

A Relatedness Analysis Tool for Comparing Drafted Regulations and Associated Public Comments

Gloria T. Lau, Kincho H. Law and Gio Wiederhold
{glau, law}@stanford.edu, gio@cs.stanford.edu
Stanford University, Stanford, CA 94305

Abstract

The process of e-rulemaking with participation from the public involves a non-trivial task of sorting through and organizing a massive volume of electronically submitted comments. This research proposes to make use of available Information and Communication Technology (ICT) to help describe the relationship of public comments to policy drafts and deliberations. Based on previous work on regulatory management and comparisons, a relatedness analysis tool has been prototyped and applied to compare drafted regulations with the associated public comments. An example using a drafted regulation on rights-of-way access and the comments received by the Access Board is employed to illustrate the prototyped analysis tool. The drafted regulation and public comments are first converted into XML format, which is well suited for handling semi-structured data such as legal documents. Feature extraction is performed to identify important domain knowledge. The resulting XML versions of the drafted regulation and public comments are compared using not only a traditional term match but also a combination of feature matches, and not only content comparison but also structural analysis. This comparison framework helps review of comments with respect to provisions in the draft. Examples of results are shown to illustrate the use and limitations of ICT to support policy making.

1. Introduction

The making of government regulations represents an important communication between the government and citizens. During the process of rulemaking, government agencies are required to inform and to invite the public to review proposed rules. Interested and affected citizens then submit comments accordingly. E-rulemaking redefines this process of rule drafting and commenting to involve the public more effectively in the making of regulations. Electronic media, such as the Internet, provide a better environment for the public to comment on proposed rules and regulations. For instance, email has become a popular communication channel for comment submission. Based on the review of public comments received in part from the electronic agora, government agencies revise the proposed rules.

The process of e-rulemaking generates a large number of public comments that need to be reviewed and analyzed along with the drafted rules. With the increased connectivity provided by the Internet, government agencies are required to handle a growing amount of data from the public. For example, the *Federal Register* documented a recent case where the Alcohol and Tobacco Tax and Trade Bureau received over 14,000 comments in 7 months, the majority of which were emails, on a flavored malt beverages proposal. The call for public comments included the following statement: “All comments posted on our Web site will show the name of the commenter but will not show street addresses, telephone numbers, or e-mail addresses.” However, due to the “unusually large number of comments received,” the Bureau later announced that it was difficult to remove all street addresses, telephone numbers and email addresses “in a timely manner.” Instead, concerned individuals were asked to submit a request for removal of address information as opposed to the original statement posted in the call for comments. As such, an “effortless” electronic comment submission process turned into a huge data processing problem for this government agency.

As noted by Coglianese, Information and Communication Technology (ICT) can potentially help streamline the development of regulatory policy in several new directions. One suggestion is to integrate rules with other laws, such as using ICT to

“link all the traces of a rule’s history, both back to the underlying statute and back to past or related rules, facilitating improved understanding of legal requirements.”¹ Previous work has shown that such an application of ICT is indeed possible. A framework for comparisons between regulatory documents from multiple sources has been developed, with successful examples of related provisions automatically linked.² Based on the developed framework, this paper demonstrates another ICT application to support rulemaking.

This paper first discusses the technology behind a demonstrative relatedness analysis system that compares government regulations from different sources. We then show the application of this system to e-rulemaking to compare drafted regulations with associated public comments. We demonstrate the automated sorting of public comments with respect to drafted rules with which interested users can review related rules and comments. Rule makers also can use this tool to locate relevant public comments among thousands received. Several examples of results obtained using this tool will be shown to illustrate potential improvements as well as limitations of the use of ICT in this rulemaking scenario. Finally, observations drawn from this prototype application of comparisons between drafted rules and public comments are presented.

Apart from the application of a relatedness analysis system on e-rulemaking, there are many research works that focus on other aspects of ICT application on the making of law. For instance, Kerrigan developed an information infrastructure to promote regulatory compliance, which could potentially help users to reflect on the feasibility of different rules.³ Gardner addressed the open texture problem or, in other words,

¹ Cary Coglianese, “Information Technology and Regulatory Policy: New Directions for Digital Government Research,” *Social Science Computer Review* 22, no. 1 (2004): 88.

² Gloria T. Lau, Kincho H. Law, and Gio Wiederhold, “Similarity Analysis on Government Regulations,” in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 111-117 (Washington, D.C., 2003).

