

Regnet – Regulatory Information Management Framework for Enhancing Access to Regulations, Compliance Checking and Assistance

Presented by

Kincho H. Law, Professor of Civil and Env. Engineering, Stanford University
Shawn Kerrigan, Ph.D. Student, Civil and Env. Engineering, Stanford University
Bill Labiosa, Ph.D. Student, Civil and Env. Engineering, Stanford University
Charles Heenan, Student, School of Law, Stanford University

Presented at

National Compliance Assistance Providers Forum
December 3-6, 2002
San Antonio, Texas

Cosponsored by

U.S. Environmental Protection Agency
and Texas Commission on Environmental Quality

Disclaimer

U.S. Environmental Protection Agency (U.S. EPA) solicited from compliance assistance providers presentations aimed at sharing expertise, building skills and networking. The following presentation is intended as a resource for providing assistance regarding compliance with environmental regulations. U.S. EPA neither endorses nor assumes responsibility for the accuracy and completeness of non-EPA materials contained herein. EPA does not necessarily endorse the policies or views of the presenters, and does not endorse the purchase or sale of any commercial services or products mentioned in this presentation.

REGNET: An Infrastructure for Regulatory Information Management and Compliance Assistance

Kincho H. Law
Prof., Civil and Env. Engr.

Jim Leckie
Prof., Civil and Env. Engr.

Barton Thompson
Prof., School of Law

Gio Wiederhold
Prof., Computer Science

Shawn Kerrigan

Bill Labiosa

Gloria Lau

Haoyi Wang

Jie Wang

Civil and Env. Engr.

Pooja Trivedi

Li Zhang

Liang Zhou

(former students)

Computer Science

Charles Heenan

Researcher, Law Student

Stanford University, Stanford, CA 94305



The **Public** and **Scientific** Problem

- **Regulations are established to protect the public**
- **Regulations constrain businesses' actions**
- **Many organizations participate to set and use regulations**
- **Interpretation of regulations is costly and inconsistent**

- **Regulations are voluminous, often incomplete, sometimes conflicting**
- **Regulations are written in natural language**
- **The objects and interests being regulated are often encoded**
- **Many sources of supportive documents – interpretative documents, guidelines, etc..**

Motivation

The complexity, diversity, and volume of federal and state regulations:

- Require considerable expertise to understand
- Increase the risk of companies failing to comply with environmental regulations
- Hinder public understanding of the government

How would IT help

- to make “applicable” regulations easily accessible?
- to assist parties involved in regulation compliance?

REGNET Project

(sponsored by Digital Government Program, National Science Foundation)

Objective

To enhance regulation management, access and the regulatory compliance process through the use of information technology

Application Focus

Environmental Regulations:

- *Federal CFR Title 40: Protection of Environment*
- *40 CFR 279: Standards For The Management Of Used Oil*
- *40 CFR 141: National Primary Drinking Water Regulations*

Others : State Regulations

Illinois Title 35: Environmental Protection

New York Title 6: Environmental Conservation Rules and Regulations

REGNET Research Goals

- **Research questions**
 - **What is an appropriate model for a information management system for compliance assistance?**
 - **How to build such a system**
 - **How to deal with the conflicting objectives?**
- **Research goal**
 - **Developing information management frameworks that can facilitate public access to regulations, improve the efficiency of regulation compliance and facilitate the compliance process.**



Research Tasks

- **Repositories:** Infrastructure for online repository of regulations and translating texts into processable form and facilitate access
- **Access Tools:** Access of the regulation text and related information
- **Ontology Development:** Formalize terms and meanings to help development of logical rules about relationships in the regulations and among the different regulations
- **Integrated Access:** Retrieval of regulations based on the content or relationships between the regulations
- **Analysis Tools:** To validate and improve the quality of the ontology and to check the content of regulations within a domain or across different domains of federal, state and local regulations.
- **Compliance Checking Assistance:** To develop the means to interface the regulations with usage.

Current Features of REGNET

- Developed a hierarchical structure to organize documents
- Regulations structured using XML
- Methodology developed to classify documents for different purposes with a singled repository
- Included a variety of documents – guidelines, background information, etc...
- Developed a logic-based compliance assistance methodology and prototype
- Integrating compliance assistance with document repository

REGNET

About

People

Publications

Presentations

Search 40 CFR

Resources

Links



CLASSIFICATION STRUCTURES

DISCLAIMER: These classification structures, and the classification rules they contain, are works-in-progress as part of an academic research effort at Stanford University. Further, the text version of 40 CFR that is searchable on this site was downloaded in January of 2001 for research purposes only.

****THIS SITE DOES NOT CONTAIN
THE CURRENT VERSION OF 40 CFR.

For counsel regarding environmental compliance, consult an attorney.

[40 CFR INDEX OF TERMS](#)

(An alphabetical index of terms in 40 CFR)

[Index to the National Primary Drinking Water Regulations](#)

(An alphabetical index of terms in 40 CFR 141)

[The Periodic Table of Elements](#)

(A classification of terms in 40 CFR according to the Periodic Table of the Elements)

[On the Topics of Regulation, Pollution, and Waste](#)

(A classification of terms in 40 CFR on the topics of regulation, pollution, and waste)

[EPA List of Extremely Hazardous Substances](#)

(A classification of terms in 40 CFR according to the EPA's list of Extremely Hazardous Substances)



Internet

RAS

Regulation
Assistance
System

[main](#)
[settings](#)

Compliance check in progress..

