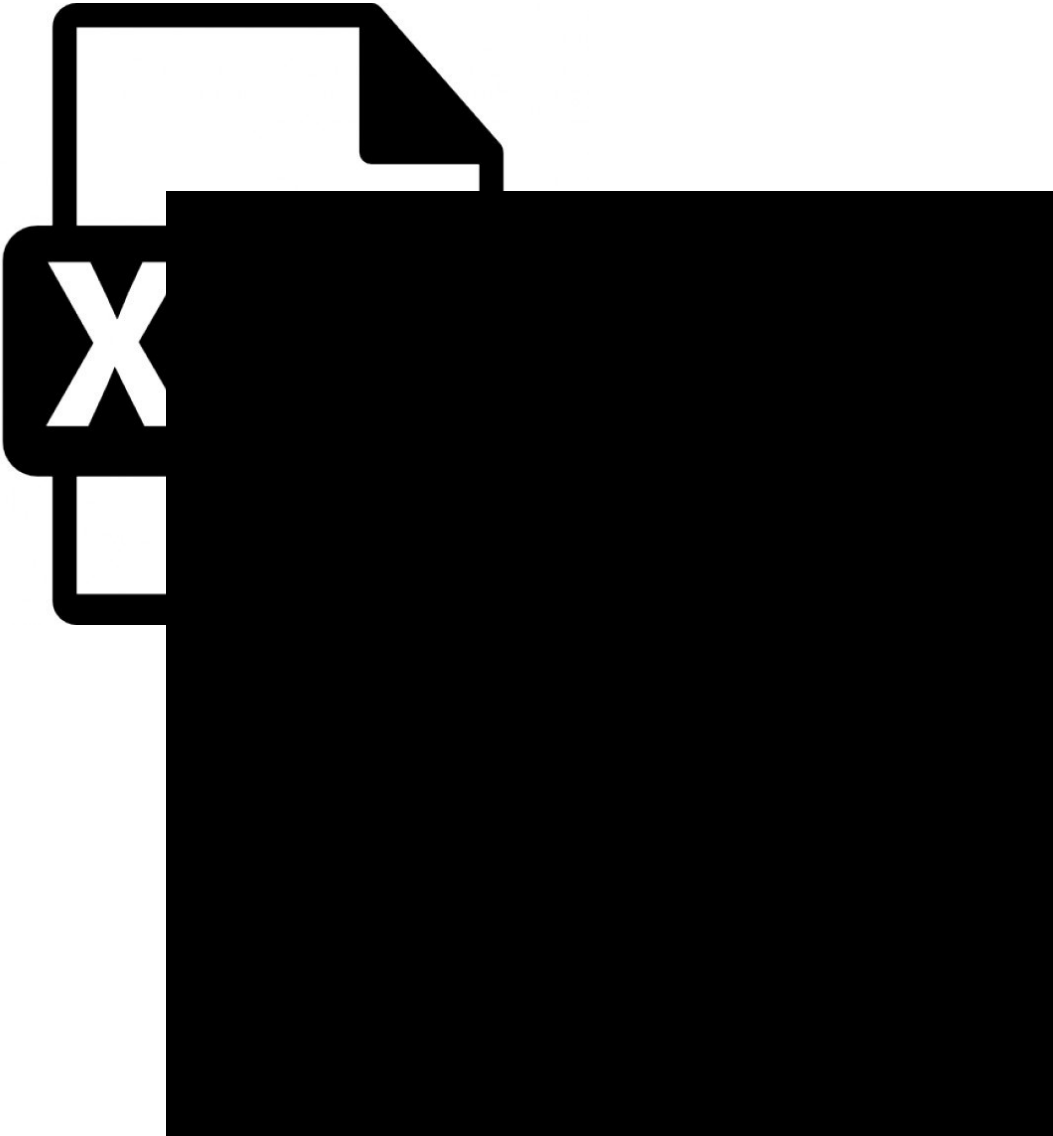
Left-hand bar with information

Inline choice interactions

The screenshot shows a website interface with a purple header and a main content area. The main content area features a list of four multiple-choice questions in French. Each question has a dropdown menu with a green checkmark or a red X indicating the correct or incorrect answer. The questions are: 1. 'Wat zeg je als je iemand begroet?' (answer: 'bonjour'), 2. 'Wat zeg je als je afscheid neemt van iemand?' (answer: 'au revoir'), 3. 'Wat zeg je als je iemand bedankt?' (answer: 's'il vous plait'), 4. 'In het Frans heb je twee manieren om *alstublieft* te zeggen: *voilà* en *s'il vous plait*. Wat zeg je als je iets geeft?' (answer: 'voilà'). The interface also includes a navigation bar at the bottom with 'Vorige' and 'Volgende' buttons, and a 'Inleveren' button. A blue circle highlights the inline choice interaction area.

