

TV and the Social Web in the NoTube Project

Libby Miller¹, Dan Brickley², Vicky Buser¹, Yves Raimond¹, and Members of the NoTube project³

¹BBC, UK

²Vrije Universiteit, Amsterdam

³<http://notube.tv>

February 11, 2010

This is a draft shorter version of the NoTube deliverable D7c.1. It may be updated in the future.

Three principles:

Conversation is king: content is just something to talk about.¹

The TV should be part of the Web, not a walled garden.

Connecting the TV to the Web doesn't have to mean showing the Web on the TV screen.

As part of the NoTube project we aim to show how these can be implemented using open APIs and formats, guided by practical scenarios, initially using a software media centre controlled by a handheld companion device.

The result will be software, APIs and formats that can enable contributions to the 'NoTube Network': a unified global network of annotations, ratings, comments, descriptions, related links and tags, attached to unique, re-referencable identifiers for programmes, segments, series, brands and channels, and that can provide new means for users to inform, educate and entertain themselves online.

Crucial to success will be the open APIs, formats and protocols we will use and create, so that the principles of an open, Web-connected TV can be applied to multiple environments, including watching on media centres, Web-based TV and mobile video.

1 The NoTube Network

Particular pieces of broadcast video content are not usually available to people in different geographical areas at the same time (generally they are at the mercy of various content-based agreements, scheduling times may be different, timezones different and so on). The same is true of many video on demand systems such as Hulu or Joost.

Nevertheless people in different geographical areas can and (we believe) should be able to create annotations ('User Generated Content') against these programmes. That is, they should all be able to refer to the same thing even if they cannot or do not want to watch it all at the same time.

Annotations could be as simple as 'I watched this', or could be tags, reviews, ratings, or anything else people want to say about pieces of media content. Although these kinds of functionalities exist currently in many TV set top boxes, video players and media centres, they usually lock the data away and it does not

¹<http://www.boingboing.net/2006/10/10/disney-exec-piracy-i.html>

flow back into the Web and so is lost and cannot be reused. There is no consistent way to control or access this information.

Below we show a plan for how that data can be made useful and reuseable by connecting the TV to the Web more closely.

Scenarios for Personalized TV and the Social Web

The NoTube project is based around three usecases. In this part of the NoTube Project, we will contribute to the ‘NoTube Network’ guided by scenarios from one usecase (‘TV and the Social Web’), showing how data might flow between the Web and the TV in both directions.

1.1 Scenario 1: Recommendations on my TV based on my web behaviour

Jana wants to see recommendations based on her social activity on her TV when she gets home at night. She talks a lot on twitter and facebook about what she watches in the context of her online social life, and so do her friends, and doesn’t see why she should have to explicitly tell any system what her preferences are. She wants to see recommendations clearly featured on the user interface of her media centre.

1.2 Scenario 2: Do you want to know more?

While watching on her TV Jana sometimes would like more information about a programme. She’d like to be able to mark a programme to come back to it later and find out more about it. She doesn’t necessarily want to have her laptop open all the time during this and neither does she want to interfere with the playing of the programme too much as she often watches TV with other people in the same room.

1.3 Scenario 3: Add to playlist

While at work, when he gets a link sent to him via Twitter or Facebook, Stephen would like to add programmes to his playlist on the media centre for watching later.

1.4 The Scenarios and the ‘NoTube Network’

These are scenarios that will only be of immediate relevance to a relatively unusual group of people at the current time: those who commonly watch TV on their media centre, and who watch a lot of on-demand media and who heavily use social networking sites. However it is possible to relate them to the majority of users as well, who are currently only consuming broadcast content.

In Scenario 2, *wanting to know something more about a programme and remembering it for later* is the key idea, a desire common to many TV watchers, regardless of the technologies they use. In Scenario 1, *filtering interesting programmes and bringing them to the user’s attention* is the most important aspect, and one becoming increasingly relevant to all TV watchers as channels proliferate. In Scenario 3, *saving items to watch later* is conceptually similar to recording programmes on video or DVD. Finally, using a setup that replicates at least the feeling of watching TV in a living room allows us to address the most common way of watching TV.

The APIs and data formats - the components that allow us to make the prototype but also specifically allow us to filter an EPG, export activity feeds, control a television and so on - are the pieces that can be reused in other user situations and can form part of the ‘NoTube Network’.

2 Our Assumptions

2.1 Roughly Replicate the Traditional TV Watching Format

Although TV-watching often takes place alone, nevertheless it is also a social activity, encompassing direct interaction with other individuals while either watching in the same place or separated by distance and communicating online. For the first prototypes we focus on cases where there may be someone else in the same room, and so the user doesn't necessarily want to change channel or otherwise disrupt the viewing experience on the TV screen.

2.2 Conversation is King

*"Conversation is king. Content is just something to talk about."*² Whether social activity is occurring online or offline it's a crucial part of TV, and once communication is online, there needs to be a way for people to refer to the content in emails, social networking sites, texts and so on. Therefore we require ways to refer accurately, uniquely and dereferencably to items of interest, including programmes, contributors, genres, series, brands.