(from provision 40.cfr.279.12.b)

Question: Is the used oil used as a dust suppressant?

Please select an answer:

- yes
- no
- I don't know

Version 0.5

§ 279.12 Prohibitions.

(a) Surface impoundment prohibition. **Used oil** shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under parts 264 or 265 of this chapter. [40.cfr.264](#) [40.cfr.265](#)

(b) Use as a dust suppressant. The use of used oil as a dust suppressant is prohibited, except when such activity takes place in one of the states listed in § 279.82(c). [40.cfr.279.82.c](#)

(c) Burning in particular units. Offspecification **used oil** fuel may be burned for energy recovery in only the following devices:

- (1) Industrial furnaces identified in § 260.10 of this chapter
- (2) Boilers, as defined in § 260.10 of this chapter
- (i) Industrial boilers located on the site of a facility where substances are transformed into new products, including the component parts of products, and as a result of such use is contaminated by physical or chemical impurities.
- (ii) Utility boilers used to produce electric power, steam, heated or cooled air, or other gases or fluids for sale; or
- (iii) **Used oil**-fired space heaters provided that the burner meets the provisions of § 279.23. [40.cfr.279.23](#)

Used oil means any oil that has been refined from crude oil, or any synthetic oil, that has been used and as a result of such use is contaminated by physical or chemical impurities.

[40.cfr.260.10](#)

ss where substances are transformed into new
esses;

(from provisi

Question: Do
used oil gene

Please select
 yes
 no
 I don't

[main](#)
[settings](#)

Version 0.5

§ 279.23 On-

**(a) The heat
oil generator**

(b) The heat

**(c) The comb
May 3, 1993**

Suggested st

[ambient air o](#)

UsedOilSu

TOP > G-H > he

Related Terms

- ambient air
- do-it-yourself
- fuel specificati

DOCUMENTS

Showing 1 to 10 o

1. <http://171.64.5>
(Cited as: 1995 v
August 25, 1991)
2. <http://171.64.5>
(Cited as: 1986 v
TRANSMITTAL
3. <http://171.64.5>
(Cited as: 1991 v
29, 1991 Mr. R
4. <http://171.64.5>
(Cited as: 1988 v
September 22,
5. **USED OIL** (C
United States
Module Introd
6. <http://171.64.5>
(Cited as:

Environmental Protection Agency
Office of Solid Waste and Emergency Response

OSWER Directive

*1 August 25, 1995

Norman H. Nosenchuck, P.E.
Director
Division of Hazardous Substances Regulation
New York State Department of Environmental Conservation
50 Wolf Road
Albany, New York 12233-7250

Dear Mr. Nosenchuck:

This letter is in response to your letter of May 24, 1995 concerning burning used oil in space heaters found in the Federal Used Oil Manager

In your letter you request EPA's interpretation of the regulatory exemption for burning used oil in space heaters in three specific used oil collection scenarios all involve a county highway maintenance garage accepting additional used oil to burn in a on-site space heater. The used oil is collected by the County maintenance facilities, County-run Do-it-yourself (DIY) collection facilities, and from businesses. Specifically, your letter requests clarification of whether the County can burn these three sources of used oil in their space heaters for the space heater exemption in 40 CFR 279.23(a)(1)).

Presentation Outline

- Short prototype demonstrations
- Methodology to establish concepts and classification structures in the regulatory documents
- “Logic-based” compliance assistance system

Purpose : Feedbacks and Suggestions

Document Repository and Access

Demonstration Session I

- SemioTagger: Concept categorization approach
- Simple Index for Drinking Water Regulations, 40 CFR 141
- Simple Index for Drinking Water Background documents, USEPA Office of Ground Water and Drinking Water's web-available material

Regulation Assistance System

Demonstration Session II

- Display regulations with meta-data
- Compliance example
- Non-compliance example

