

Looking for Rumpelstiltskin

SPINNING STRAW INTO GOLD



PRESENTED BY:
Mary McRae
Orbis Technologies

The Project

HEALTHCARE INFORMATION PROVIDER

- 20+ product offerings; released three new offerings in the past 12 months
- tens of thousands of fairly short articles (1 or 2 paragraphs to several pages in length)
 - Microsoft Word, HTML, HTML5
- proprietary systems used by product teams; nearing end of life

CUSTOMERS

- healthcare facilities, insurers, and practitioners/clinicians
- constantly looking for new content, new ways to deliver content, and increase end-user engagement (consumer, patient, caregiver)

The Project (cont'd)

TO MEET CURRENT AND FUTURE DEMANDS:

- eliminate departmental silos
- create a unified content architecture
- support content reuse across product offerings, and at a more granular level
- migrate unstructured and semi-structured content formats to XML

The Project ... in other words:

INPUT



DESIRED OUTPUT

```
<topic class="TOPIC-Section" id="topic_3">
  <title class="Section">What are high-risk situations with firearms?</title>
  <body>
    <p>Awareness of situations that can increase your child's risk of exposure to unintentional firearms injuries may help prevent them. High-risk situations include the following:</p>
    <ul>
      <li>
        <p>
          <b>Accessible firearms in the home.</b>Parents often underestimate their child's ability to gain access to the firearm in the house, or even the child's ability to pull the trigger.</p>
      </li>
      <li>
        <p>
          <b>Accessible firearms at another home.</b>Even if your home does not contain a firearm, your neighbor, friend, or relative may have an accessible firearm in the house.</p>
      </li>
      <li>
        <p>
          <b>Adolescent boys.</b>The majority of children killed accidentally with firearms are boys. The American Academy of Pediatrics (AAP) considers homes with adolescent boys at a higher risk for accidental injury and death from firearms. :</p>
        </li>
      </ul>
    </body>
  </topic>
```

Choosing an Architecture



create and maintain content independent of the particular product(s) in which it is currently used



reuse content at a more granular level than article



a standards-based tag suite that would support the existing content library with little to no customization.

How Do We Choose?

EXISTING CONTENT:

- headings (typically 2-3 levels)
- list items (numbered, unnumbered, rarely nested)
- an occasional simple table
- often an image or two
- no footnotes
- no citations (at least not in the content itself)
- no linking within or across articles

FUTURE STATE:

- modular content
- semantic enrichment
- on-demand output to customer-facing applications or systems
- AND still need to support content creation in Word

DECISION:

- desire to reuse content below the article level
- availability of the DITA Open Toolkit and DITA for Publishers toolsets,
- DITA was chosen as the best option.

How Hard Can It Be?!

New-Style Ham and Cheese

(#20905)

Ingredients

24 wonton wrappers
1/4 pound lean, low-sodium smoked ham
4-ounce log of fresh goat cheese, such as Montrachet
4 tablespoons spicy fruit chutney

Directions

Lightly spray two mini-muffin pans with cooking spray. Fit the wontons into the muffling the edges to make a well for the filling. Chop the ham and divide among cups. Divide cheese among the cups. Top each cup with about a half teaspoon of at 375 degrees for seven to 10 minutes; edges should be lightly browned and fill with fresh fruit on skewers.

Makes 24 ham and cheese cups.

Nutrition Facts

Each cup contains about 69 calories, 2 grams protein, 1 gram fat, 4 mg cholesterol, 0 grams fiber, and 112 mg sodium.

Recipe: Autumn Vegetable Spaghetti

(#35324)

This lively vegetarian pasta dish contains no added fat or oil, is low in calories, and is good hot or cold.

2 C small yellow onions, cut in eighths
2 C chopped, peeled, fresh, ripe tomatoes (about 1 lb)
2 C thinly sliced yellow and green squash (about 1 lb)
1 1/2 C cut, fresh green beans (about 1/2 lb)
1/2 C water
2 T minced, fresh parsley
1 clove garlic, minced
1/4 t chili powder
1/4 t salt
to taste black pepper
1 can (6 oz) tomato paste
1 lb uncooked spaghetti
1/2 C grated Parmesan cheese

1. Combine first 10 ingredients in large saucepan; cook for 10 minutes, until tomato paste. Cover and cook gently, 15 minutes, stirring occasionally until are tender.
2. Cook spaghetti in unsalted water according to package directions.
3. Spoon sauce over drained, hot spaghetti and sprinkle Parmesan cheese.