³ Shawn L. Kerrigan, “A Software Infrastructure for Regulatory Information Management and Compliance Assistance” (Ph.D. diss., Stanford University, 2003).

incomplete definition of many legal predicates of the law.⁴ It is suggested that “framers of legal rules have often abandoned clear directives in favor of open textured rules.”⁵ Conceivably, Gardner’s development could help rule makers detect intentional or unintentional open textured rules during the process of rulemaking. Many others have attempted the application of artificial intelligence (AI), and in particular knowledge-based systems, to the law⁶. As such, the emergence of e-rulemaking coupled with the growing power of computers provides a very rich platform for research.

2. Automated Sorting of Public Comments with Respect to Drafted Rules

A relatedness analysis system previously developed for regulatory comparison is enhanced to help screen and filter public comments. By comparing a set of drafted rules

⁴ Anne von der Lieth Gardner, *An Artificial Intelligence Approach to Legal Reasoning* (Cambridge, MA: The MIT Press, 1987).

⁵ Donald H. Berman and Carole D. Hafner. “The Potential of Artificial Intelligence to Help Solve the Crisis in Our Legal System,” in *Applied Artificial Intelligence: A Sourcebook*, ed. Stephen J. Andriole and Gerald W. Hopple (New York: McGraw Hill, 1989): 499.

⁶ Trevor J.M. Bench-Capon, ed., *Knowledge-Based Systems and Legal Applications* (San Diego: Academic Press Professional, 1991).

Stefanie Brüninghaus and Kevin D. Ashley, “Improving the Representation of Legal Case Texts with Information Extraction Methods” in *Proceedings of the Eighth International Conference on Artificial Intelligence and Law* (St. Louis, Missouri, 2001), 42-51.

James Osborn and Leon Sterling, “JUSTICE: A Judicial Search Tool Using Intelligent Concept Extraction” in *Proceedings of the Seventh International Conference on Artificial Intelligence and Law* (Oslo, Norway, 1999), 173-181.

Erich Schweighofer, Andreas Rauber and Michael Dittenbach, “Automatic Text Representation, Classification and Labeling in European Law” in *Proceedings of the Eighth International Conference on Artificial Intelligence and Law* (St. Louis, Missouri, 2001), 78-87.

Peter Wahlgren. *Automation of Legal Reasoning: A Study on Artificial Intelligence and Law* (Deventer, The Netherlands: Kluwer, 1992).

with the associated public comments using the relatedness analysis system, the public comments are automatically sorted with respect to related provisions in the draft. The source of data is from the US Access Board, which released a newly drafted chapter for the ADAAG,⁸ titled “Guidelines on Accessible Public Rights-of-Way.”⁹ This draft is less than 15 pages long. Over a period of four months, the Board received over 1,400 public comments representing around 10 megabytes of data, where some comments are longer than the original draft. To facilitate understanding of the comments with reference to the draft, a relatedness analysis is performed on the draft chapter and the comments. In this section, we discuss the technology behind this automated sorting of comments, followed by an illustration of the use of this framework.

Relatedness analysis among regulations and supplementary documents should desirably identify elements in documents that are alike and/or connected by a discoverable relation. To locate related material among regulations and public comments, we use certain characteristics of regulations that add knowledge to the comparison. In particular, most regulations or drafted rules are domain-specific, focusing on a narrowly-defined issue or area of interest. The drafted chapter from the Access Board is an example that focuses exclusively on accessible public rights-of-way. Another property of regulations is their natural hierarchical organization and referential structure. Provisions are structured in a parent-child relationship to reflect contextual coherences. Provisions are also heavily referenced and linked to one another. Thus, the computational properties of regulations, defined as the regulatory structure and any available domain knowledge, can be utilized in the comparisons between the drafted rules and public comments.

The first stage of a relatedness analysis is document parsing. The drafted rules and the public comments are compiled in the same machine-understandable format before an analysis is performed. The Access Board posted both the drafted chapter and the comments in free-form HyperText Markup Language (HTML). However, HTML is not suited for representing domain knowledge and organizational structure of regulations,

both of which are important elements in a relatedness analysis. For this task, we use the eXtensible Markup Language, XML, as the representation format.

