Why is Accessibility Metadata Proving Difficult?

Liddy Nevile
Motile Research
liddy@motile.net

Abstract

Accessibility metadata is simply metadata that describes the accessibility of resources and services, usually those on, or available through, the web. Awareness of widespread web content inaccessibility led to work being done to develop guidelines for authors and others to make sure that content would be more accessible to those with special access needs, especially those with disabilities who were being disenfranchised by their lack of access to the web. Currently, work is being done to find ways of signalling the degree of accessibility of resources, and ways of matching resources to searches and people. In addition, accessibility metadata could be used to repair some inaccessibility problems on the fly. This paper describes some of the work being done and the problems that have contributed to make the progress comparatively slow.
Keywords: *Accessibility, metadata, Dublin Core, DC-accessibility, Evaluation and Report Language, EARL, people with disabilities, guidelines, W3C, IMS.*

1. Introduction

Accessibility metadata is, put simply, metadata that describes the accessibility of resources and services, usually those on or available through, the web.

Web content accessibility became a topic in the mid nineties. It was realised that much of the content of the new 'web' was not accessible to people who did not use standard web GUI browsers, the same technology that was making the web attractive and available to naïve computer users. Many people complained, at that time, that they could not download the 'much too big' files, or that the colours were not consistent, but a whole swag[1] of people suddenly found that the very technology that had enabled them to rejoin society was suddenly alienating them. In particular, blind people, people with motor coordination problems, in fact, many people including those who could not use a mouse on a computer screen for one reason or another, were suddenly not able to use their computers as their life-style-support machines. Additionally, people who depended, for one reason or another, on screen readers were often being read content that was unrecognisably jumbled, according to the GUI layout specifications of the author.

The World Wide Web Consortium (W3C, [1]) responded by establishing a Web Accessibility Initiative (WAI, [2]) program to work on what was making web content inaccessible. Since that time, W3C WAI have developed extensive guidelines as to how to make content accessible to all devices, so that those using non-standard combinations of hardware and software could access web content, or content available through the web, if their devices were standards compliant. This work is undertaken under the banner of the Web Accessibility Initiative, is open to all, and is international, partly due to its special funding structures. The immediate aim was to avoid content that would be totally inaccessible to some, before working on making all content more generally accessible.

The W3C WAI works on how to make offending content accessible, often by repairing it. They concentrate on what accessibility would be but also on the authoring and user access tools. The WAI Authoring Tools Accessibility Working Group [3] has emphasised how to make authoring tools, for instance, productive of accessible content, even when the author is not aware of what is necessary. Such emphases were chosen to get the greatest benefit to the greatest number as quickly as possible, in the realisation that authoring was becoming more complex and more people would soon be using authoring tools.

Repairing an inaccessible page; identifying inaccessible content; techniques for making accessible content, are all under control, if not yet completely documented. What is required now, is work on metadata to perform a number of roles. Of course, discovery is a primary goal. Finding a resource or service is an on-going problem on the web, and all the usual difficulties operate when people with special needs use

---

[1] Australian expression meaning what is rolled up and carried around as 'home' by tramps. It usually contains everything but the kitchen sink.

the web. Everyone has a need for information that suits their purposes at the time they seek it. This may occur in special circumstances, such as when people working underground in protective clothing, perhaps because they are working in a mine, need to access information (possibly how to deal with a leak), without using their hands, and so without keyboards. These users would possibly need to be able to use their voice-controlling software to use the command-key navigation of a web page. They will need to know if the page is properly constructed so that such navigation is possible. If it is not well-constructed, they may need to know:

- how it is constructed so they can determine if they will be able to get to the information in some compromised way, or
- if they can expect some transformation application to make it accessible or finally,
- if there is no hope of access for them.

The challenge is to find a suitable way of expressing and disseminating accessibility metadata and to make it available as soon as possible.

In this paper, we consider how the W3C Web Content Accessibility Guidelines (WCAG, [4]) and other guidelines, in a sense derived from the WCAG, have helped towards the problem of identifying how to produce metadata about resources and service accessibility, and what problems remain. Primarily, the author asserts that the rush to solutions[2], without time to establish a clear set of requirements, has left organisations interested in this metadata with the very difficult task of trying to fit requirements to solutions – known in the software industry generally as a serious nightmare situation! In many cases, the actual requirements will be known only at the time, and this makes it additionally difficult. Additional resources and style sheets, for instance, may also need to be retrieved to accompany the original resource.

2. Accessibility

2.1. Accessibility Information

Assessment of accessibility usually requires an assessor to identify what types of content are contained within a particular resource, and therefore which of the total array of guidelines apply, to what standard and thus, if the resource as a whole is compliant. At best, assessors then make an assertion about the accessibility or otherwise of a resource. (This is just another example of a situation in which it is important for the consumer of that assessment or evaluation to know who made it).