Juiste antwoord Mijn antwoord

Support for Multiple Output Channels



A content solution has to deal with:

- High volume
- Different product families with
 - Different structural models
 - Varying markup demands
- Complex interactions and other data structures
- Reuse
- Multiple output channels



Solution properties?

Not:

- Multiple content-models
- Multiple input tools/configurations
- Multiple output toolchains
- Content Silo's



But:

- Single content-model
- Single input tool with minimum number of configurations
- Minimum number of output toolchains
- No content Silo's



Solution? Multi-Layered Content Modelling

Secondary layer: Product Family Customization

Base layer: Generic Content Model

Similar Approaches

- DITA Specializations
- RelaxNG Customizations
- Epischemas
- Schematron

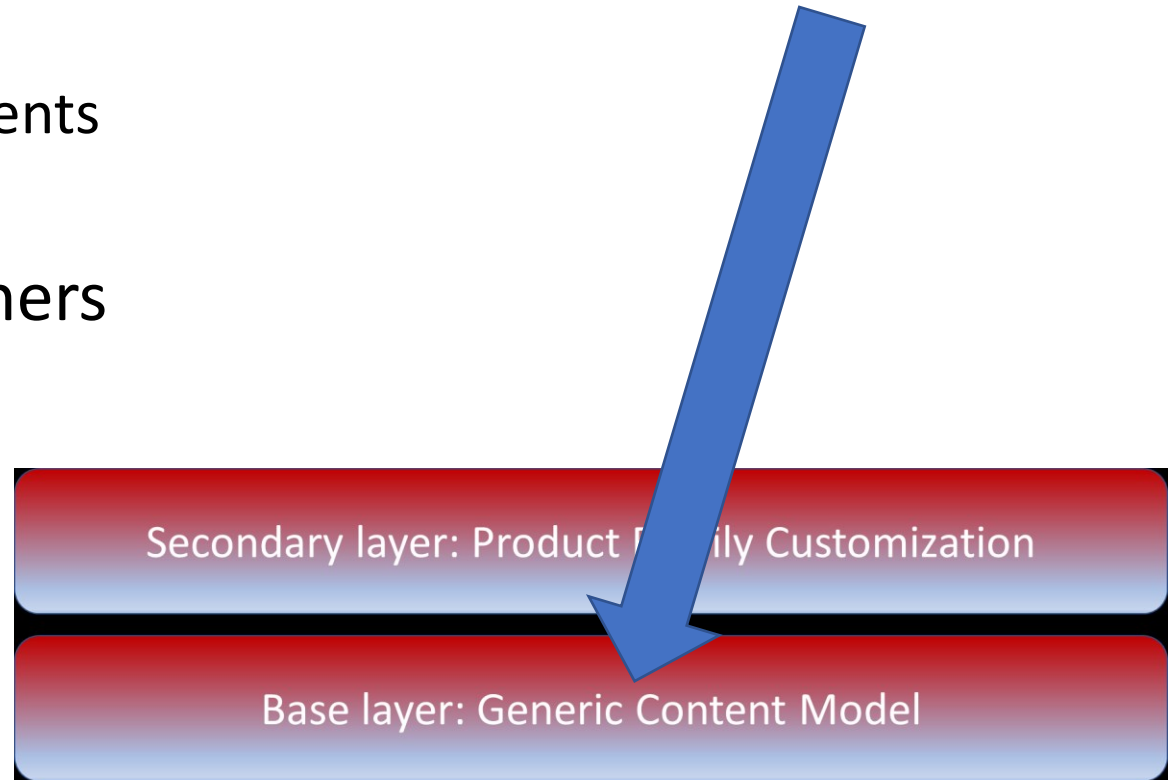


These are all technically based languages and approaches, about XML itself

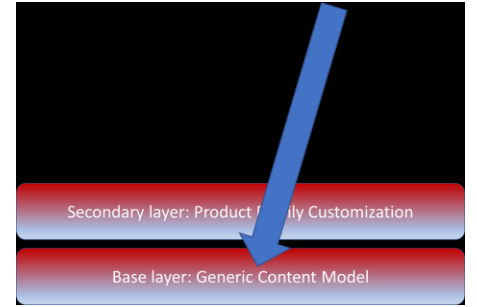
We wanted something more functional and domain driven