To convert HTML documents into XML, a parser is developed to encapsulate each provision or comment as a single XML element. The collective set of 1,400 public comments is parsed as one XML document containing 1,400 elements, each corresponding to a single comment. The parser also recreates the structure of the drafted chapter automatically through pattern matching. For instance, Section 1109.2 from the draft is created as a child XML element of Section 1109. References are automatically extracted as well. Public comments possess no specific structure that needs to be extracted. Figure 1 shows an excerpt from the XML version of the draft that illustrates the XML elements and the structure. By using this parser, we obtain two XML documents, one representing the drafted chapter and one representing the set of public comments.

```
<regulation id="rights-of-way draft">
  <regElement id="adaag.1101" name="Application and Administration">
    ...
  <regElement id="adaag.1109" name="On-Street Parking">
    <regElement id="adaag.1109.1" name="General">
      <regText> Car and van on-street parking spaces shall comply with 1109.
    </regText>
    </regElement>
    <regElement id="adaag.1109.2" name="Parallel Parking Spaces">
      <regText>
        An access aisle at least 60 inches (1525 mm) wide shall be provided at street
        level the full length of the parking space. The access aisle shall connect to a
        pedestrian access route serving the space. The access aisle shall not encroach
        on the vehicular travel lane.
        EXCEPTION: An access aisle is not required where the width of the sidewalk
        between the extension of the normal curb and boundary of the public right-of-
        way is less than 14 feet (4270 mm). When an access aisle is not provided, the
        parking space shall be located at the end of the block face.
      </regText>
    </regElement>
    <regElement id="adaag.1109.3" name="Perpendicular or Angled Parking Spaces">
      ...
    <regElement id="adaag.1109.4" name="Curb Ramps or Blended Transition">
      ...
    </regElement>
  </regulation>
```

Figure 1: Drafted regulation in XML format

The next step in document preparation is feature extraction. Feature extraction is a form of pre-processing, for example, combining input variables to form a new variable. Features often are constructed by hand, based on some understanding of the particular problem being tackled.⁷ In our framework, important features representing available domain knowledge are extracted and incorporated into the XML documents. Examples include concepts, measurements, and definitions. Feature extraction is performed semi-automatically using a combination of handcrafted rules and text mining techniques.⁸ Once the two XML documents are refined with the extracted features, they can be compared using the relatedness analysis framework.

The relatedness analysis framework compares each XML element from one document with each XML element from another document. In the example case, each provision from the drafted chapter is compared with each of the 1,400 public comments. To compare provisions with comments, a similarity score is computed per pairs of provisions and comments based on the computational properties, including feature matching and structural matching as defined earlier. Here, the basic procedure for the relatedness analysis is discussed.⁹

We define feature matching as the computation of relatedness between two elements, based on their shared features using the vector model.¹⁰ A vector representing

⁷ Christopher M. Bishop, *Neural Networks for Pattern Recognition* (New York: Oxford University Press, 1995).

⁸ Marti A. Hearst, “Untangling Text Data Mining” in *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics* (University of Maryland, 1999), 3-10.

⁹ Gloria T. Lau, “A Comparative Analysis Framework for Semi-Structured Documents, with Applications to Government Regulations” (Ph.D. diss., Stanford University, 2004).

¹⁰ Gerard Salton, ed., *The SMART Retrieval System: Experiments in Automatic Document Processing* (Englewood Cliffs, NJ: Prentice Hall, 1971).

different features is developed for each element of comparison, and the *relatedness* between two elements is defined to be the cosine distance between two vectors. We employ a vector space transformation,¹¹ i.e., a mapping onto an alternate space prior to a cosine computation between two vectors, to incorporate available domain knowledge into the analysis. The importance of domain knowledge can be illustrated with an example. In the area of accessibility, a domain expert clarified that “[t]he terms ‘lift’ and ‘elevator’ although synonymous in definition in normal English usage have evolved into specific references in North America.”¹² It is clear that domain knowledge is not replaceable by common sense or dictionary knowledge. Therefore, feature matching between two XML elements reflects how much resemblance can be inferred between the pair of elements based on their shared features, such as domain knowledge.

Apart from feature comparisons, structural matching aims to reveal hidden similarities that are embedded in the organizational structure of regulations. The hierarchical and referential structures of regulations are incorporated into the relatedness analysis. Neighboring provisions are compared to identify similarities that are not apparent through a direct element-to-element comparison. Referenced provisions can be compared using an analogous approach, similar to citation and link analysis.¹³ Together,

Gerard Salton and Michael J. McGill, *Introduction to Modern Information Retrieval* (New York: McGraw-Hill, 1983).