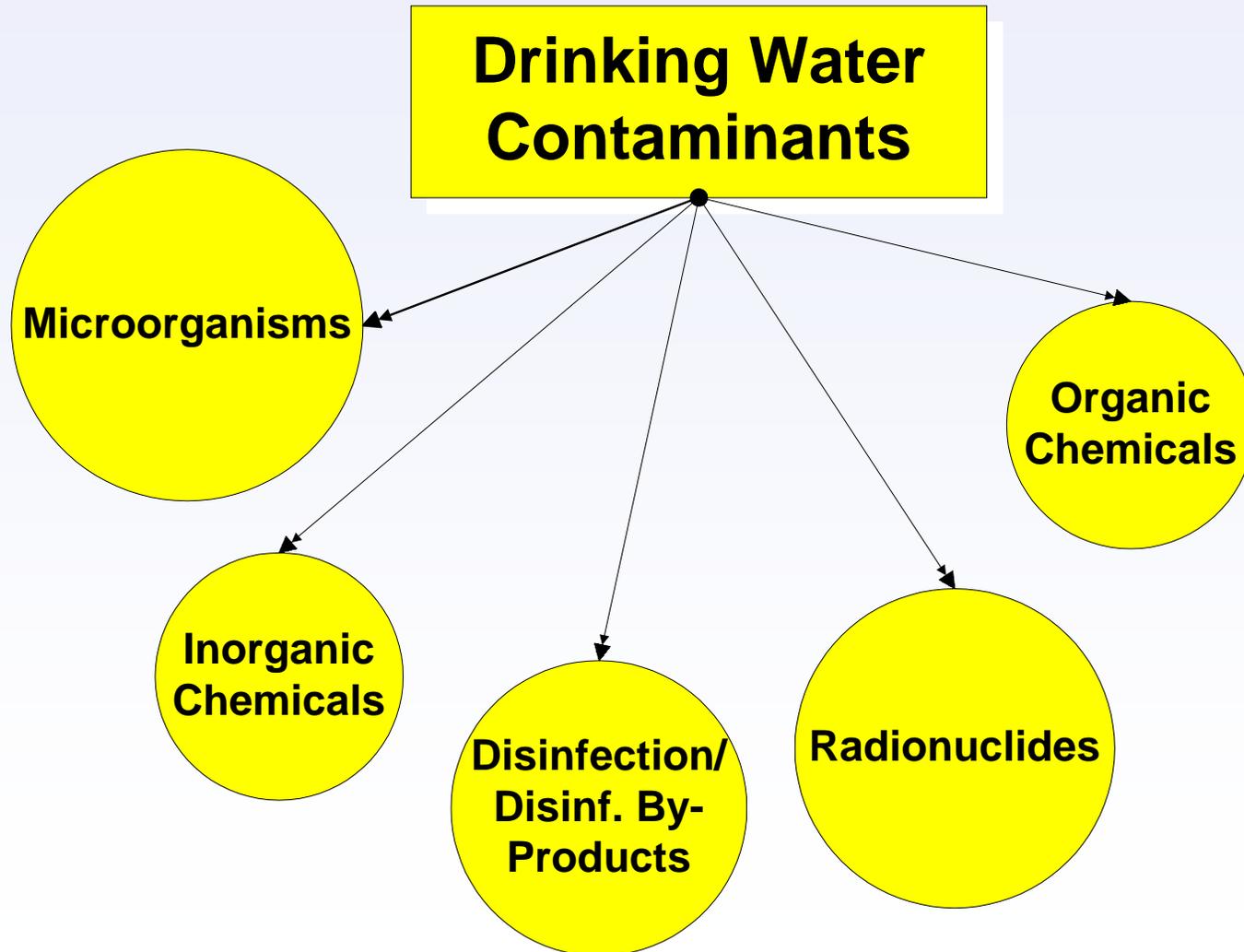
Document Repository and Access: Concept Categorization Approach

