

SemanticWeb Enabled Web Accessibility Evaluation Tools

Shadi Abou-Zahra
World Wide Web Consortium (W3C)
2004, Route des Lucioles BP 93
06902, Sophia-Antipolis, France
+33.67.104.80.43
shadi@w3.org

ABSTRACT

Evaluating Web sites for accessibility remains an effort intensive process. Potentially, evaluation tools can significantly improve the efficiency and quality of Web accessibility evaluations but the currently available tool market only provides little or no consistency in the reliability and performance amongst these tools. In fact, in some cases evaluation tools may be confusing or misleading to users with little or no experience in Web accessibility. This paper will highlight how the utilization of Semantic Web technologies in evaluation tools can facilitate the exchange of evaluation data between tools and hence provide new approaches to support designers, content authors, programmers, quality assurance reviewers, project managers, or other users in accomplishing their respective tasks during the development and maintenance of accessible Web sites. Furthermore, this paper will also highlight some additional usages of Semantic Web enabled Web accessibility evaluation tools beyond the scope of evaluation processes.

Keywords

Accessibility, Automated Evaluation, Conformance Testing, Quality Assurance, Validation, Semantic Web

1.INTRODUCTION

The W3C/WAI Web Content Accessibility Guidelines (WCAG) addresses accessibility requirements for people with disabilities. It is accompanied by a set of *Techniques* documents which explain to Web developers how to implement accessibility features in different Web technologies such as (X)HTML, CSS, SVG, or SMIL. In order to determine conformance of Web sites to these guidelines, numerous accessibility checks need to be made on each page. For large or sophisticated Web sites, comprehensive conformance evaluations can be time consuming and costly. This is one of the main motivations for developing the *Evaluation and Report Language* (EARL); a platform independent and vendor neutral data format to help leverage the capabilities and quality of Web accessibility evaluation tools which potentially reduce the time and effort required to evaluate Web sites for accessibility.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

W4A at WWW2005, 10th May 2005, Chiba, Japan
Copyright 2005 ACM 1-59593-036-1/05/05...\$5.00.

2.ANATOMY OF EARL

The *Evaluation and Report Language* (EARL) is roughly a collection of *Classes* (comparable to *Objects* in object-oriented programming languages) and *Properties* (comparable to *Attributes*), as well as basic relationships between these *Entities*. These terms are expressed using the Resource Description Framework (RDF) which does not actually define any domain specific vocabulary but rather a basic set of expressions to allow the definition of vocabularies. It is similar to XML which does not define application specific vocabulary but rather a formal grammar for the syntax of these; only that RDF also represents semantic knowledge related to the corresponding vocabulary.

2.1Data Model

There are four core classes in EARL which define the basic data model of the language, these are expressed below:

Assertor

Context information about the test and contains information such as who (person) or what (tool) carried out the test, when the test was carried out, or information about the platform and configuration in which the test was executed.

Subject

The subject which is being evaluated is usually a Web page but could also be anything else such as a piece of software code, a document, or an object which is not available on the Web (e.g. a fax machine).

Testcase

The test criteria against which a subject is being evaluated could be any (referenceable) specification, guidelines, a single test from a test suite, or some other test case.

Result

The outcomes of a test conducted on a subject can contain information about the success or fail of the test, a confidence indicator for the obtained results, the reason for failing, or other attributes.

2.2Inheritance

RDF allows classes and properties to be subclassed into further entities, while still retaining backward compatibility. This happens through simple inference of the class type through inheritance. For example, a fax machine manufacturer may want to subclass the *Subject* to a new class called *FaxMachine1* in order to identify a specific model of fax machines which is currently being tested. Outside this company, other systems can still deduce that *FaxMachine1* is a subject (because its parent in the hierarchy is a subject) without really knowing much more

about what *FaxMachine1* actually is. So, while EARL defines a minimum set of entities required to describe test results, it does not confine developers from extending these core classes with domain specific terminology. In fact, RDF (and hence EARL) ensures the compatibility of the test results despite such extensions.