Base Layer: Generic Content Model

- Bespoke
 - Functional and technical requirements
 - Allow for secondary layer
- Stole a lot of good ideas from others
 - DITA maps/topics
 - TEI structures
 - QTI interactions
- Validation included
 - Schemas
 - Schematron



Base Content Model Example 1: Structure



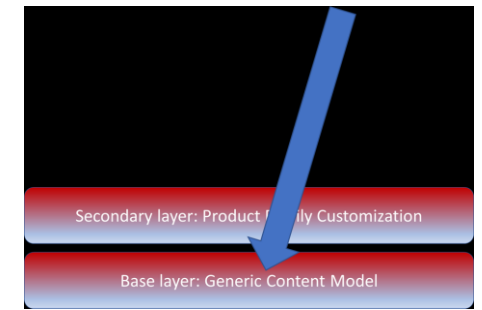
Not:

```
<chapter>
  ...
  <subsection1>
    ...
    <subsection2>
      ...
      </subsection2>
    </subsection1>
</chapter>
```

```
<section class="chapter">
  ...
  <section class="subsection1">
    ...
    <section class="subsection2">
      ...
      </section>
    <section>
      ...
    </section>
```

Minimal number of structuring elements (~1), specialize using @class

Base Content Model Example 2: Markup



Not:

```
<h1>A <i>header</i> that is <b>bold</b></h1>
```

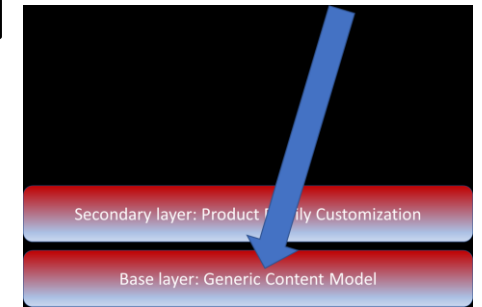
```
<p class="header1">A <em class="italic">header</em> that is  
<em class="bold">bold</em></p>
```

Minimal number of markup elements (~12), specialize using @class

Base Content Model Example 3: Metadata

```
<multiple-choice-interaction id="mc-1">  
  ...  
</multiple-choice-interaction>
```

```
<metadata-set about-idrefs="mc-1" class="interaction-info">  
  <properties>  
    <property property-name="nr-of-retries">  
      <value>2</value>  
    </property>  
    <property property-name="score">  
      <value>6.5</value>  
    </property>  
  </properties>  
</metadata-set>
```



Metadata organized in sets

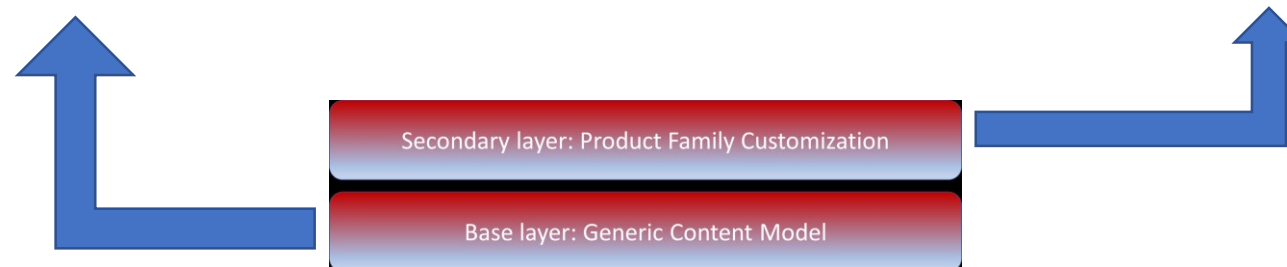
The Layered Approach

What	Base layer	Secondary layer
Structure	<ul style="list-style-type: none">• Any structure to any depth• Any class	<ul style="list-style-type: none">• Define the section classes that can be used• Define how sections with these classes can be nested• Limit the number of occurrences of sections with a certain class