¹¹ Gene H. Golub and Charles F. Van Loan, *Matrix Computations* (Baltimore: The Johns Hopkins University Press, 1983).

¹² David C. Balmer, “Trends and Issues in Platform Lifts,” (presented at Space Requirements for Wheeled Mobility Workshop, Buffalo, NY, October 9-11, 2003).

¹³ Sergey Brin and Lawrence Page, “The Anatomy of a Large-Scale Hypertextual Web Search Engine.” *Computer Networks and ISDN Systems* 30, no. 1-7 (1998): 107-17.

Eugene Garfield, “New International Professional Society Signals the Maturing of Scientometrics and Informetrics.” *The Scientist* 9, no. 16 (1995): 11.

Larry Page and others, “The PageRank Citation Ranking: Bringing Order to the Web,” Stanford University, <http://dbpubs.stanford.edu:8090/cgi-bin/makehtml.cgi?document=1999/66>.

feature comparisons, hierarchical, and referential structure matching define the basis of our relatedness analysis for regulations.

The results of a relatedness analysis are related pairs between the provision from the draft and individual comments. Figure 2 shows the developed framework where users are given an overview of the draft along with related comments. Industry designers, planners, policy makers, as well as interested and affected individuals are potential users who can benefit from the exploration of relevant provisions and comments provided by this framework.

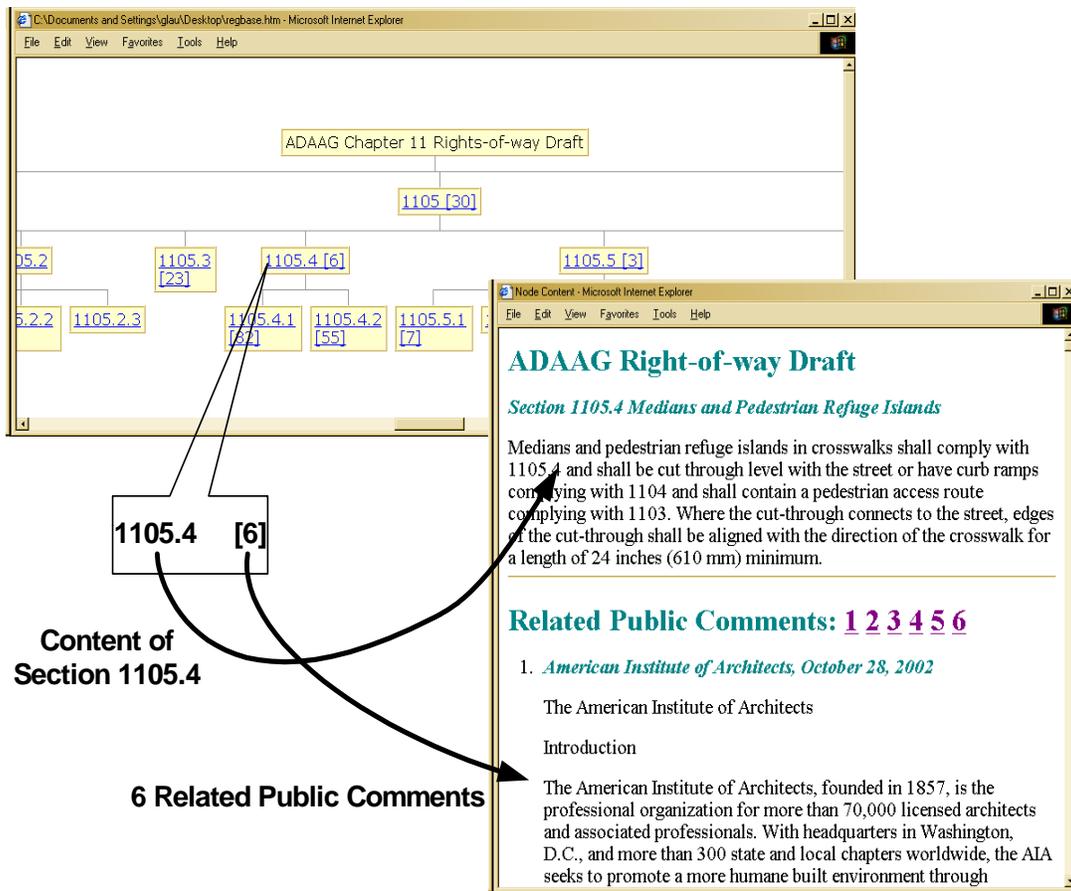


Figure 2: Comparisons of drafted rules with public comments in e-rulemaking

As shown in Figure 2, the drafted regulation appears in its natural tree structure with each node representing sections in the draft. Next to the section number on the

node, for example, Section 1105.4 is a bracketed number that shows the number of related public comments identified. Users can follow the link to view the content of the selected section in addition to its retrieved relevant public comments. This prototype demonstrates how a regulatory comparison system can help improve the e-rulemaking process where one needs to review drafted rules based on a large pool of public comments.