Bill Labiosa

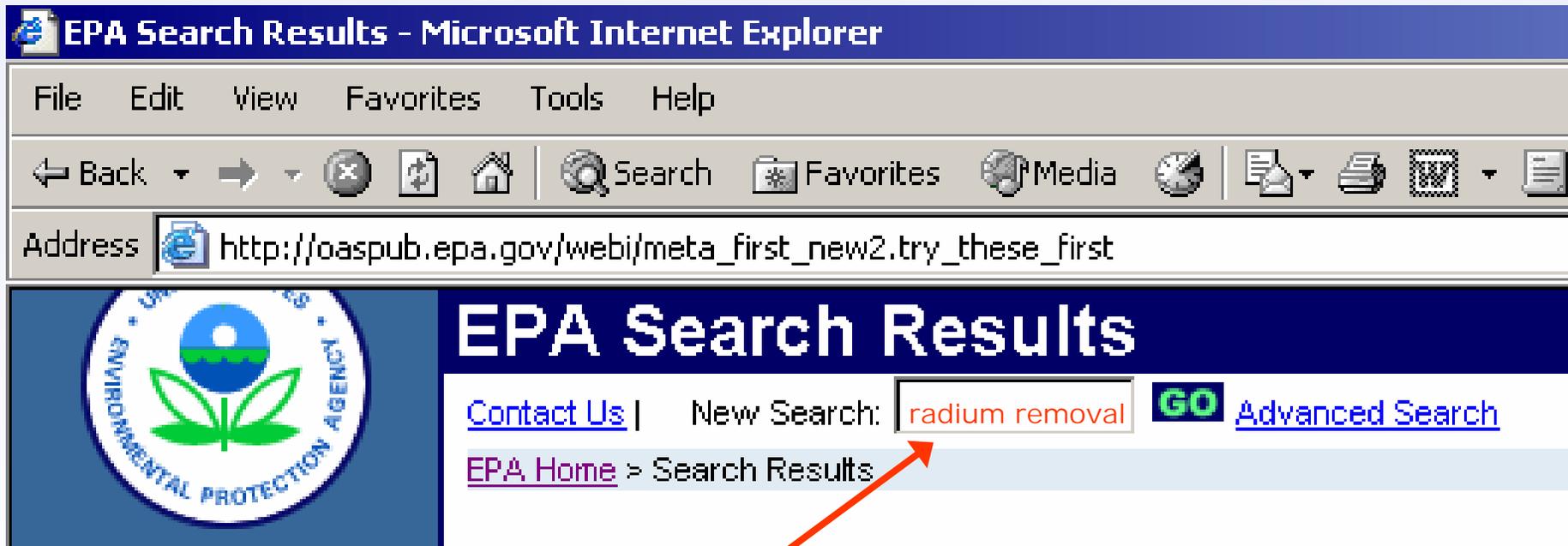
Engineering Informatics Group

Stanford University

Example Taxonomy: Drinking Water Contaminants



Current Approach: Using “EPA Search”



The screenshot shows a Microsoft Internet Explorer browser window titled "EPA Search Results - Microsoft Internet Explorer". The address bar displays the URL: http://oaspub.epa.gov/webi/meta_first_new2.try_these_first. The main content area features the EPA logo on the left and the heading "EPA Search Results" in large white text on a dark blue background. Below the heading, there is a search bar containing the text "radium removal" in red. To the right of the search bar is a green "GO" button and a link for "Advanced Search". Below the search bar, there are links for "Contact Us", "EPA Home", and "Search Results". A red arrow points from the text "Search Term: 'radium removal'" below the screenshot to the search bar.

Search Term:
"radium removal"



U.S. Environmental Protection Agency

EPA Search Results

Contact Us | New Search:
EPA Home > Search Results

Small System Compliance Te

The 1996 SDWA contains over 7

URL: <http://www.epa.gov/OGWV>

Rank

- 1 [United States Response to the Commission's Ninth Biennial Report on Great Lakes Water Quality](#)
The Commission's Ninth Biennial Report on Great Lakes Water Quality identifies 10 persistent toxics from the Commission's 1996 SDWA contains over 7000 chemicals.
URL: <http://www.epa.gov/glwq>
- 2 [Small System Compliance](#)
The 1996 SDWA contains over 7000 chemicals.
URL: <http://www.epa.gov/OGWV>
- 3 [Meeting Summary: Radionuclides in the City of Jefferson Davis](#)
This is a summary of the radionuclides in the City of Jefferson Davis.
URL: <http://www.epa.gov/s>
- 4 [United States Response to the Commission's Ninth Biennial Report on Great Lakes Water Quality](#)
The Commission's Ninth Biennial Report on Great Lakes Water Quality identifies 10 persistent toxics from the Commission's 1996 SDWA contains over 7000 chemicals.
URL: <http://www.epa.gov/glwq>
- 5 [EPA Publication Titles Sorted Alphabetically](#)
The following list is available for each publication title.
URL: <http://www.epa.gov/c>



minants Regulated Before 1996

Format

- | | |
|---|--|
| United States Response to the Commission's Ninth Biennial Report on Great Lakes Water Quality | |
| and virtually eliminating, | |
| Small System Compliance | |
| Meeting Summary: Radionuclides in the City of Jefferson Davis | |
| United States Response to the Commission's Ninth Biennial Report on Great Lakes Water Quality | |
| and virtually eliminating, | |
| or alphanumericly by EPA | |

- EPA Newsroom
- Browse EPA
- Laws, Regulations, and Dockets
- Where You Live
- Information Sources
- Educational Resources
- About EPA
- Programs
- Business Opportunities
- Jobs
- Recursos en Espanol
- Explorers' Club For Kids!

United States
Environmental Protection Agency Office of Water EPA 815-R-98-002
4607 September 1998



Small System Compliance Technology List for the Non-Microbial Contaminants Regulated Before 1996

Full 112 page Document is returned ...

