

A Social Web for Another Billion

T B Dinesh¹, Suzan Uskudarli²

¹Janastu, 3354 KR Road, Bangalore, 560070, India

Tel: +91 80 26762963, Email: dinesh@servelots.com

²Computer Enginnering, Bogazici University, Istanbul, Turkey

Email: suzan.uskudarli@boun.edu.tr

Abstract: The issue of Web-accessibility in the era of *Web on Mobile* in India is daunting, as most users are not likely to be sufficiently literate to browse the text heavy pages. Among many considerations in addressing this issue are the identification of the requirements for content consumable for those with different abilities and the determination of an approach for creating such content. Inspired by many success stories regarding participatory approaches, we propose a *propagator* (those who re-narrate content for wider consumption) approach for the creation of print-impaired friendly content. This segment is the majority of mobile users in India. This paper presents an approach for re-narrating web content for enabling access to print-impaired web users. A system called *alipi* for such re-narration is presented with examples.

1. Introduction

The Web has been remarkable in making a great collection of diverse content available at our fingertips. The widespread availability of mobile phones with Internet access expands the possibilities even greater. Access to phones crosses most socio-economic barriers, which presents very interesting opportunities and challenges [6,7]. Perhaps the most important challenge is making the content delivered to such devices *truly accessible*. Content is accessible, not only, if it is deliverable to a device, but that it also is understandable to the user of that device. Issues like illiteracy, disability, and languages are serious barriers to true access.

Historically, web content has been predominantly textual. More recently, audio, video, and image based content is also commonly available. Nevertheless, content is still dominated by text. Even when we consider video content, there are usually descriptions and user comments associated with them. Government policy documents and acts, for example, are always large textual documents.

Given the abundance of content and content providers -- thanks to participatory platforms of Web 2.0 -- we propose an approach whereby users can provide content based on existing content. In fact, bloggers often perform such services, by writing blogs based on something(s) they read on the Web. We propose a structured manner of doing this, where the relationship between the source and target are preserved. What we propose can be termed as a social semantic web [10], a collective knowledge system [11] that builds on capacities and interest among Web users who can provide alternative narrations of content that is suitable for a specific target community need.

1.1 Oral culture versus the Internet

It is too familiar in the “South”, that a society is based on transmission of history, literature, law and other knowledge across generations - orally i.e., without a writing system. Internet has however been based on text i.e., a writing system. There has been an influx of video documentation on the Internet over the last couple of years. This effect can be attributed to a large extent on the availability of cameras on mobile phones that make it not only easy to capture but also to upload the picture or video to a site like YouTube.com.

South Asia, which has been lagging behind the West in the number of people connected to the Internet, is bearing witness to the power of the mobile device to empower millions with information, content and services. The power of mobile innovations and applications in content and services delivery is on the rise. It is time that we in South Asia, the nations and governments in the region, as well as service providers, give a thrust on mobile for delivering meaningful services, especially that could empower and enable efficient day-to-day life for the larger masses.

1.2 Non-Formal Communications and community radio

There is a profound lack of appropriate and effective learning opportunities in remote, rural and resource-poor parts of the world. Schools, by and large, do not cater to non-formal or life-long learning needs of adults or youth. Universities and colleges have at best a limited footprint in most rural and remote areas and courses are rarely framed to meet the livelihood, health or development needs of communities or their members.

Just as it is hard to imagine universities offering non-formal educational services in developing areas, it's unlikely that community groups can fill the gap on their own. Local media, community development programmes, information and communication technology (ICT) centres, development and other localised services, even when information-based, are seldom effectively structured for engaged learning. There is however untapped potential in collaboration among these groups: educational institutions, local development agents, media/ICT groups and communities.

Community Radio has been advocated to play a large role in helping develop a culture of dialogs and information sharing for localized needs.