3. Results and Observations

Several interesting results illustrate the potential impact as well as limitations of the use of a comparison framework on rulemaking. Figure 3 shows a typical pair consisting of drafted section and its identified related public comment. Section 1105.4.1 in the draft discusses situations in which “signal timing is inadequate for full crossing of traffic lanes.” Indeed, one of the reviewers complained about the same situation, where in the reviewer’s own words, “walk lights that are so short in duration” should be investigated. (See Figure 3). This example illustrates that our system correctly retrieves related pairs of drafted section and public comment, which aids user understanding of the draft. Another observation from this example is that a full content comparison between provisions and comments is necessary, because title phrases, such as “length” in this case, are not always illustrative of the content. Automation is needed as it would otherwise require a lot of human effort to perform a full content comparison for the large number of comments.

A different type of comment screening is shown in Figure 4. It is an even more interesting result in which a particular piece of public comment is not latched with any drafted section. Indeed, this reviewer’s opinion is not shared by the draft. This reviewer commented on how a visually impaired person should practice “modern blindness skills from a good teacher” instead of relying on government installed electronic devices on streets to help. This opinion is not represented in the drafted document from the Access Board, which explains why this comment is not related to any provision according to the relatedness analysis system. As shown in the two examples, by segmenting the pool of comments according to their relevance to individual provisions, our system can

potentially save rule makers a significant amount of time reviewing public comments in regard to different provisions of the drafted regulations.

ADAAG Chapter 11 Rights-of-way Draft
Section 1105.4.1: Length

Where **signal timing is inadequate for full crossing of all traffic lanes** or where the crossing is not signalized, cut-through medians and pedestrian refuge islands shall be 72 inches (1830 mm) minimum in length in the direction of pedestrian travel.

Public Comment
Deborah Wood, October 29, 2002

I am a member of The American Council of the Blind. I am writing to express my desire for the use of audible pedestrian traffic signals to become common practice. Traffic is becoming more and more complex, and many traffic signals are set up for the benefit of drivers rather than of pedestrians. This often means **walk lights that are so short in duration** that by the time a person who is blind realizes they have the light, the light has changed or is about to change, and they must wait for the next walk light. this situation can repeat itself again and again at such an intersection, which can make crossing such streets difficult, if not impossible. I was recently hit by a car ...

I am Deborah Wood. My address is 1[...].
Thank you for your consideration.

Deborah Wood.

Figure 3: Related drafted rule and public comment

The provision and its related comment shown in Figure 5 suggests that a comparison between drafted provisions and comments is indeed the right approach. This commenter started by citing Section 1109.2 in the draft, followed by a list of suggestions and questions about Section 1109.2. Our system gathered the relatedness between Section 1109.2 and this comment through different features, such as the shared phrases. This piece of comment is a representative example of a lot of comments that are written similarly: comments that are concerned about a single provision in the draft. Thus, a comparison between drafted provisions and comments is important to help users focus on the comments that are most closely related to each provision.

ADAAG Chapter 11 Rights-of-way Draft
[None Retrieved]

No relevant provision identified

Public Comment
Donna Ring, September 6, 2002

If you become blind, no amount of electronics on your body or in the environment will make you safe and give back to you your freedom of movement. You have to **learn modern blindness skills from a good teacher**. You have to practice your new skills. Poor teaching cannot be solved by adding beeping lights to every big Street corner!

... If you want blind people to be "safe" then pray we get better teachers of cane travel.

I am utterly opposed to mandating beeping lights in every city. That is way too much money to spend on an unproven idea that is not even needed.