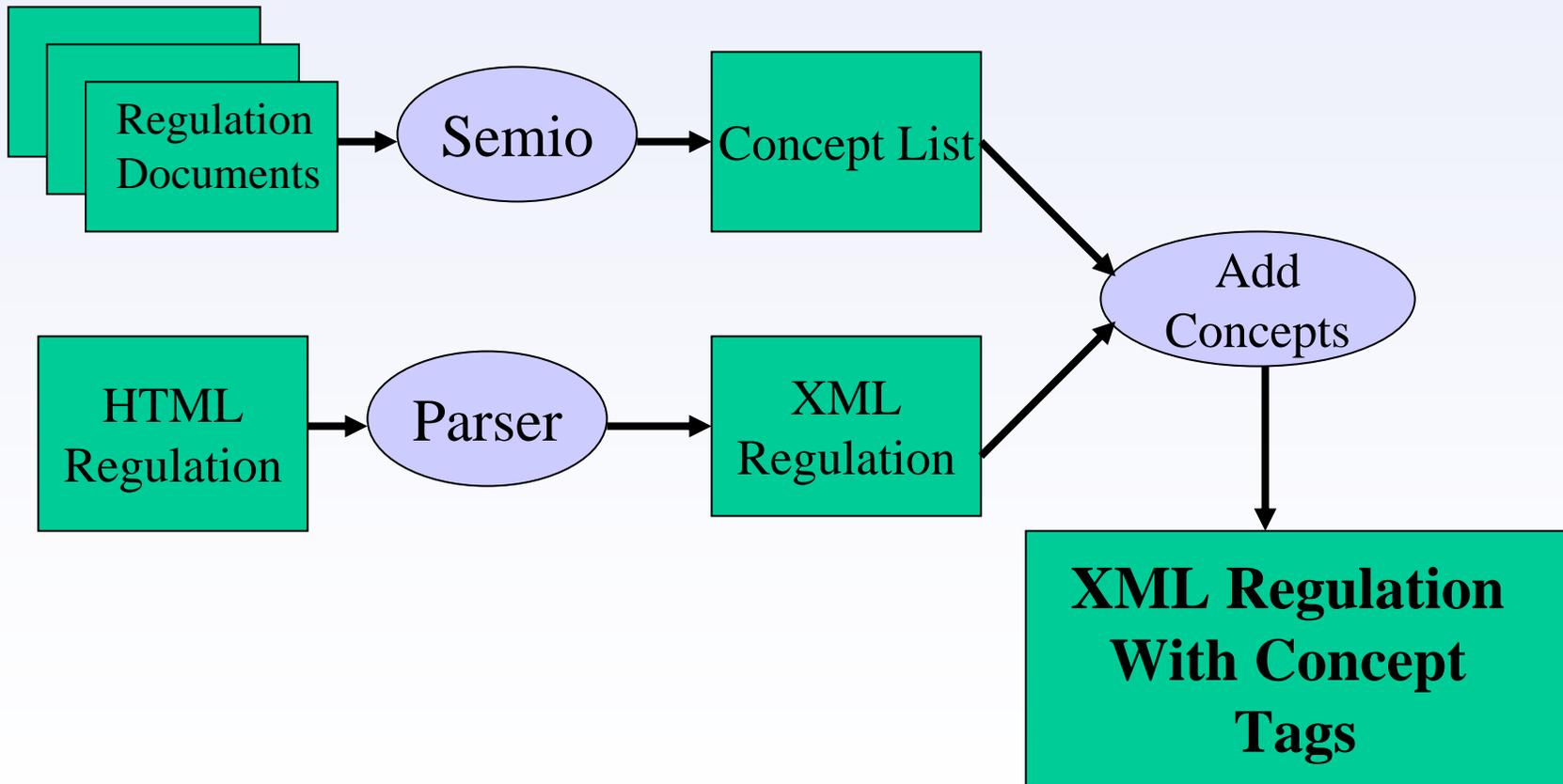
Concept Categorization

- noun phrase extraction
- noun phrase co-occurrence cycles
- hierarchy creation
- document tagging
- information retrieval interface

Document Repository and Access Demonstration Session III

- Another example: Drinking Water Contaminant Hierarchical Structure
- A bit on how the Concept Categorization approach works
- A bit on our document repository and access

Document Repository and Access



Regulatory Compliance Assistance

Shawn Kerrigan
Engineering Informatics Group
Stanford University

Background

- Current state of compliance checking:
 - Paper-based process
 - Locating and interpreting the relevant regulations is complex, even with the help of supplementary information
 - Small companies have difficulty conducting compliance checks due to lack of resources and knowledge
- Vision for future:
 - Up-to-date regulations and compliance-checking assistance procedures available online
 - Improved regulation and compliance-requirement transparency through clear presentation and linking

Research Questions

- How can we make the information and rules more accessible?
- How can we represent the information and rules in environmental regulations in a computer interpretable format?
- How can we structure this information to assist with regulation compliance checking?

General Approach

- Information Integration
- Formalization of meaning and relationships
- Regulation-centric
- Tie the information to the appropriate portion of the regulation

Regulation Assistance System (RAS)

- Provides a unifying web interface for the regulation documents and meta-data
- Demonstrates the usefulness of XML structured regulation documents with meta-data
- Works with a logic-based compliance-checking assistance system to demonstrate web-based regulation services

Regulation Parsing

- Need to transform plain text/PDF regulations into XML
- Can structure the XML to represent the hierarchical structure of the regulation

HTML to XML Regulation Parsing

The screenshot shows a web browser window displaying the Electronic Code of Federal Regulations (e-CFR) for Title 40, Chapter I, Part 279. The page header includes "Electronic Code of Federal Regulations" and "e-CFR™". Below the header, it states "THIS DATA CURRENT AS OF THE FEDERAL REGISTER DATED OCTOBER 25, 2002" and "40 CFR - CHAPTER I - PART 279". A "View Part" link is visible. The main content area shows the beginning of § 279.12 Prohibitions, with two subsections: (a) Surface impoundment prohibition and (b) Use as a dust suppressant. The browser's address bar shows "Internet".