1.3 Accessibility and the Print Impaired

Web pages dominate with text. The Web-accessibility has been traditionally addressed as a concern by countries where illiteracy has not been a major issue and therefore accessibility concerns have been for the physically disabled. However, in countries like India, majority of people are uncomfortable with text, either because they are not literate or because they are literate only in their localized language.

There are several Assistive Technologies used for web browsing such as screen readers, speech recognition, screen magnification and keyboard overlays. There are web-page authoring guidelines developed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) [1]. These guidelines help understand and implement web accessibility. However, these guidelines do not have a prescription for the print-impaired or for the needs of an “oral web”. The issue of Web-accessibility for the print-impaired can be considered as the issue of the next-phase of Internet users - the

next billion new users who may not be as literate as the Internet users until now. This is a large class of people who cannot read the content on Web-pages and will include billions of people who are illiterate but might soon find it easy to access the Web through their mobile phones. This is also an issue for inter-cultural inter-language inter-contextual communication that the current web is trying to cope with.


On the one hand, we need to look at how authoring of Web-pages (the structure) and meta-information (appropriate meta-tags) can help and on the other, how the world of Web 2.0 can nurture a community around a Web-page so that the content becomes accessible to a wider group of people than what the capacity of the original author could address. For example, the sub-titling of TED talks [2] and the Mozilla UniversalSubtitles project [3] not only support but rely on such community contributions.

2. Re-narration Web

2.1 The idea


Lets take a web page, say of fire safety, which is originally authored in English as shown below.

HOME



Fire is the rapid oxidation of a material in the chemical process of combustion, releasing heat, light, and various reaction products. Slower oxidative processes like rusting or digestion are not included by this definition. The flame is the visible portion of the fire and consists of glowing hot gases. If hot enough, the gases may become ionized to produce plasma. Depending on the substances alight, and any impurities outside, the color of the flame and the fire's intensity might vary.


A fire department or fire brigade is a public or private organization that provides fire protection for a certain jurisdiction, which typically is a municipality, county, or fire protection district. A fire department usually contains one or more fire stations within its boundaries, and may be staffed by career firefighters, volunteers, or a combination thereof.



Fire in its most common form can result in conflagration, which has the potential to cause physical damage through burning. Fire is an important process that affects ecological systems across the globe. The positive effects of fire include stimulating growth and maintaining various ecological systems. Fire has been used by humans for cooking, generating heat, signaling and propulsion purposes. The negative effects of fire include decreased water purity, increased soil erosion, an increase in atmospheric pollutants and an increased hazard to human life.


Now imagine this link is sent to a person who does not read English but can read Hindi. The reader then asks if there is a Hindi narration of this page. And Web responds by looking for available Hindi narrations out there in various blog posts and renders the page using these. Say it can look like:

HOME



Fire is the rapid oxidation of a material in the chemical process of combustion, releasing heat, light, and various reaction products. Slower oxidative processes like rusting or digestion are not included by this definition. The flame is the visible portion of the fire and consists of glowing hot gases. If hot enough, the gases may become ionized to produce plasma. Depending on the substances alight, and any impurities outside, the color of the flame and the fire's intensity might vary.

आग विभाग या आगशमक दल एक सार्वजनिक या निजी समस्था है जो आग से होने वाली दुर्घटनाओ से सुरक्षा प्रदान करती है, जो आन तौर पे एक नगर - पालिका या जिन्ना का निरिक्षण करती है | एक विभाग के सिमा में आम तौर पे एक से अधिक आग शमक स्टेशन होते हैं | इन स्टेशनो मे व्यवस्तिक आग शमक या स्पयर्सियक



गर्व करते है |

Notice that a search for alternative narratives that are suitable for the Hindi speaker (from India) has found an image of an Indian fire engine and also found a Hindi narrative for the second paragraph and re-rendered the page by substituting the alternative narratives that are more suitable for the Hindi speaker.

Similarly a Turkish or a French person could ask for re-rendering of the page and see parts of the page or the whole page - depending on the contributed Turkish or French narratives out there on the Internet.