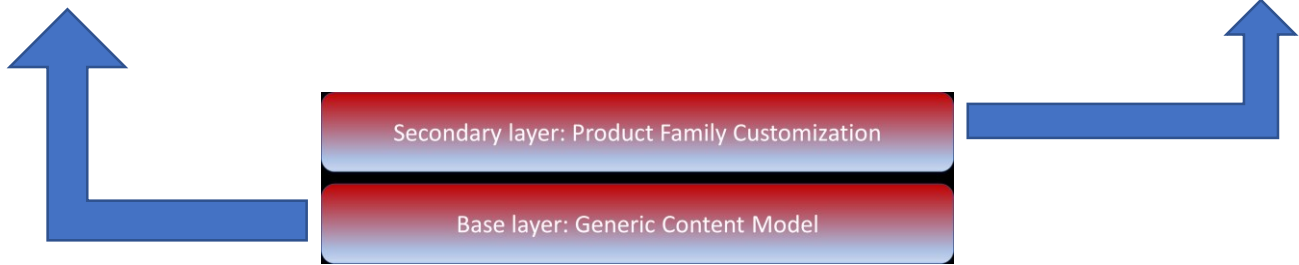
The Layered Approach

What	Base layer	Secondary layer
Structure	<ul style="list-style-type: none">• Any structure to any depth• Any class	<ul style="list-style-type: none">• Define the section classes that can be used• Define how sections with these classes can be nested• Limit the number of occurrences of sections with a certain class
Markup	<ul style="list-style-type: none">• Limited number of elements• Any class	<ul style="list-style-type: none">• Depending on where you are in the document, limit element and class usage



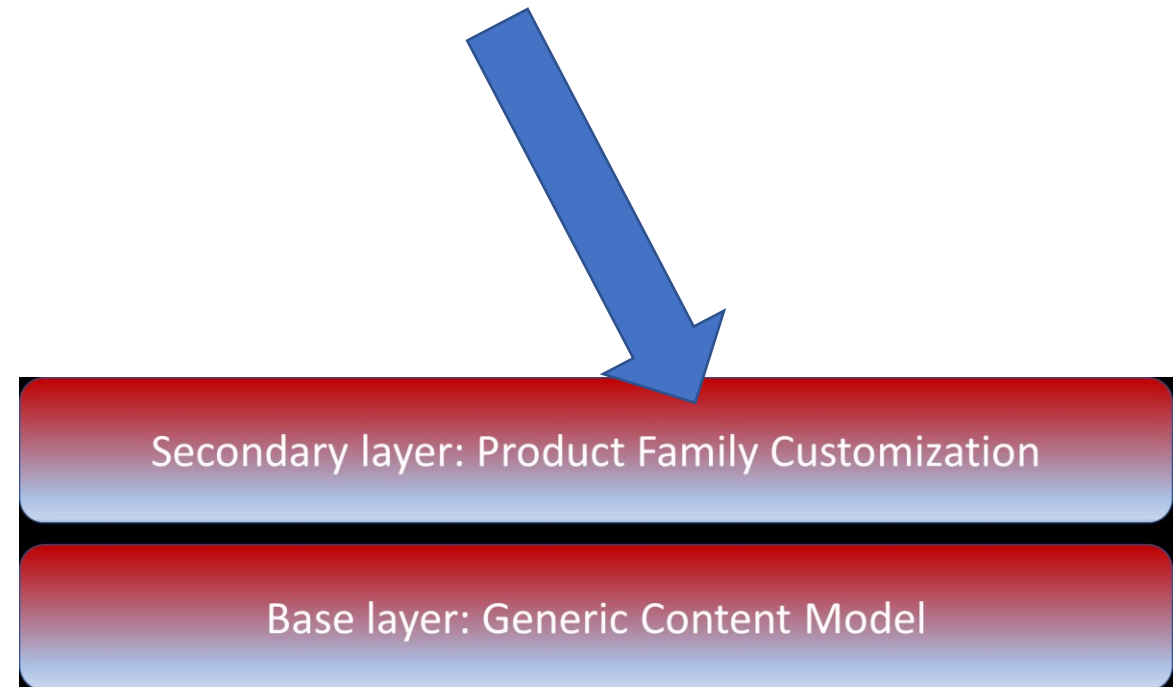
The Layered Approach

What	Base layer	Secondary layer
Structure	<ul style="list-style-type: none"> Any structure to any depth Any class 	<ul style="list-style-type: none"> Define the section classes that can be used Define how sections with these classes can be nested Limit the number of occurrences of sections with a certain class
Markup	<ul style="list-style-type: none"> Limited number of elements Any class 	<ul style="list-style-type: none"> Depending on where you are in the document, limit element and class usage
Metadata	<ul style="list-style-type: none"> Metadata-sets can be attached to elements in the contents Any property, any value type 	<ul style="list-style-type: none"> Define what is in a specific metadata-set (properties, value types, nr. of occurrences, etc.) Define to which elements you can bind a specific metadata-set