Title 40
CHAPTER I
PART 279

- [279.1](#)
- [279.10](#)
- [279.11](#)
- [279.12](#)
- [279.20](#)
- [279.21](#)
- [279.22](#)
- [279.23](#)
- [279.24](#)
- [279.30](#)
- [279.31](#)
- [279.32](#)
- [279.40](#)

Electronic Code of Federal Regulations
e-CFR™

THIS DATA CURRENT AS OF THE FEDERAL REGISTER DATED
OCTOBER 25, 2002

40 CFR - CHAPTER I - PART 279

[View Part](#)

§ 279.12 **Prohibitions.**

(a) *Surface impoundment prohibition.* Used oil shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under parts 264 or 265 of this chapter.

(b) *Use as a dust suppressant.* The use of used oil as a dust suppressant is prohibited, except when such activity takes place in one of the states



XML
Structured
Document

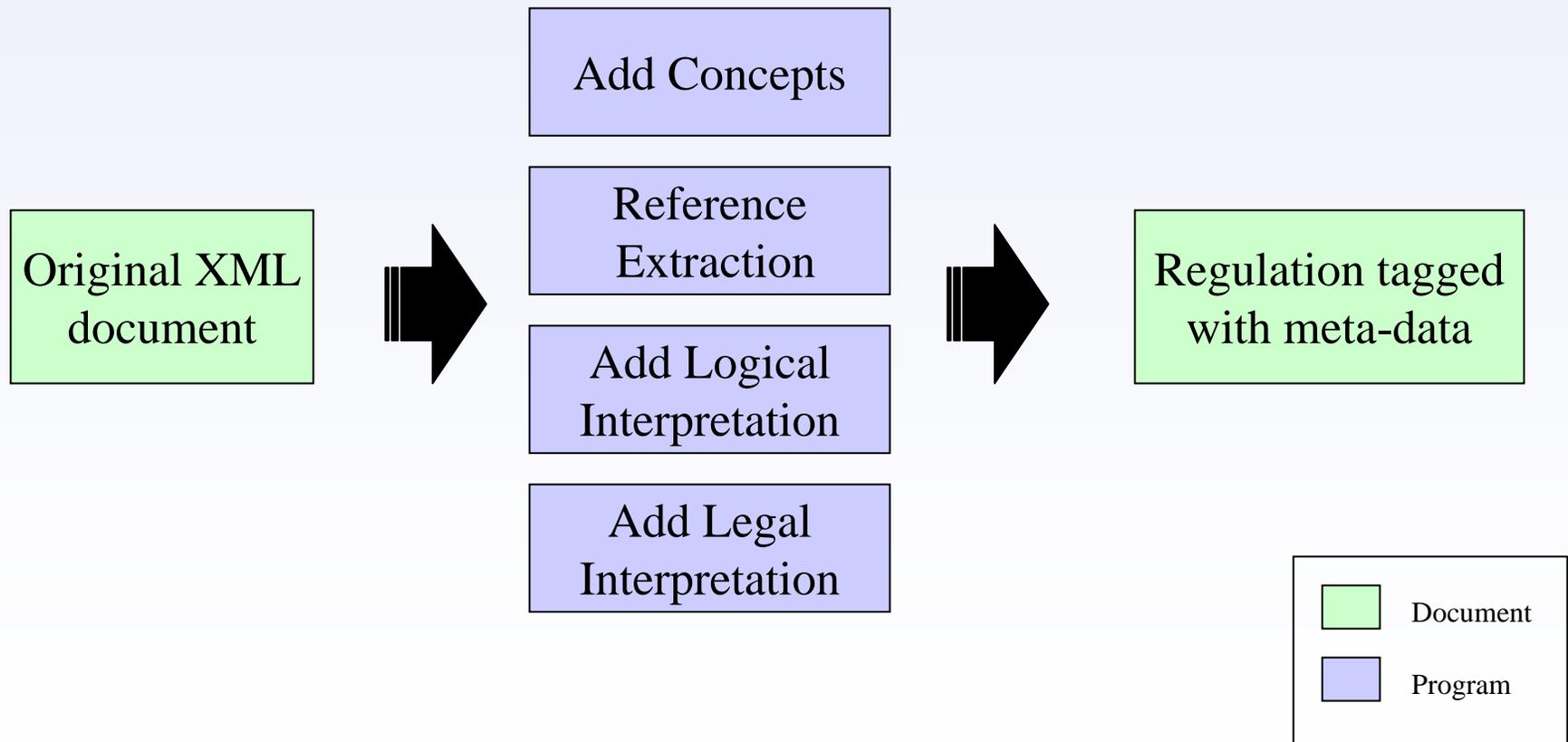
Regulation Parsing

§ 279.12 Prohibitions.

(a) Surface impoundment prohibition. Used oil shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under parts 264 or 265 of this chapter.

```
<regElement id = "40.cfr.279.12" title = "Prohibitions">  
  < regElement id = "40.cfr.279.12.a" title = "Surface Impoundment prohibition">  
    <regText>  
  
      Used oil shall not be managed in surface impoundments or waste piles unless  
      the units are subject to regulation under parts 264 or 265 of this chapter.  
  
    </regText>  
  </regElement>  
</regElement>
```

Adding Meta-Data to Regulations



Parsing References

PART 279—Standards For The Management Of Used Oil

Subpart B – Applicability

...