2.3 Connecting the TV to the Web doesn't have to mean showing the Web on the TV screen

Flowing directly from the face-to-face social user assumption, especially coupled with the difficulties of data input commonly experienced with remote devices on a TV, comes the idea that the Web is a useful companion to TV rather than the Web being on the TV screen. Therefore in this first prototype we are interested in companion devices when watching the TV, used to access the web **and** get information from the TV.

2.4 The TV is Part of the Web, not a Walled Garden

The third scenario illustrates this nicely: if the user sees something on the Web they can't watch now but are still interested to watch later when they have more time, then why should their TV not understand a link? Similarly, the plethora of APIs and protocols that underlie the modern Web and help users store their data at their convenience and where they get most benefit from it could also be applied to TV.

3 Linked Data and TV

In the project, we use Semantic Web³ vocabularies and technologies where useful, particularly Linked Data. Linked Data is rather like putting multiple databases on the web and giving each item of interest a globally unique key that allows you to make links between different databases. In the Linked Data world, these keys are also dereferencable URLs, which can give more machine-readable information about themselves when fetched, such as what they are connected to, what sort of a thing they are and what properties or attributes they have. The items of interest can be anything - people, documents, places, pictures, videos, anything that can be identified.

3.1 Linked data and Content Filters

Using Linked Data enables interesting links to be made between items in a partially automated way. For example by combining MusicBrainz with DBpedia, which is derived from Wikipedia, we may be able to find out that a piece of music is related in an interesting way to another piece of music. Some of the links between the two pieces of music might be very uninteresting ('both artists were American'), but we might

²<http://www.boingboing.net/2006/10/10/disney-exec-piracy-i.html>

³<http://www.w3.org/Talks/WWW94Tim/>

find something interesting by linking up these datasets, (such as ‘both artists worked in Detroit in the 1960s’). This is a way to suggest content of interest to someone who has just listened to music by the first artist, or to provide more information of interest about the first artist (for example as part of the ‘Do you want to know more’ scenario). The Music Bore⁴ and dbrec⁵ are examples of using this technique without user profiles.

We believe that a profile for a user stating that she is interested in particular things (such as a particular kind of music) will allow useful personalised content filters to be created, by narrowing the types of linked data connections that should be followed for this user. Further, if the user profile can state in a machine-processible fashion the preferences of the user over multiple interests in context then a smart linked data filter should be able to order the suggestions in priority of interestingness.

One of the nice features of this approach is that cross-domain recommendations can be produced, for example a TV programme could be used to recommend related music, or related documents or books. Another useful feature is that this technique can provide contentful explanations of why the user might be interested in a piece of content, as opposed to the ‘black box’ explanations of Amazon-style collaborative filtering techniques (e.g. ‘You might like Y because other people who bought X also bought Y’).

One of the outputs of the NoTube project is planned to be research in applying complex profiles to linked data content filters of this kind.

3.2 User profiling from social activity data

Many organisations gather data on individuals, for example supermarket clubcards, Web search engines, and broadcast organisations. They tend to be interested in people matching statistical trends that they can use to target advertising or products to specific individuals, and in particular to target people with items they actually want to purchase or use. Here we are interested in something similar: to suggest content to people that they are actually interested in. The means we have chosen is to harvest the data about individuals on the Web **with their explicit consent** to create machine-readable profiles that they control. The part of NoTube implementing this is the ‘Beancounter’, which enables users to securely add social networks they are members of to generate a profile and edit the result, so they get suggestions about the things they are interested in.

3.3 BBC Programmes Linked Data

As part of its website the BBC aims to produce a URL for each programme created, available in HTML and RDF formats, and linked to other items of interest, such as series, genre and contributors, using the BBC Programmes Ontology⁶ and in some cases MusicBrainz and DBpedia identifiers. In doing so it forms part of the Linked Data Cloud⁷. This is separate to its iPlayer catchup service, although where video of the programme is available it will be linked from or embedded in the programme page.

This makes for a useful testbed for these ideas: without adding to the BBC data, we can easily create bookmarking systems, because there has a URL for most recent programmes - for the BBC, TV already links to the Web. In addition, because the BBC Programmes Ontology is an FRBR-based vocabulary, it distinguishes between items of interest at the ‘Brand’ (e.g. Dr Who), ‘Series’ (e.g. Dr Who Series 8), ‘Programme’ (Dr Who Series 8, episode 5), ‘Version’ (Dr Who Series 8, episode 5 with signing). It also describes when it was broadcast, the genre it is placed in. and in some cases who the contributors were. This means that extremely simple recommendations such as ‘you might like this episode of Dr Who because you liked the series Dr Who Series 8’ can be made using this data alone, using the identifiers and the broadcast schedule.