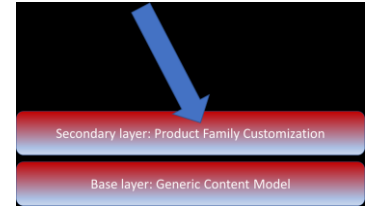


Secondary Layer Implementation: DSVL (Domain Specific Validation Language)

- **Not** a generic validation language
 - Works only for valid base layer contents
 - Language based on features of the base model
- Serves many masters:
 - XML validation
 - Editor
 - Toolchain
 - Content specialist
- Limited functionality, to what is most needed
- Semantic, less technical, oriented



DSVL Structure Definition



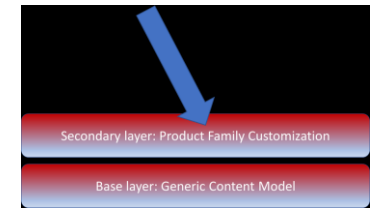
```
<section class="chapter">
...
<section class="subsection1">
...
<section class="subsection2">
...
</section>
<section>
</section>
```

Secondary layer defines:

- Nesting
- Classnames
- Occurrences
- ...

```
<section-definition class="chapter" max-occurs="10">
...
<section-definition class="subsection1">
...
<section-definition class="subsection2" min-occurs="0">
...
</section-definition>
</section-definition>
</section-definition>
```

DSVL Markup Definition



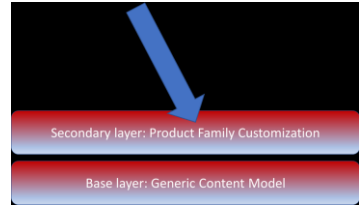
```
<p class="header1">A header that is  
  bold</p>
```

Secondary layer defines:

- Elements
- Allowed classes for elements
- Word and character counts
- ...

```
<element-definition element-names="p" class="header1"  
                    word-count-max="15"/>  
<element-definition element-names="em" class="italic"/>  
<element-definition element-names="em" class="bold"/>
```

DSVL Metadata Definition



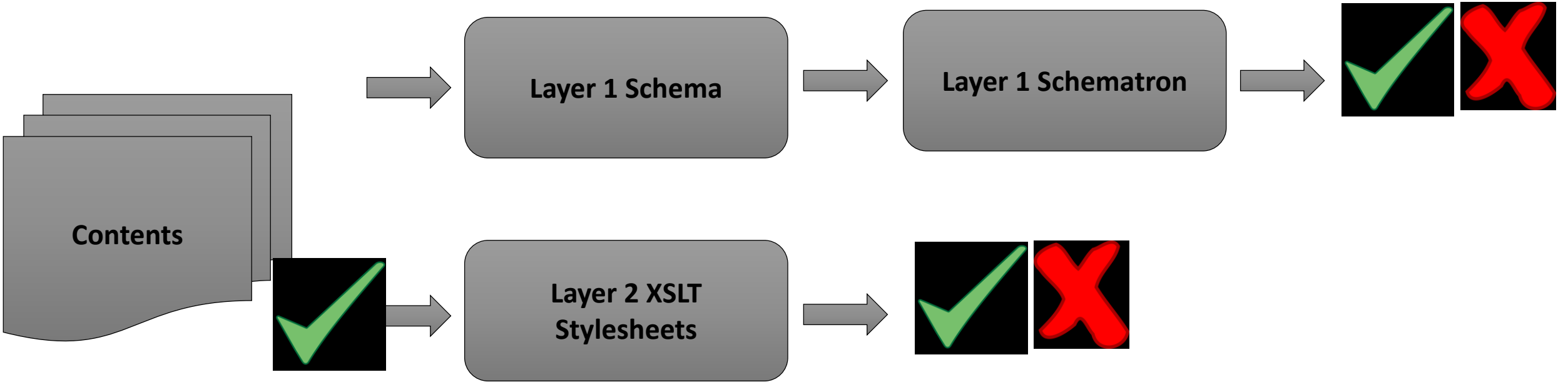
```
<metadata-set about-idrefs="mc-1" class="interaction-info">
  <properties>
    <property property-name="nr-of-retries">
      <value>2</value>
    </property>
    ...
  </properties>
</metadata-set>
```