When single resources are to be evaluated, a report on their compliance is usually produced. As the num-

ber of resources increases, and the frequency of their evaluation, and the increase or otherwise of accessibility of collections becomes of concern, metadata management becomes an issue. Not only will people want metadata to discover the accessible resources, or to transform them appropriately, in addition document management agencies will want records and reports automatically generated about the accessibility of the resources, possibly integrated into their document management systems.

People can have special needs because they are temporarily in a situation, for instance one that makes their hands unable to work the keyboard, or the noise level so high they cannot hear a screen reader, or otherwise. In order to help people with special needs, it is necessary to identify their needs and requirements.

The most useful way to do this is to develop a normative set of requirements and then have users or agents select from that. This approach has been used: the banking industry, for example, has worked to accommodate selected people with special needs at their Automatic Teller Machines. The people fit particular profiles and are issued with smart cards that adapt displays for this purpose. Determining a comprehensive normative list is a non-trivial exercise, however.

Once such a list has been determined, the W3C WAI guidelines, for instance, can be matched to the requirements and relevant sub-sets of accessibility determined. This again is a major task, given that there are hundreds of checkpoints involved in compliance with the W3C WAI guidelines to achieve a general accessibility rating. The good news is that many of these checkpoints can now be tested automatically, at least to the level of partial compliance, or possibly more usefully, complete non-compliance. Some criteria, such as the provision of a text alternative to an image, is failed if there is nothing but not compliant unless what is there is intelligent.

2.2. Accessibility Solutions

The main accessibility solutions take a number of forms but three are notable:
- Guidelines as to what should and should not be published, or more precisely how content should be published, e.g., an image should not be published unless it is accompanied by an ALT (text alternative) tag, and, if the image conveys significant information that is required for comprehension of other content, it should be accompanied in addition by a long description, and that should be done in one of two ways, and so on.
  These guidelines, in the case of W3C WAI especially, have been developed after significant consultation and account for a range of access devices that may be in use, catering simultaneously for people with different devices and, incidentally, with a range of disabilities. In the case of W3C, these

---

[2] An activity in which the author has been engaged for a number of years.

guidelines are not model or agent specific, and are written to be robust under evolving conditions of the web and associated tools.

- Checklists [5] have been developed to allow assessors to work with some sort of consistency. These checklists provide more or less detail but aim to clarify what might be considered compliance, again differing in nature according to whether this information is to be understandable to a naïve assessor or only useful to an expert.

- Technique documents [6] aim to show that it is possible, and hopefully practical, for developers to author resources that will comply with the guidelines. They can be thought of, at one extreme, as proof-of-concept documentation, as there may be more ways of achieving the goal, but at the other extreme, these are tools for accessible content creation, The range and availability of such collections of techniques is extensive.

Finally, there are many situations in which organisations are developing what might be called accessibility policies. Making resources fully accessible can be burdensome, and may not always be appropriate. Organisations around the world are working on what is feasible in their context, what they will set as their local standard. Such policies often work by selecting criteria from the W3C WAI list. The guidelines do not go so far as to provide a specific, actionable set of requirements that can be immediately included in development specification documents. But more problematic is that neither do they provide for absolute testing of compliance: compliance is subjective, and for many of the guidelines, can only be assessed by an expert. Ultimately, of course, accessibility is totally subjective, and depends upon a user at the time. It should be remembered that accessibility as described in guidelines does not guarantee useability; that too has to be assessed subjectively by humans.

2.3. Accessibility Requirements

Imagine the needs of a deaf person.

A deaf person does not have hearing problems with content that does not have sound associated with it but in many cases, the spoken language used by the surrounding community is a second language for the deaf person. Such a person might be used to sign language, and so be a non-native speaker, and therefore reader, of text that is included in a resource. In some cases, text at all is not helpful to the deaf person, and they may require the text to be *translated* into sign language for them. The requirements for accommodating a deaf person are not simple, as shown, but depend upon the strategies adopted by the deaf person to operate with their disability.

Deciding whether a resource is suitable for a deaf person, if done by a third person, is a form of censorship. In many cases, this is not appreciated by people with disabilities: they would prefer to know what it is

with which they might have difficulty, and then decide how much effort to make and what compromises are acceptable to them in the particular circumstances.

Even where it is necessary according to business or functional specifications for a resource to be classified as suitable for deaf people, as might need to happen if an organisation's accessibility policy is to ensure its resources are accessible to deaf people, it will not be a straight-forward matter. Accessibility technical requirements specify, *in technical terms*, what resources need to do and usually the developers have to determine the most appropriate way to achieve these aims. This is the ideal situation and likely when professional developers are at work, using the most rigorous techniques. It is not what happens in most cases in practice.