Donna Ring

Figure 4: A piece of public comment not related to the draft

Based on the observation made from the example shown in Figure 5, there seems to be room for improvement for an e-rulemaking portal. The public might find it helpful to submit comments on a per provision basis, in addition to a per draft basis. With the available technology, it should be possible to develop an online submission system that allows for both types of comment submission. It saves participants the time it would take to paraphrase or cite the provision concerned. It also saves rule makers the time necessary to locate related comments either through human effort or by using an automated system. Comments submitted on a per draft basis can still be analyzed and compared with the entire draft to identify any relevant provisions. On a side note, this commenter also suggested that it is important to forward the comment to the right person. An extension of this relatedness analysis framework could be developed to inform automatically any assigned personnel in charge of reviewing the provision within government agencies.

Apart from correctly identifying comments that are related to different provisions, limitations of our system have also been observed. Section 1109.2 is related to another

comment as shown in Figure 6. The relatedness is revealed through the shared features between Section 1109.2 and the comment, which includes a direct quotation and revision of Section 1109.2. The identified relatedness is correct; however, suggested modifications and revisions of provisions cannot be detected automatically. In essence, our current system is able to uncover the relatedness but not the revised version of provisions embedded in the comments. To locate precisely the revisions suggested in the comments, one can potentially perform linguistic analysis to compute differences between the drafted version and the suggested version. This is assuming that the suggested revision does not differ significantly from the draft such that patterns can still be matched.

Finally, Figure 7 shows a piece of public comment that is not identified as relevant to any provision in the draft. This reviewer commented on the general direction and intent of the draft, which explains why our system failed to sort this comment into any provision. Furthermore, this particular result suggests that a comparison between provisions and comments might not be enough. One could use the same analysis framework to compare comments with one another. For instance, this reviewer supported the positions of the American Council of the Blind (ACB) and the Washington Council of the Blind (WCB). While our system failed to associate this comment with any provision, comments submitted by ACB and WCB might give a clue to where this comment should belong. Essentially, clustering of comments alone could be as handy as the illustrated clustering of comments and provisions.

ADAAG Chapter 11 Rights-of-way Draft
1109.2 Parallel Parking Spaces

An access aisle at least 60 inches (1525 mm) wide shall be provided at street level the full length of the parking space. The access aisle shall connect to a pedestrian access route serving the space. The access aisle shall not encroach on the vehicular travel

lane.

EXCEPTION: An access aisle is not required where the width of the sidewalk between the extension of the normal curb and boundary of the public right-of-way is less than 14 feet (4270 mm). When an access aisle is not provided, the parking space shall be located at the end of the block face.

Public Comment

Norman Baculinao, P.E., PTOE, August 26, 2002

1109.2 Parallel Parking Spaces. An access aisle at least 60 inches (1525 mm) wide shall be provided at street level the full length of the parking space. The access aisle shall connect to a pedestrian access route serving the space. The access aisle shall not encroach on the vehicular travel lane.

EXCEPTION: An access aisle is not required where the width of the sidewalk between the extension of the normal curb and boundary of the public right-of-way is less than 14 feet (4270 mm). When an access aisle is not provided, the parking space shall be located at the end of the block face.

1. This section needs to be clarified, i.e., where is the access aisle located? that is, "will it be on the driver side or passenger side?"

2. The following is more of a question/concern about this requirement:

In downtown areas where parking is premium, this requirement will make it very difficult ...

3. The requirement for the exception is install the parking stall at the end of the block ...

I would really appreciate, if you **could forward this comments to the right individual and hopefully get a response back**. Please feel free to call me for any clarifications regarding this comments.

Sincerely,

Norman Baculinao, P.E., PTOE

Traffic Engineering Manager

Department of Transportation, City of Pasadena

Figure 5: Comment intended for a single provision only

ADAAG Chapter 11 Rights-of-way Draft

1109.2 Parallel Parking Spaces

An access aisle at least 60 inches (1525 mm) wide shall be provided at street level the full length of the parking space. The access aisle shall connect to a pedestrian access route serving the space. The access aisle shall not encroach on the vehicular travel

lane.

EXCEPTION: An access aisle is not required where the width of the sidewalk between the extension of the normal curb and boundary of the public right-of-way is less than 14 feet (4270 mm). When an access aisle is not provided, the parking space shall be located at the end of the block face.

Public Comment

Bruce E. Taylor, P.E., October 25, 2002

Re: Request for Comments on the Draft Guidelines for Accessible Public Rights-of-Way.

The Oklahoma Department of Transportation has reviewed the proposed draft guidelines for ...