Yield: 9 servings

Serving size: 1 cup spaghetti and 1/2 cup sauce with vegetables

Each serving provides:

Calories: 279
Total fat: 3 g
Saturated fat: 1 g
Cholesterol: 4 mg
Sodium: 173 mg

Apple-Cranberry Salad Toss

(#26797)

Ingredients

1 head romaine lettuce
1 Granny Smith apple
1/2 cup walnuts, chopped
1/2 cup dried cranberries
1/4 cup red onion, sliced
1/2 cup low-fat balsamic vinaigrette dressing

Directions

1. Wash lettuce and tear into small pieces. Wash, core, and slice apple.
2. Toss lettuce, apple, walnuts, cranberries, and onion in a large bowl.
3. Add dressing and toss to coat.
Serve immediately.

Nutrition Facts

Serves 8

Amount Per Serving
Calories 110

	% Daily Value (DV)*
Total Fat 6 g	9%
Saturated Fat 0.5 g	3%
Trans Fat 0 g	
Cholesterol 0 mg	0%
Sodium 190 mg	8%
Total Carbohydrate 17 g	6%
Dietary Fiber 3 g	13%
Sugars 10 g	
Protein 2 g	
Vitamin A	148%
Vitamin C	8%
Calcium	4%
Iron	6%

*Percent Daily Values are based on a 2,000-calorie diet.

Raspberry Vinaigrette

(#26878)

- 3/4 cup frozen raspberries, thawed
- 1 tbsp. olive oil
- 2 tbsp. cider vinegar
- 1 tsp. sugar
- 1/4 tsp. garlic salt
- 1/4 tsp. oregano
- 1/4 tsp. ground black pepper

Place all vinaigrette ingredients in a blender, blend until smooth. Serve with your favorite salad and enjoy!

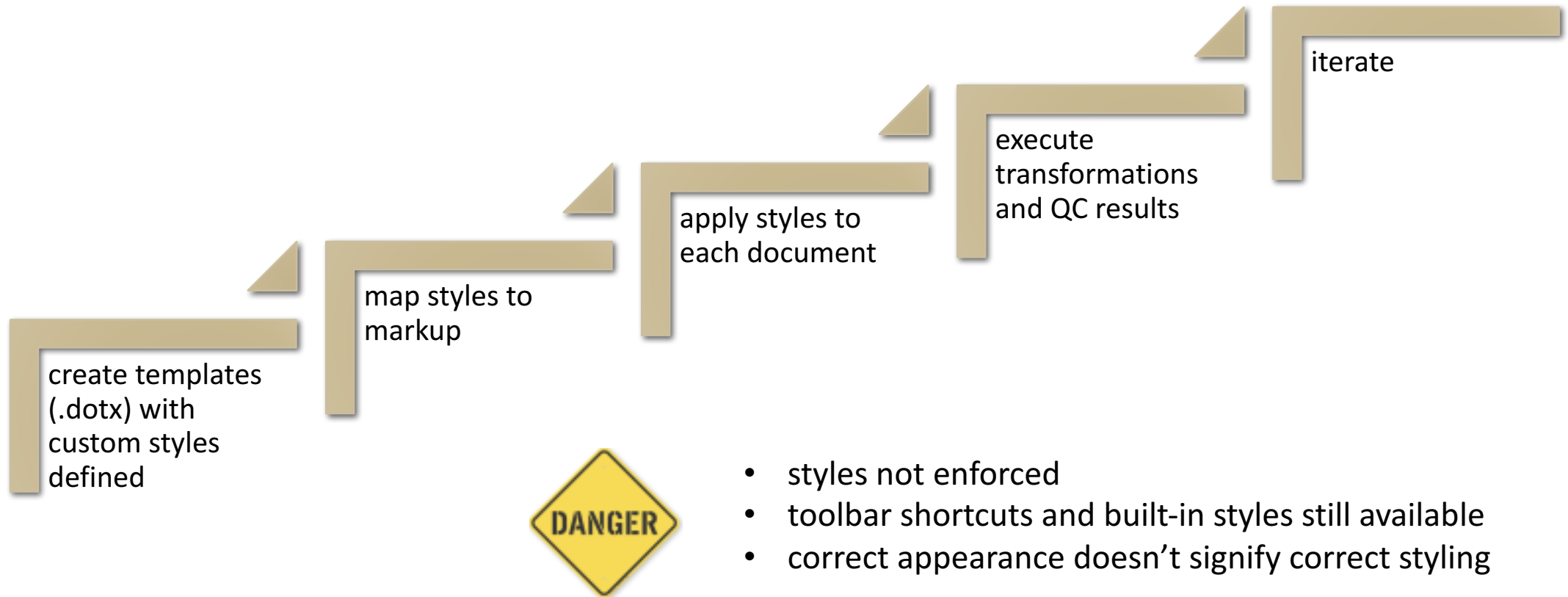
(Note: Vinaigrette can be made ahead and refrigerated.)