2.3 Compatibility

Because EARL is implemented as an RDF Schema it is automatically part of the larger Semantic Web community and can benefit from existing standards, tools, and implementations. For example to describe persons, EARL enabled Web accessibility evaluation tools can use widely accepted and deployed vocabularies such as *Dublin Core* or *FOAF*. Test results expressed in EARL can be processed according to readily available ontology languages such as *DAML* or *OWL*. Finally, EARL also inherits compatibility features to numerous RDF query languages such as *RDQL* or *SPARQL*.

3. USE CASES FOR EARL

The Evaluation and Report Language (EARL) is merely a standardized format for describing test results, not necessarily restricted to Web accessibility. However, this generic design of EARL allows it to be flexible enough for many usages in the field of Web accessibility evaluations and beyond. The following will describe some of the use cases of EARL in the context of Web accessibility.

3.1 Combine Reports

Web accessibility evaluation tools vary greatly in their capability to test WCAG Checkpoints. For example, while some evaluation tools have more advanced color contrast analysis algorithms, others perform better in text analysis. EARL provides a standardized data format which allows test results from fully-automated, semi-automated, or manual evaluation tools to be collected into a single repository. This allows reviewers to easily integrate several evaluation tools into the review process and maximize the number of checks which are evaluated with the assistance of tools.

3.2 Compare Test Results

EARL allows the test results from different Web accessibility evaluation tools to be compared against each other. For example, test results comparison tools could allow reviewers to prefer the results from certain evaluation tools for specific WCAG Checkpoints more than others. Such tools could then verify and infer assertions based on the test results collected from different tools as well as the user preferences. This minimizes the rate of false negatives (not identifying existing errors) and false positives (incorrectly reporting non-existing errors) in evaluation reports.

3.3 Benchmarking Tools

Because EARL test results are expressed in a uniform format, the EARL results generated from Web accessibility evaluation tools could also be compared to a set of results which are known to be correct. For example, evaluation tool developers could compare the output of their tools to the WCAG test suites. This may reduce misinterpretations of the guidelines and facilitate more consistency and quality across the tools.

3.4 Data Mining

A standardized format such as EARL encourages the development of data processing tools that analyze, sort, prioritize, or infer test results according to different policies. For instance, it may sometimes be desirable to sort the test results according to their corresponding severity (for example by matching them to the respective WCAG Priorities). In other cases, the relative cost of repair for an accessibility barrier may be the key by which a tool may be configured to sort the test results by. Data processing tools can also output their reports in EARL format to allow cascades of EARL enabled tools with different specializations.

3.5 Reporting

Test results can contain comprehensive information for different end-users. For example line numbers and detailed error messages for Web developers, or less verbose technical detail for project managers and executives. Repair suggestions and educational resources may sometimes be helpful to educate developers new to Web accessibility, but may also be tedious for more experienced ones. The well defined structure of EARL allows customized data views to be made from the same set of test results in order to suite the preferences of the end-users.

3.6 Integration into Authoring Tools

EARL implicitly provides a standardized, vendor neutral interface between Web accessibility evaluation tools and authoring tools. Instead of generating reports with test results, authoring tools could directly process these machine readable results and assist Web developers in finding and fixing errors through appropriate prompts and dialogs. This features also benefits evaluation tool vendors who want to focus on developing specialized accessibility checking algorithms rather than on implementing full-blown tools with user interfaces; these modules could easily export their results to EARL enabled authoring tools.

3.7 Integration into Web Browsers

Web accessibility evaluation tools could add significant features to Web browsers by assessing the accessibility of the content and providing detailed results. Web browsers could then compare the user preferences to the encountered accessibility results and readapt the content accordingly. For example, a Web browser may be configured to re-render complex tables or to suppress moving content. On each of these occasions, evaluation tools could provide information about the existence and location of such content on Web sites.