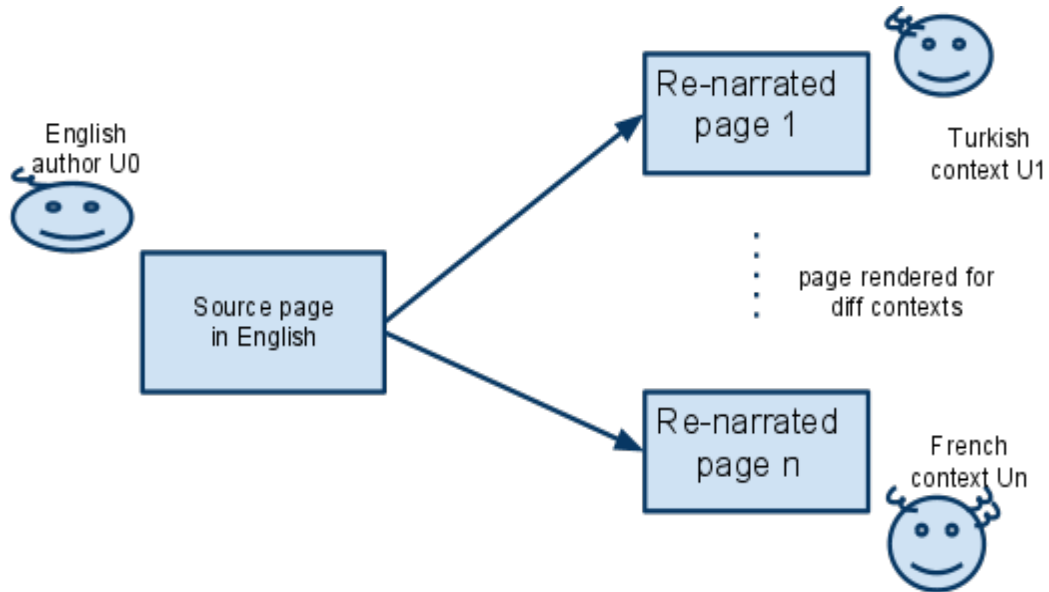


Figure 1. Re-narration scenario. Many re-narrations are rendered as alternative web pages to the source page.

Figure 1 describes the general process of how various page observers can see more appropriate renditions of the original page that serves as the source for alternative renditions.

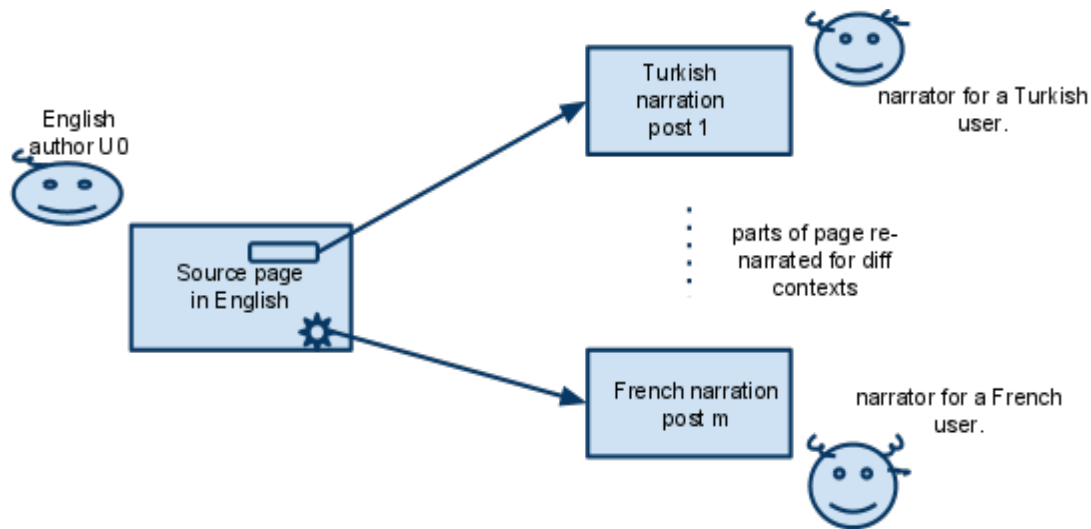


Figure 2. Re-narration authoring scenario. Many alternative narrations for diff contexts can be authored as contributions.