Further, Section 1109.2, Parallel Parking Spaces, states;

An access aisle at least 60 inches (1525 mm) wide shall be provided at street level the full length of the parking space. The access aisle shall connect to a pedestrian access route serving the space. The access aisle shall not encroach on the vehicular travel lane. EXCEPTION: An access aisle is not required where the width of the sidewalk between the extension of the normal curb and boundary of the public right-of-way is less than 14 feet (4270 mm). When an access aisle is not provided, the parking space shall be located at the end of the block face.

Flexibility should be afforded the Engineer to allow off-street accessible parking, where available, in a reduced vehicular environment common to most minor streets adjoining heavily traveled thoroughfares. The Department would propose that the requirements of Section 1104.12 requiring one compliant parking space per block face, be removed, and **Section 1109.2 be revised to read;**

An access aisle at least 60 inches (1525 mm) wide shall be provided at street level the full length of the parking space. The access aisle shall connect to a pedestrian access route serving the space. The access aisle shall not encroach on the vehicular travel lane. EXCEPTION: An access aisle is not required where the width of the sidewalk between the extension of the normal curb and boundary of the public right-of-way is less than 14 feet (4270 mm). When an access aisle is not provided, the parking space shall be located at the end of the block face or on adjacent connecting streets.

The Department appreciates the opportunity to comment on the Draft Guidelines for Public Access. Should you have questions or comments, please advise.

Sincerely,
Bruce E. Taylor, P .E.
Chief Engineer
Oklahoma Department of Transportation

Figure 6: Suggested revision of provision in comment

ADAAG Chapter 11 Rights-of-way Draft
[None retrieved]

No relevant provision identified

Public Comment

Douglas L. Hildie, September 13, 2002

I am responding to a request from a fellow member of the blind community in this nation. She, and I, are members of the American Council of the Blind (ACB), its state affiliate the Washington Council of the Blind (WCB), and local chapters in our communities. **I support the positions of ACB, WCB**, and many people who are blind that, failure of national, regional, and local government to provide for the require and implement rational policies and practices resulting in the installation of tactile warnings and audible pedestrian signals at intersections would be unjustified and unjustifiable.

... It is obvious, I believe, that blind people are not "all the same", any more than any group of individuals is "all the same". It is true for "sighted people", and for "blind people", that some will have varying degrees of functional ability. But, contrary to the ideological perspective being foisted upon the public at large by a foolish few in the broader community of blind persons, people who are blind cannot do everything ...

Douglas L. Hildie

Figure 7: Comment on the general direction of draft

4. Conclusions

E-rulemaking defines the process with which electronic media are used to aid traditional rulemaking. In particular, government agencies are required to invite public comment for newly drafted rules. Electronic media provide an easy-to-access environment for the public to submit comments. On the other hand, an increasingly unmanageable amount of electronic data, in the form of public comments, can be easily created. There is a need for an analysis tool to help rule makers and interested and affected individuals review drafted rules along with the received public comments.

We applied a relatedness analysis framework that compared drafted regulations and public comments to illustrate the use of ICT on rulemaking. An automated sorting of public comments with respect to provisions in drafted rules was performed. Drafted regulations and public comments were first parsed into a consolidated XML format. We extracted relevant features, such as key concepts that represent domain knowledge, from the XML documents. Parent-child and referential structures were also extracted. We

then performed a relatedness analysis based on the computational properties of the documents, namely domain knowledge and structures.

The most challenging task was the development of a comparison algorithm that accounts for the computational properties of regulatory documents. To illustrate the usage of these computational properties in comparison, we showed several examples of results using this system. There were pairs of provision and comment that were correctly identified as related to one another. Limitations were observed, in which comments that dealt with the general intent of the drafted rules were difficult to match. It is conceivable that more pairs of “related” provisions and comments can be retrieved by relaxing the matching algorithm, such as lowering the threshold similarity score. However, this can lead to an overwhelming number of “related” matches, which might not aid in the understanding of the draft and associated public comments as intended.

Potential future research directions include automated forwarding of comments to corresponding personnel in agencies, as well as automated clustering of comments. Linguistic analysis could help identify suggested provision revisions embedded in comments. An online comment submission portal, allowing for commenting per provision in addition to the existing per draft basis, could be valuable. Other applications of ICT on e-rulemaking, such as a compliance assistance system as mentioned in the introduction section, may also open routes for future research.

5. Acknowledgments

This research project is sponsored by the National Science Foundation, Contract Numbers EIA-9983368 and EIA-0085998.