And, as shown above, there are so many variables and dependencies that it may be better for the deaf person to be enabled to say what they want, by choosing from what is available, than to be subjected to some automatic feeds.

Such considerations often lead to a call for description of people's needs and sometimes, on to descriptions of people's disabilities. Again, this is not considered an appropriate approach in all circumstances, especially as it can easily degenerate into breaches of privacy, and error.

3. Resource Matching

3.1. Matching Requirements to Solutions

Metadata that describes the accessibility or otherwise of a resource in particular circumstances is likely to include a lot of information and not be the sort of thing a person will just read, in the way they might have read a paper library catalogue record. It is not likely that something as simple as 'accessible' or 'not accessible' will be meaningful.

In fact, in order to promote accessibility, and reward those who tried, W3C introduced a scheme of levels of accessibility, A, AA and AAA. Unfortunately this proved over-simplistic. Compliance with all but one small detail in a set of criteria meant failure to that level, even if there was also compliance with almost all the other criteria necessary for compliance with a higher level. This proved discouraging and not helpful. A new system has been proposed for the future. Other organisations promoting accessibility have tried rewarding authors with icons [7] but these have been abused by ill-intentioned and misused by ignorant[3] people to the point where the icons lack credibility. Anyway, they lack precision and are not very useful to users with specific needs.

---

[3] A typical example is offered where people try to evaluate sites using the Bobby test and do not read the fine print and assume their site is accessible when Bobby just says it would be accessible if …

This leaves a situation where applications are required to handle all the accessibility metadata. In such a case, inter-operability is also important and as part of that, the semantics and syntax used for the metadata. In addition, as the quantity of metadata increases, due to continual testing and other document management activities, metadata management will become important.

Fortunately, the work of W3C on syntax in the context of Resource Description Framework (RDF, [8]), used already by many for metadata, has led to a useful development that will help solve some of these problems. Evaluation and Report Language (EARL, [9]) is a derivative of RDF and has the form:

A asserts that X was compliant with Y on Z date

Such a statement leaves it up to the user, or agent, to decide on the trustworthiness of such a statement but makes it possible for a computer to interpret the statement and, if necessary, act on it. In such a case, the computer may be using discovery software but equally, transformation or other applications or even searching for complementary resources, such as captions files.

3.2. Matching Resources and Services to Users

Bringing together users and resources, as has been mentioned, is a bit like dealing with censorship. How best to do it is probably more of a political issue than a technical issue, but nonetheless difficult for that. Technically, the choices relate to the difference between server-side and client-side solutions. Servers can restrict what is made available to users:
- during the discovery process;
- at delivery time, or even;
- from within a composite resource that has alternative content that is intended to be varied according to a request of the user agent seeking it.

In some situations, users will not want to receive content they cannot use because they will have telecommunication constraints that make redundancy expensive in terms of time, money, and in other ways. In other situations, possibly the same person will want to receive complete resources, like everyone else, in order to maintain privacy about their needs.

User agents, acting on the client-side, can modify requests before making them to a server, or filter what is received and present only portions to the user.

Client-side technology that immediately pre-dated RDF, designed for such a purpose, was the Platform for Internet Content Selection (PICS[4], [10]). To use PICS, a user selects from a range of options made available on a form by their user agent, and that

information is converted into a numerical representation to be used by the user agent to control user presentations. PICS was extended to include semantic and structured values and values for repeated variables and transformed into RDF.

PICS remains, however, as a technology that might be useful in this context. The IMS Global Project (IMS, [11]), a consortium of those interested in developing tools for education, especially in situations where a student's path through the use of resources and assessment points is to be monitored, have adopted the approach of using a 'Learner Information Profile'. It is the IMS' intention to add some accessibility requirements into this profile. This will, somehow, be matched with resource accessibility metadata that will be attached to the resource.

3.3. Developing the Vocabularies

Finally, just as with any other classification system, different organisations will have local purposes and requirements and will want their metadata to support those needs. This means that there is a distinct likelihood of the different agencies wanting to use different sets of elements to make up their vocabularies for their metadata. One of the factors that makes it 'hard' to work in the field of accessibility metadata is that there are not already accepted and tested vocabularies. There are not even keywords in common usage. Accessibility metadata is a new idea.

W3C WAI, having the most comprehensive set of criteria for accessibility evaluation, is working towards a numerical schema for identifying compliance with individual criteria. Such a schema could provide a useful resource for others who then could merely pick and choose among the criteria for their local purposes. It would also promote inter-operability of accessibility metadata.