Source: cheesemplyatiz.gov

The less structure, consistency, and semantic identification in the source content, the more challenging it is to create structurally and semantically rich content in the resulting output



Some Types of Straw Are Better Than Others: Unstyled Microsoft Word Documents



Some Types of Straw Are Better Than Others: HTML

HTML+-LITE (CUSTOM SCHEMA)

```
<h1>Nonmelanoma Skin Cancer: Statistics</h1>
```

```
<h1></h1>
```

```
<p><b>Nonmelanoma Skin Cancer: Statistics</b></p>
```



- <h#> usage inconsistent
- @data-
- odd structures (b/img, ul/ul)

HTML5

```
<a data-content-type="Article" data-content-slug="autosomal-recessive-inheritance" data-bucket-slug="adult-diseases-and-conditions-v0">Autosomal Recessive Inheritance</a>
```

```
<ul>  
<li><strong>Extra oxygen.</strong> This will make up for the decreased breathing ability of your child's lungs.</li>  
<li><strong>Medicines.</strong> Your child's provider may give him or her the following medicines:</li>  
<ul>  
<li>Bronchodilators to open the airways</li>  
<li>Steroids to reduce inflammation</li>  
</ul>  
<li><strong>IV (intravenous) fluids and nutrition.</strong> This will help your baby and his or her lungs grow.</li>
```

Some Types of Straw Are Better Than Others: Metadata

Review Group	Specialty	Subspecialty
Adolescent Medicine	Family Medicine	Adolescent Medicine
Alcohol/Drugs/Tobacco	Psychiatry and Neurology	Addiction Psychiatry
Allergy/Immunology	Allergy and Immunology	General
Arthritis/Rheumatology	Rheumatology	General

Category	Specialty	Subspecialty
Children	Pediatrics	General
Depression	Mental Health	General
Diabetes	Endocrinology	Diabetes and Metabolism
Domestic Violence	Mental Health	General
Ear, Nose & Throat	Otolaryngology	General

Some Types of Straw Are Better Than Others: JSON

FROM THIS:

Getting the Right Nutrients

Code: CHOLESS_03A

Condition: Diabetes

MLU Type: Click & Learn

Header

URL Code:

Category: Management

Display Banner: Yes

Intro: You know that managing your fat intake is important. At the same time, you need to make sure you consume the right mix of nutrients to stay healthy. When you are planning your meals, think about the foods and drinks you are putting on your plate and in your cup.

Directions: CLICK THROUGH THI ACTIVITY TO LEARN HOW YOU CAN MAKE SURE YOU EAT HEALTHY AND STAY WELL.

Note:

Body

Hotspot Opacity (%): true

myplate_green_mlu5_CL.png

Image Size: big

1. MLU Item (Title): Protein

Item Defaults

Hotspot Top Loc (%): 50

Hotspot Left Loc (%): 55

Hotspot Size: 80

Item Activated

Image Ref: iStock_16140658_mlu5_CL.png

TO THIS:

Getting the Right Nutrients

You know that managing your fat intake is important. At the same time, you need to make sure you consume the right mix of nutrients to stay healthy.

When you are planning your meals, think about the foods and drinks you are putting on your plate and in your cup.