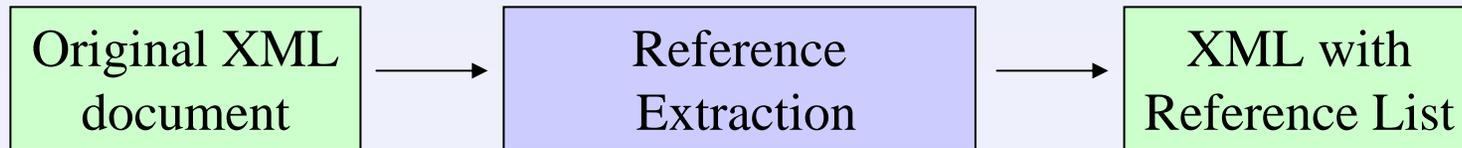
§ 279.12 Prohibitions.

- (a) Surface impoundment prohibition. Used oil shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under **parts 264 or 265 of this chapter**.
- (b) Use as a dust suppressant. The use of used oil as a dust suppressant is prohibited, except when such activity takes place in one of the states listed in **§ 279.82(c)**.
- (c) Burning in particular units. Off-specification used oil fuel may be burned for energy recovery in only the following devices:
- (1) Industrial furnaces identified in **§ 260.10 of this chapter**;
 - (2) Boilers, as defined in **§ 260.10 of this chapter**, that are identified as follows:
 - (i) Industrial boilers located on the site of a facility engaged in a manufacturing process where substances are transformed into new products, including the component parts of products, by mechanical or chemical processes;
 - (ii) Utility boilers used to produce electric power, steam, heated or cooled air, or other gases or fluids for sale; or
 - (iii) Used oil-fired space heaters provided that the burner meets the provisions of **§ 279.23**.
 - (3) Hazardous waste incinerators subject to regulation under **subpart O of parts 264 or 265 of this chapter**.

§ 262.11 Used Oil Specification.

.....

Parsing References



Before:

```
<regText>
(a) Surface impoundment prohibition. Used oil shall not be managed in surface
impoundments or waste piles unless the units are subject to regulation under parts 264 or
265 of this chapter.
</regText>
```

After:

```
<regText>
(a) Surface impoundment prohibition. Used oil shall not be managed in surface
impoundments or waste piles unless the units are subject to regulation under parts 264 or
265 of this chapter.
</regText>
<reference id = "ref.40.cfr.264" />
<reference id = "ref.40.cfr.265" />
```

What is a “Concept”?

- Examples:
 - emission requirement
 - leaked hazardous substance
 - disposal of solvents
 - principal hazardous constituent
- Why are they useful?
 - identify similar regulations even when they do not reference each other
 - provide a “context” for the regulation provision

Regnet Taxonomy

Taxonomy - 40CFRv4 The Topic of Regulation (Powered by Semio) - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Copy Paste

Address <http://eil.stanford.edu/regnet/> Go Links

40CFRv4 The Topic of Regulation (Powered by Semio)

SEARCH HELP

STANFORD UNIVERSITY

TOP > On the Topic of Pollution and Waste (523)

SUBCATEGORIES

| | |
|--|--|
| ▶ Pollution (365) | ▶ Liquid Waste (158) |
| ▶ Solid Waste (123) | ▶ Emissions (354) |
| ▶ Exhaust (160) | ▶ Hazardous Waste (99) |
| ▶ Waste Management (79) | ▶ Contamination and Decontamination (93) |
| ▶ Incineration (98) | ▶ Dumping (30) |
| ▶ Waste Disposal (67) | ▶ Industrial Waste (24) |
| ▶ Leaching and Leachate (21) | ▶ Municipal Waste (30) |

CONCEPTS

| | |
|--|---|
| ◆ accumulated waste [1] | ◆ aftertreatment device [4] |
| ◆ aftertreatment system [1] | ◆ agricultural waste [9] |
| ◆ alkaline waste [1] | ◆ allowable waste feed [2] |
| ◆ animal waste [8] | ◆ annual waste quantity [1] |
| ◆ asbestos waste container [1] | ◆ asbestos waste shipment [1] |
| ◆ asbestos waste [2] | ◆ asbestoscontaining waste material [1] |
| ◆ ash waste [1] | ◆ bagging of waste [1] |

Internet

Tagging with Concepts

```
<regText>
```

Used oil shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under parts 264 or 265 of this chapter.

```
</regText>
```

```
<regText>
```

Used oil shall not be managed in surface impoundments or waste piles unless the units are subject to regulation under parts 264 or 265 of this chapter.

```
</regText>
```

```
<concept name = “waste pile” />
```

```
<concept name = “surface impoundment” />
```

XML Embedded Logic

Rule logic represents the rules specified by the regulation:

40.CFR.279.12.b – Use as a dust suppressant:
“The use of used oil as a dust suppressant is prohibited...”

```
<logic_sentence>  
  all _o (usedOil(_o) -> -(dustSuppressant(_o))).  
</logic_sentence>
```