As we noticed in the page rendition for the Hindi reader, only part of the original page was re-narrated to a Hindi context. This means that alternative narrations were contributed for only parts of the original page. As Figure 2 shows, alternative narratives can be contributed to parts of a page by narrators who have an interest in narrations for a certain user group.

Technically, this is achieved by annotating a blog post as an alternative narration of a particular page of a source page. For example, [a11y.in/a11ypi/idea/firesafety.html](http://www.a11y.in/a11ypi/idea/firesafety.html) is the source page shown above. As explained in the idea page [a11y.in/a11ypi/idea](http://www.a11y.in/a11ypi/idea) [9], a blog post that is meant to be an alternative narrative saves as meta information a reference to the part of the source page that is the source for the alternative narrative. In the Hindi blog case the meta information was part of the paragraph tag in the blog and has the attributes **foruri** and **rec**:

```
<p id="hi" foruri="http://www.a11y.in/a11ypi/idea/firesafety.html:div1" rec="lang:hi">
आग वभिाग या आगशमक दल एक सार्वजनकि या नजिी समस्था है जो आग से होने वाली दुर्घटनाओं से सुरक्षा
प्रधान करती है, जो आम तौर पे एक नगर - पालकिा या जलि्ला का नरिक्षण् करती है। एक वभिाग के समिा में आम
तौर पे एक से अधकि आग शमक् स्टेशन होते हैं। इन् सटेशनो मे व्याव्सयकि आग शमक या स्वयंसेवक कार्य करते हैं।
</p>
```

This information is analogous to backlinks in blogs. However, the semantics attributed in this case is that the text in Hindi is an alternative narrative for the text at “id” div1 in the source page <http://www.a11y.in/a11ypi/idea/firesafety.html> and the “rec” attribute indicates that this narrative is **recommended** for Hindi speaking contexts.

2.2 Adaptation to a Mobile context

These contributed alternate narratives, in addition to *associations* [4] also make it easier to deliver a web-page as a slide show or effectively as an audio/video rendition, on smaller screens such as those of mobile phones and tablets. For instance, the fire safety page mentioned above can be rendered by serializing the associated information fragments on a mobile device. When a person on a mobile device receives a link to the above page on fire safety and the person can only understand Kannada, the page can be rendered like a slide show with an audio option. A snap shot can look like this:



This Kannada narrative is picked up by contributed narratives available on the Web. This content for this slide is available as a blog, which provides an alternative fire engine picture for Karnataka visitors and Kannada replacement paragraph for Kannada speakers. Similarly, the audio rendition of this is picked up from another narrative. Such a blog would look like:

Blog of Aravind

post saturday



ಅಗ್ನಿ ತಾಮಕ ವಿಭಾಗ ಒಂದು ಸಾಮಾಜಿಕ ಅಥವಾ ವಯಕ್ತಿಕ ಸಂಸ್ಥೆ ಆಗಿದ್ದು ಇದು ಅಗ್ನಿ ನಂದಿಸುವ ಕಾರ್ಯವನ್ನು ಮಾಡುತ್ತದೆ. ಎಲ್ಲ ಜಿಲ್ಲೆಯ ಉರಗಲು ಇ ಸಂಸ್ಥೆಯನ್ನು ವಲಗೂಂಡಿರುತ್ತವೆ ಪತಿವಂದು ವಿಭಾಗ ಒಂದೆಂತ ಹೆಚ್ಚು ತಾಕೆಗಳನ್ನೂ ತನ್ನ ಸೀಮೆಯಲ್ಲಿ ಹೂಂದಿರುತ್ತದೆ ಪತಿ ತಾಕೆಯು ನುರಿತ ಸಿಬ್ಬಂದಿ ಹಾಗೂ ಸ್ವಪೇರಿತ ವ್ಯಕ್ತಿಗಳನ್ನು ವಲಗೂಂಡಿರುತ್ತದೆ.