CLICK THROUGH THIS ACTIVITY TO LEARN HOW YOU CAN MAKE SURE YOU EAT HEALTHY AND STAY WELL.



Vegetables

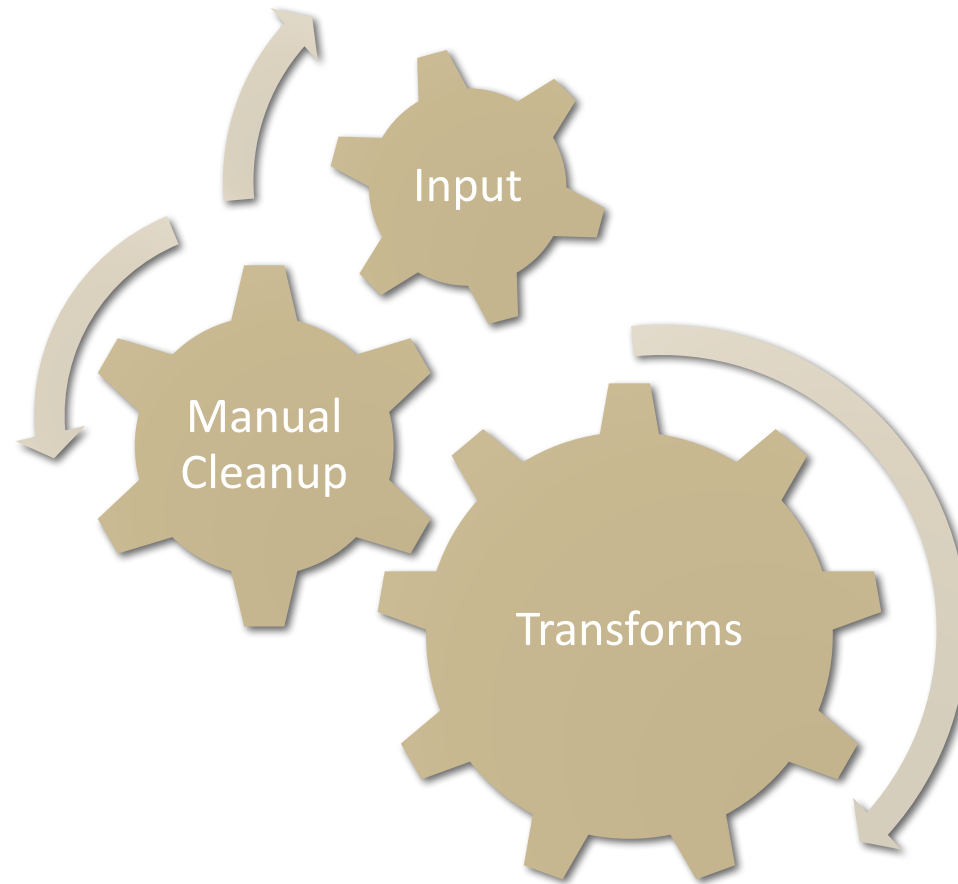
- Vegetables are great sources of vitamins, minerals, fiber and the disease-protecting phytochemicals. They also contain plant stanols and sterols that work much like soluble fiber in lowering cholesterol.
- Choose non-starchy vegetables, which contain lots of fiber over starchy ones.
- Be sure to fill $\frac{1}{2}$ of your plate with them.

FEEDBACK

JSON



Getting to Gold



Is it Really Gold?

24K (100%)	the gold standard. all fragments have been properly identified (i.e. bibliographic entries, intra-document cross-references, footnote references/footnotes, external websites, semantic enrichment, other application-specific information).
18K (75%)	all existing content (elements and attributes) correctly transformed; hierarchical structures accurately reflected in the markup.
10K (41.7%)	minimum acceptable level. XML instance must be valid, and no displayable content unintentionally dropped. markup may not be correct.

In Conclusion ...

REMEMBER:

- Up-transformations are hard.
- Just because there's a DTD or schema in place, doesn't mean it's suitable for up-translation.
- Just because the content looks the same, doesn't mean it's styled/tagged the same.

ADVICE:

- Gather as many samples as you can.
- Each exception becomes the new norm.
- XSpec is your friend.