Secondary layer
defines:

- Metadata-set class
- Property
 - Names
 - Data types
 - ...

```
<metadata-set-definition class="interaction-info">
  <properties-definition>
    <property-definition property-name="nr-of-retries">
      <value-definition datatype="integer"/>
    </property-definition>
    <property-definition property-name="score">
      <value-definition datatype="double"/>
    </property-definition>
  </properties-definition>
</metadata-set-definition>
```

Usage: Validation



- Chunk 1ea32587-26e7-4998-93a2-bae18a8d0657:

```
chunk [class="digital-non-assignment-container-chunk" code="201702011600_NN_1_PLPTST01-c00116" id="1ea32587-26e7-in feature"]
```

```
chunk-body [id="15e21e1a-b79b-4e6c-e8f2-c0b53bf3891d"]
```

```
p [id="2a67a05f-7525-4309-c5bc-a2109d62e511"]
```

```
Test met nesten features van dezelfde soort.
```

```
div [class="feature1" id="2154e478-2220-4e1b-9cf3-e7ee35159f08"]
```

```
p [id="4e6ed706-e978-49d0-d52e-c071624a4d00"]
```

```
Dit is feature 1.
```

```
div [class="feature1" id="6733fe93-40b1-47d7-e15f-c0e31ba17f33"]
```

```
[undefined-element] (warning): Element div with class "feature1" not allowed here
```

```
p [id="2509d2e3-a204-4db5-8378-7561d84285a0"]
```

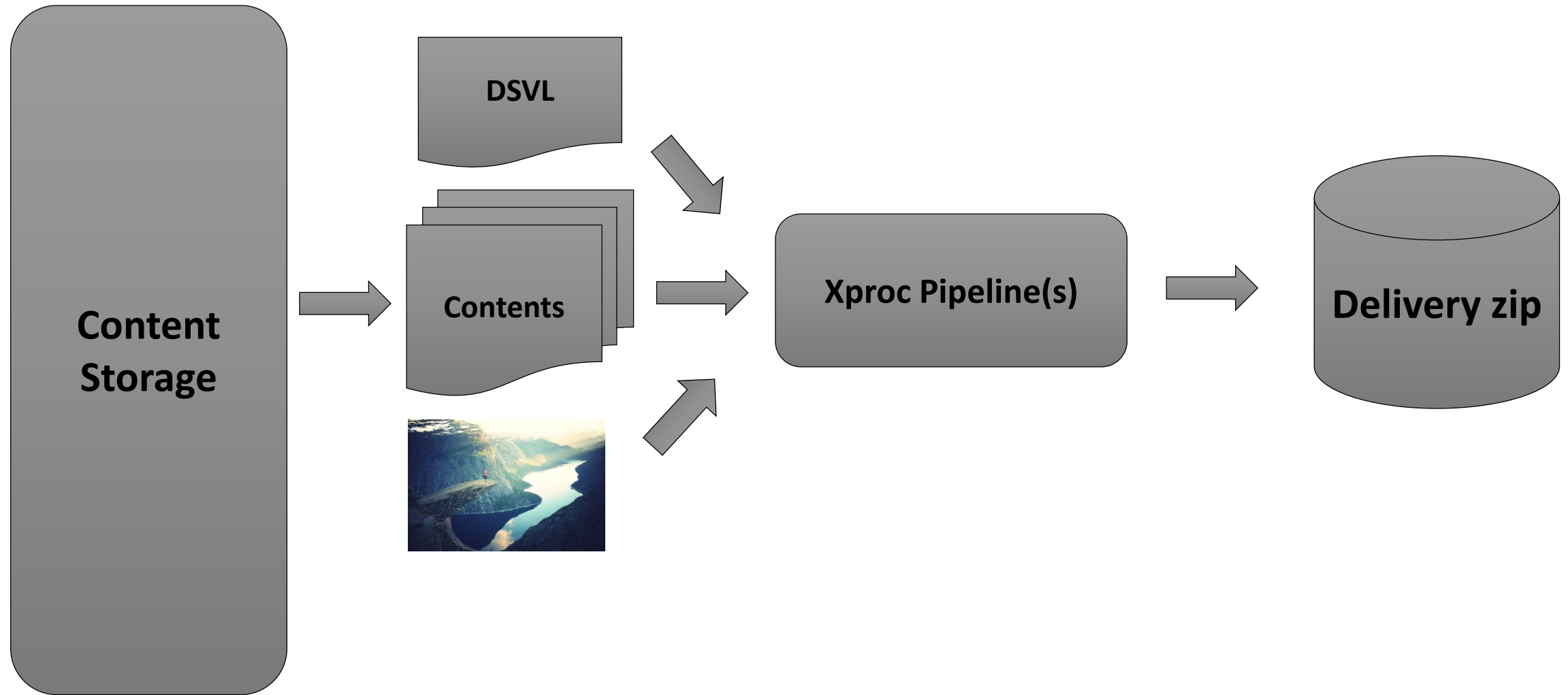
```
Met daarin een tweede feature 1. Aanpassing in context zorgt ervoor dat dit een division wordt.
```