Looking at the HTML source, we see the meta links that help a web crawler identify the re-narration relationship between the original fire safety page and the contents of the blog for a target community [8].

```
<h2>Blog of Aravind</h2>
<h3>post saturday</h3>

<br />
<p id="kn"
foruri="http://www.ally.in/allypi/idea/firesafety.html:div1"
rec="lang:kn">ಅಗ್ನಿ ಅಮಕ ವಿಭಾಗ ಒಂದು ಸಾಮಾಜಿಕ ಅಥವಾ ವಯಕಿಕ ಸಂಸೆ ಅಗಿದು
ಇದು ಅಗ್ನಿ ನಂದಿಸುವ ಕಾರ್ಯವನ್ನು ಮಾಡುತ್ತದೆ. ಎಲ ಜೆಲೆಯ ಉರುಗಳು ಇ ಸಂಸೆಯನ್ನು
ವಲಗೂಂಡಿರುತವೆ. ಪ್ರತಿವೆಂದು ವಿಭಾಗ ಒಂದೆಕಿಂತ ಹೆಚ್ಚು ಅಕಗಳನ್ನೂ ತನ ಸೀಮೆಯಲ್ಲಿ
ಹೂಂದಿರುತದೆ. ಪ್ರತಿ ಅಕೆಯು ನುರಿತ ಸಿಬ್ಬಂದಿ ಹಾಗು ಸ್ವೇಚ್ಛೀರಿತ ವ್ರತಿಗಳನ್ನು
ವಲಗೂಂಡಿರುತದೆ.</p>
<br />
```

Traditionally, bloggers have been a dynamic network of *citizen journalists* who publish various interpretations and opinions on the Web - this work is frequently about relating interesting work by rewriting it for their readers. These days we see this activity more in smaller grain on Facebook and Twitter like Web 2.0 platforms. The association between the original page (source) and re-narrated page (target) is primarily in the author's and readers' mind with perhaps an additional link to the source page and/or some quotes. The explicit relations between items are typically not indicated. For example, a reference to a particular image, paragraph, or audio fragment are not indicated. Also, it would be hard to specify these meta relationships with out some assistance by "blogging" tools.

There are specific benefits for having such references in an explicit format rather than implicit or non-specific. For example, it is possible to contextualize the source content such as for re-narration and other community annotations i.e., for *propagation* of a specific purpose, a meta annotation of an intent.

3. Objectives and Conclusions

3.1 Current directions

A goal of Alipi is to enable localization and contextualization of laws and policy documents that concern the citizens of India, so that these documents become available on the mobile phones of the many print-impaired people. Towards this, we have authoring guidelines that document authors can use [4]. Then we have developed the re-narration model so that an effective process can be initiated via the communities of interest or through those who have a mandate towards such activity. Filters [4] help identify such communities of interest in certain context. For example, in the case of government documents that are put online, it may be natural for the authorities to announce the authorized re-narrators on their web-sites. This can become a directive to the Alipi's narration recommendation algorithm using which only the official translations or localization/contextualization are provided as choices to a user. The ReNarrationAct case study on alipi.janastu.org shows a document regarding the law related to minimum wages and using an Android based phone to demonstrates that the print-impaired community of domestic workers can now "browse" (listen-to/read-Kannada-version-of) this document using the Alipi toolbar on the Firefox browser.

3.2 Implementation and tools

Alipi model development is supported by prototype work that gave us the comfort of its feasibility over time. One of the tools is a Firefox browser extension that helps in selecting elements of a web-site that are of interest and assist in re-authoring the page on the client side. The changed content can then be blogged some where. The same extension can also help a person browsing the web in choosing alternative narrative when they are available. The Alipi toolbar is supported by a web-service that simulates a web-search index that responds to a query for re-narrations available for a web-site (an url). The alipi toolbar for Firefox and its avatar for an Android based mobile device are available as git-hub repositories from alipi.janastu.org.