3.8 Integration into Search Engines

Similar to Web browsers, search engines could also make use of EARL reports to respond to user requests according to their preferences. However, search engines may prefer to out-source such features to third party evaluation services that publish accessibility reports for Web sites according to specific guidelines (for example national policy requirements). This is another example of where EARL benefits from the powerfull querying mechanisms provided by the Semantic Web.

4. CHALLENGES AND RESEARCH

While EARL is maturing to become a stable and widely deployed standard, there are still some challenging research

questions which the W3C/WAI Evaluation and Repair Tools Working Group (ERT WG) is currently working on:

4.1 Pointing to Web Content

Making assertions about *Web Content* is not as trivial as it may seem. Web content changes constantly, making URI pointers invalid soon after their creation. Moreover, while some WCAG Checkpoints only refer to specific attributes such as the ALT attribute of HTML IMG elements, other Checkpoints refer to consistent navigation across whole Web sites.

While XPointers may provide the necessary granularity to reference Web content, it has many limitations; for example it only works for XML documents. Other discussions and approaches include digests or hashes to increase the robustness and persistency of content pointers; or time stamps and fuzzy logics are also being investigated. Generally, in all these methods there is a tradeoff between complexity and robustness of the solutions.

4.2 Reasoning Results

In order to make an assertion on a specific WCAG Checkpoint, it is usually necessary to execute several sub-tests. Evaluation tool vendors constantly develop new tests to improve the performance of their tools. Is it possible for evaluation tools to reason their assertions on a Checkpoint level while possibly referring to proprietary testing algorithms?

RDFS or OWL models could provide necessary mappings between different tests. For example, a vendor may develop a proprietary test that is equivalent to a test described in the WCAG test suite. By only using a symbolic URI for this proprietary test, the evaluation tool could provide more information about the result of an assertion. However, the exact mechanism for mapping and reasoning results still remains a research question.

5. CONCLUSION

EARL provides a mean for different types of tools to exchange test results amongst each other. These tools include (but are not limited to) Web accessibility evaluation tools, authoring tools, user agents, search engines, transformation tools, and data processing tools. EARL facilitates the interoperability of these

tools and their integration into current environments. For Web developers, the integration of Web accessibility evaluation tools into existing development environments such as editors or content management systems may reduce the time and effort required to carry out comprehensive evaluation reviews. For Web users, the integration of evaluation tools into search engines or browsers can significantly enhance their experience on the Web.

While existing implementations of EARL underline its success, EARL is still at a relatively early stage without a main stream support in Web accessibility evaluation tools. There are also some key challenges and research questions open which need to be addressed and resolved before EARL can become a stable W3C standard. The Evaluation and Repair Tools Working Group (ERT WG) is actively developing EARL in coordination with several related domains within and outside W3C in order to deliver a mature standard which can be of relevance outside the realm of Web accessibility.

6. REFERENCES

- [1] W3C Web Accessibility Initiative (WAI)
<http://www.w3.org/WAI/>
- [2] Evaluation and Repair Tools Working Group (ERT WG)
<http://www.w3.org/WAI/ER/>
- [3] Evaluation and Report Language (EARL)
<http://www.w3.org/TR/EARL10/>
- [4] Resource Description Framework (RDF)
<http://www.w3.org/RDF/>
- [5] Web Ontology Language (OWL)
<http://www.w3.org/2004/OWL/>
- [6] Web Content Accessibility Guidelines (WCAG)
<http://www.w3.org/TR/WCAG10/>
- [7] W3C Semantic Web Activity (SW)
<http://www.w3.org/2001/sw/>
- [8] W3C Quality Assurance Activity (QA)
<http://www.w3.org/QA/>
- [9] Dublin Core (DC)
<http://www.dublincore.org/>
- [10] Friend-Of-A-Friend (FOAF)
<http://xmns.com/foaf/0.1/>