Usage: Editor Configuration

The screenshot shows the editor configuration interface for 'Editing chunk mcsa'. The top navigation bar includes 'START', 'CHUNK', and 'VIEW' tabs. Below this is a toolbar with various editing tools: Interaction, Formatting, Inlines, Subscript, Superscript, Link, Blocks, Divisions, Lists, Track changes, Find, and Preview. A secondary toolbar contains Tabel, Assets, Math, and Feedback. The main workspace is divided into several sections: 'ASSIGNMENT CHUNK' with a 'TIP' box containing the text 'Tip toegevoegd met Divisions - tip. Deze tip staat boven het roze assignment-veld.'; an 'ASSIGNMENT' box with a 'MULTIPLE CHOICE INTERACTIE' containing the question 'Met welke zin kun je aan je luisteraars laten merken dat je ergens enthousiast over bent?'; and an 'INTERACTION-CONTAINER' with a 'MULTIPLE CHOICE INTERACTIE' containing a question 'Met welke zin kun je aan je luisteraars laten merken dat je ergens enthousiast over bent?' and a radio button option 'Alleen deze week nog voor een lagere prijs!'. A 'FEEDBACK' box is also visible. On the right side, there is a vertical sidebar with tabs for 'PROPERTIES', 'METADATA', 'REVIEW', and 'VALIDATION'. The bottom status bar shows 'Assignment chunk', 'Body', and 'No content issues'.

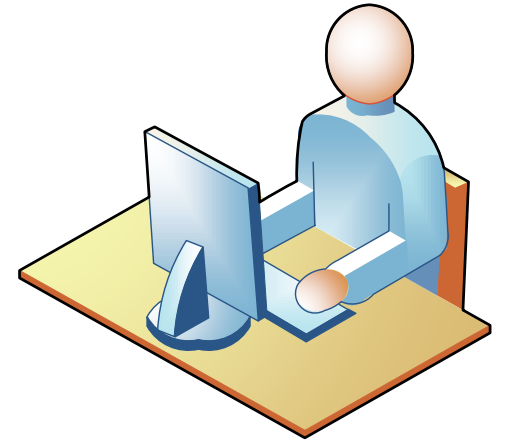
- Structure editor home grown
- Content editor Fonto
- Both adapt to DSVL in use

Usage: Delivery toolchain



DSVL Creation

- Train DSVL Designers
 - XML Skills
 - Basic XSLT skills
- DSVL Designers talk with publishers and editors
- DSVL Designers create:
 - DSVL
 - Sample contents
 - Rendering (preview)
- Evaluate



DSVL Lessons and Takeaways

- Make your DSVL a true **DSVL**
 - Reduce impedance between key users, DSVL designers and base model
- Don't overengineer the DSVL
 - Don't re-invent schema languages
 - Leave some to business rules
- Don't underengineer the DSVL
 - Allow for some abstraction
- Allow for refactoring



Implementation Lessons and Take Aways

- You need a way to validate content
 - For checking hand-crafted content
 - For validating your tools
- Current implementation in IDEs takes an extra step
- Re-use still not seamless
 - But definitely easier



Co

Mul

- Yo
- All

Allo

- Cu
- Ad
- De



Modelling can be a useful solution when:
families to support
nt models

Other components more easily
th less trouble

:



Thanks!

XATAPULT

CONTENT ENGINEERING

Erik Siegel
erik@xatapult.nl