Computer-Useable Metadata

Another quality of accessibility metadata that is not unique but is challenging, is that this metadata is designed to be used by both computers and people. So far, there are not applications waiting to retrieve inaccessible resources to turn them, on-the-fly, into accessible resources. The enabling technology is already developed for this, however. In determining the metadata formats to be used, it is obviously important to have this potential in mind, but as it is not clear what will be required, or more practically, used, it is difficult to decide what to do about it.

4. Metadata Associations

4.1. Dublin Core and Accessibility Metadata

Finally, determining how the Dublin Core Metadata Element Set (DCMES, [12]) can be used to provide accessibility metadata is not yet settled.

---

[4] Later, PICS was associated with censorship, most particularly because it could be used at the proxy level, and so has not been a 'popular' technology.

One of the charter aims of the DC-Accessibility Interest Group [13] is:

- to determine the relationship between accessibility (evaluation and repair) descriptions and DC descriptions - and report on an appropriate way to represent accessibility information in DC.

Specific questions to be answered are:

1) Is accessibility a resource discovery issue?
2) What is the relationship between accessibility (W3C's EARL) descriptions and DC descriptions?
3) Is it sensible to embed one in the other?
4) Could one provide, as part of DC RELATION, information about the relationship between equivalent resources?
5) Should EARL schemata be recommended?

There is little debate about the value of accessibility metadata for discovery although there is some about how the requirements should be matched with the resource's qualities. Accessibility descriptions will not be about the content of the resource, but rather one of its qualities. This suggests that it is not DC-description type information, and so would not be well located there. Nor does this information make a lot of sense as DC-format information. It is not about the format, but how it is used.

The question of whether DC-relation is a good place for information about equivalent resources for use by those who cannot access the primary resource is not so clear. If a resource is well-developed, from the accessibility perspective, it will have information or alternative resources correctly embedded or associated with it, so the use of DC-relation will be the same in this context as others, to identify something that may be an interesting alternative resource. Where the original resource is not well-formed, and the user is dependent upon the replacement of one element or all of it by another, not originally associated with the primary resource, this information should be in the metadata relating to the primary resource. Otherwise, where what is required is the use of a different style sheet, for instance, this will need to be retrieved and the original resources related to it instead of whatever was originally proposed. In either case, the accessibility of the resource will be changed by the effect of this metadata, and the post-hoc provision of the equivalent element. This information should be closely associated with the description of the accessibility of the primary resources, as it affects it. For this reason, it makes more sense to have accessibility and alternative and/or equivalent resource elements co-located. So it is probably better to have a separate accessibility element with all this information in it.

EARL descriptions will be RDF expressions but deciphering metadata in RDF that is well-formed with clear parsing rules should not present a problem in a DC environment. One argument for embedding EARL statements, or other accessibility descriptions, in DC metadata is the potential for wide-adoption of accessibility metadata if it is associated with the incredibly wide-spread DCMES.

If DCMES is seen as a suitable affiliation, or association for accessibility metadata, the question remains as to how this will work out in practice. It can not be mandated that people must produce and use accessibility, or any other, metadata. Nor can it be predicted with accuracy whether they will. It is hoped, however, that if the Dublin Core Metadata Initiative can take the lead on this, with the collaboration of other interested bodies, that whatever is chosen as the DC-accessibility solution, will be useful to and popular among others concerned with this issue.

4.2. The way Forward …

What seems possible is for those working on accessibility metadata to work together. 2002 has been a year of integration and currently W3C, IMS and DC-Accessibility have brought together their activities to the individual benefit of all groups, as well as for the general problem.

EARL is reaching maturity, and expected to be promoted as a W3C recommendation by the end of 2002. IMS Metadata specifications, particularly the Learner Information Profile, designed to be a metadata management tool for tracking student progress and needs, is reaching maturity and also expected to be ready for recommendation by the end of 2002.

5. Conclusion

Although all metadata standards are 'hard' to develop, particularly as their global uptake depends upon local utility, as well as a number of factors to do with inter-organisational and international cooperation, accessibility metadata standards, in particular, are 'hard' to achieve.

References

[1] http://www.w3.org/
[2] http://www.w3.org/wai/
[3] http://www.w3.org/wai/au/
[4] http://www.w3.org/WAI/Resources/#gl
[5] see eg http://www.w3.org/WAI/Resources/#ch
[6] see eg http://www.w3.org/WAI/Resources/#te
[7] see eg http://bobby.cast.org/
[8] http://www.w3.org/RDF/
[9] http://www.w3.org/2001/03/earl/
[10] http://www.w3.org/PICS/
[11] http://www.imsproject.org/
[12] http://dublincore.org/
[13] http://dublincore.org/groups/access/