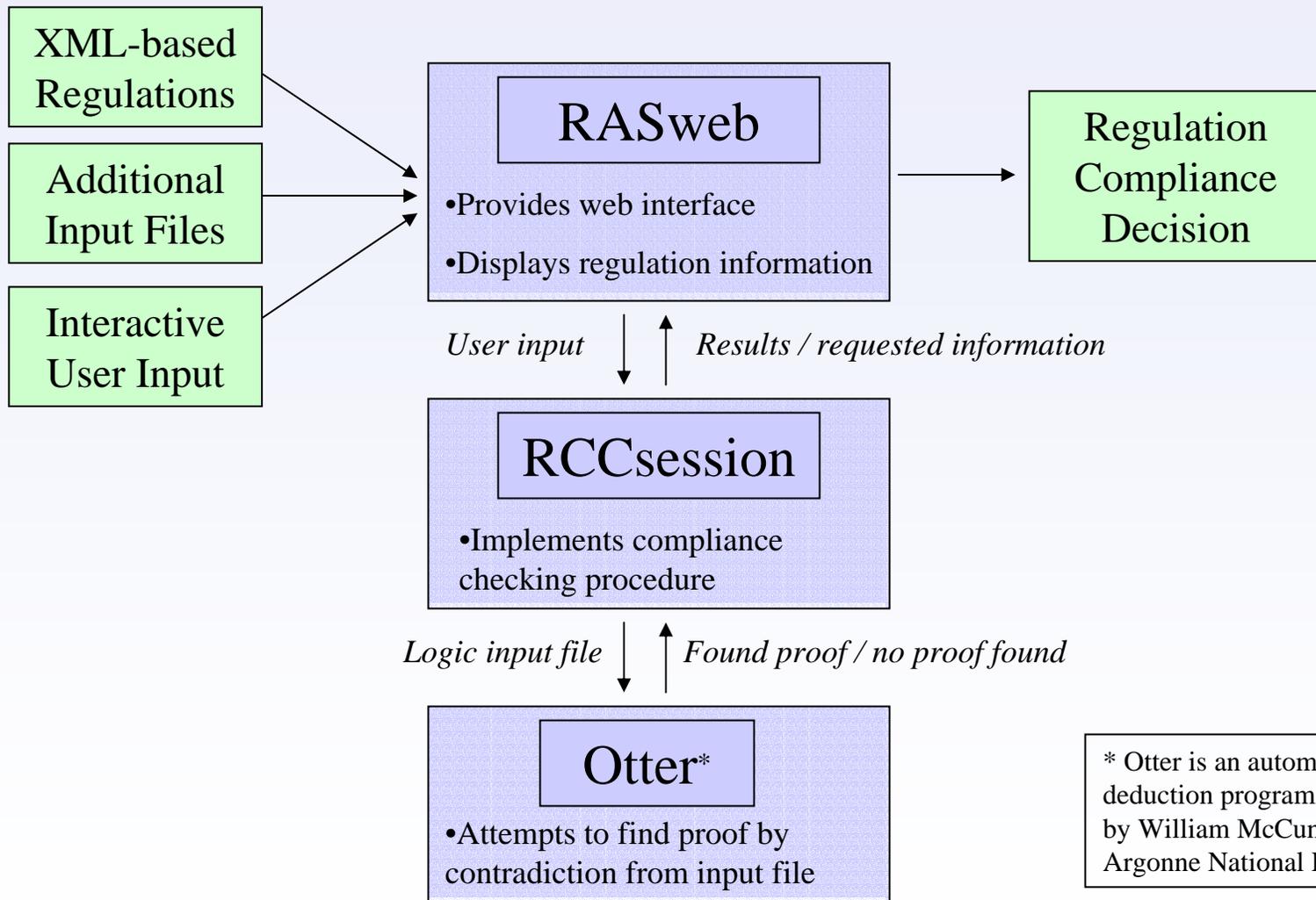
Option elements define the user interface:

```
<logic_option>  
  <question>  
    Is the used-oil used as a dust suppressant?  
  </question>  
  <logic_opt answer = "yes">  
    (usedOil(oil1) & dust_suppressant(oil1)).  
  </logic_opt>  
  <logic_opt answer = "no">  
    (usedOil(oil1) & -(dust_suppressant(oil1))).  
  </logic_opt>  
</logic_option>
```

Control statements specify processing instructions for compliance-checking:

```
<control>  
  <goto target = "ref.40.cfr.279.65" />  
  <switchTo target = "ref.40.cfr.279.73" />  
</control>  
  
<control>  
  <end target = "ref.40.cfr.279.12" />  
</control>
```

RAS System Structure



Demonstration Session IV

- Use of control elements
- Use of “I don’t know” to check multiple paths

Summary

- Can decompose regulations into a structured XML document
- Adding rich meta-data about regulations enables more sophisticated interaction with the documents
- Automated assistance with environmental compliance-checking may be possible

Thank You!

Questions?

Discussion Questions

- How will such a system be useful?
- What are examples of how you could use such a system?
- What would make the system more useful?
- Do you have suggestions for people/fields we should contact that might be interested in what we are doing?
- How are the problems addressed currently dealt with?
- What are some existing technologies we should investigate?
- What are recommendations for issues we should address?
- What might be complementary tools to develop next?