3.3 Conclusions and Future Work

Accessibility provisions are all about transformation of information from one form to another which is often different from the original form of dissemination-intent. The re-narration web is about explicitly discussing and clarifying that this provisioning is fundamental and therefore needs to be reasoned and relaxed. The re-narration is also essential for the literate to follow something that is published in another context or a foreign language.

Print-impaired users are people able to use their vision and their hearing capabilities, but have difficulties accessing written text. A non-disabled user navigates within the web page and understands its structure instantly by relying on image connotations or paragraph titles for example. It is however frustrating for a print-impaired user to use assistive technologies such as content readers in order to understand the page structure: using an auditory description is not adapted to their needs since they can see and would rather rely on their vision than their hearing. Another barrier is the language. In fact, if the spoken/written language is not familiar to the user, it would not help them understand what is going on. Thus, the idea of Alipi accessibility guidelines is to allow a lay out a web-page's content in a certain way that allows print-impaired users to understand its structure by indicating connectedness of fragments in a page. Alipi tools aim to assist in sufficiently re-narrating the structure of the page so as to render it in a print-impairment friendly manner.

There are many aspects of the proposed approach that need to be pursued. Assessing trust of content authors requires a trust model for contributors. When there are more than one re-narrations a choice of which one to render must be made. Such choices can be based on author rating or user choice. Another option is to present alternative views for which a ranking among choices is required. Again, rating or user choice is relevant for this purpose. For user preferences friendship networks based on "like", "follow" or such relations could be utilized. Such friendship networks would need to be handled.

And last but not least, we need to study the various forms of narrations and what makes this popular and useful for delivery over the mobile phones and particular for the print-impaired.

Acknowledgements:

This work is partially funded by The Information Society Innovation Fund (ISIF.asia, 2010/2011) and the B.U. Research Fund (5079).

References

- [1] Web Accessibility Initiative, <http://w3.org/WAI> (accessed Nov. 2011)
- [2] TED - Ideas Worth Spreading, The TED Open Translation Project, <http://www.ted.com/OpenTranslationProject> (2008)
- [3] Mozilla Drumbeat, Mozilla Universal Subtitles, <https://www.drumbeat.org/en-US/projects/universal-subtitles> (2010)
- [4] T. B. Dinesh, AlipiReport - accessibility for the print impaired <http://j.mp/alipi2011> (Aug. 2011)
- [5] T. B. Dinesh, Alipi - God's Own Content. Authoring Guidelines, tools and other re-narration web case studies at <http://alipi.janastu.org> (2010)
- [6] Smyth, Kumar, Medhi and Toyoma, Where There's a Will There's a Way: Mobile Media Sharing in Urban India, <http://research.microsoft.com/pubs/122816/p753-smyth-mediasharing.pdf>
- [7] Silvia Masiero, Reframing Empowerment: A new Paradigm for ICT4D, iSCHANNEL 6(1): 6-13, (2011), www.lse.ac.uk/collections/informationSystems/iSChannel/iSChannel2011/iSChannel_Vol6_Iss
- [8] Choppella and Agarwal, Ontology services for identifying target communities; Draft 2011
- [9] T. B. Dinesh, Demonstration of Alipi re-narration Web idea, <http://a11y.in/a11ypi/idea/> (2011)
- [10] Social Semantic Web, In Wikipedia. Retrieved Nov 28 2011, http://en.wikipedia.org/wiki/Social_Semantic_Web
- [11] Tom Gruber, Collective knowledge systems: Where the Social Web meets the Semantic Web, Web Semantics: Science, Services and Agents on the World Wide Web, Volume 6, Issue 1, February 2008, Pages 4-13.