⁴http://www.bbc.co.uk/blogs/radiolabs/2009/07/the_music_bore.shtml

⁵<http://dbrec.net/>

⁶<http://www.bbc.co.uk/ontologies/programmes/>

⁷<http://linkeddata.org/>

3.4 Data Enhancement

For more interesting and for non-BBC content suggestion and filtering there is substantial work to be done in

- Concept identification: creating new links between items of interest and other vocabularies, using natural language processing or ontology matching
- Matching different identifiers for the same item. In the example of the database on the Web, where there are different unique keys coming from different databases which refer to the same thing, we need to find them and create links between them. An example might be a BBC programme URL for Murder She wrote, Series 12, episode 4 and the equivalent IMDB URL.
- Matching of schedule and TVAnytime broadcast identifiers and Web programme identifiers
- Matching user channel or programme availability to any given user, for example using schedule data and location data

while for the bookmarking and more information scenarios, there is a need for

- Generated or existing URLs for channels and programmes

4 Tasks and Implementation

The prototype implementation will use a companion device (an iPhone in this case) to control a software media centre (MythTV) and also receive information from the media centre to allow the user to plan viewing without disturbing others (e.g. by browsing available programmes on the device). The user can also find out more information about a programme using the device, or access existing social networks from it in order to talk about the programme.

We are currently at the proof-of-concept stage of implementation.

4.1 Scenario 1: View Recommendations in an EPG

The screenshots in Figures 1 and 2 show recommendations appearing as a changed UI in MythTV frontend media centre and its web-based navigator. This shows that the UI can be altered, and in building it we discovered the fields we needed for the output of the recommendations engine: a consistent channel identifier and an exact date-time value.

4.2 Scenario 2: Bookmark and Tag a Programme from TV to the Web

The screenshots in Figures 3 and 4 show our proof-of-concept we used a text-based XMPP interface via MythTV to create bookmarks in the bookmarking site delicious.com. XMPP is used where http would also work because we want to explore XMPPs built-in ‘friending’ mechanism in later iterations for security and privacy. The user would not be expected to type the commands in manually in the finished interface.

4.3 Future Work

The implementation is ongoing, and we hope to show functioning and usable implementations of all three scenarios by May 2010. Later we will place particular emphasis on its **security** and **privacy** aspects. In particular:

- Securely pairing devices to media centres
- Securely accessing private data from social networking sites
- User experience of privacy in the Beancounter and elsewhere

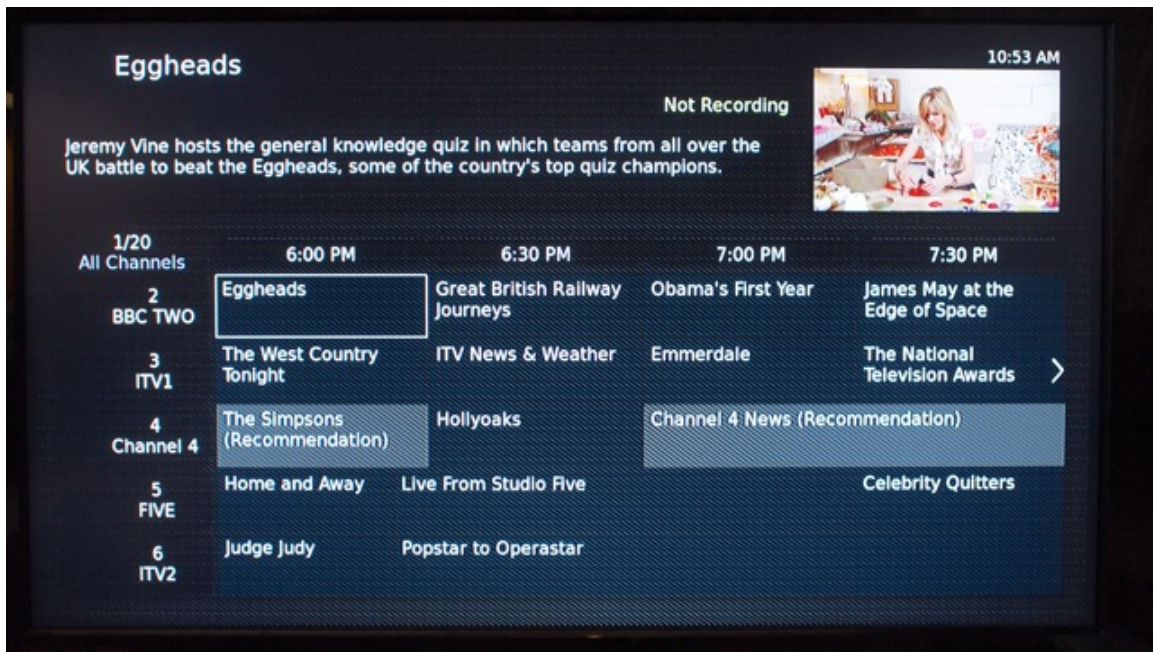


Figure 1: Recommendations appearing in MythTV frontend EPG

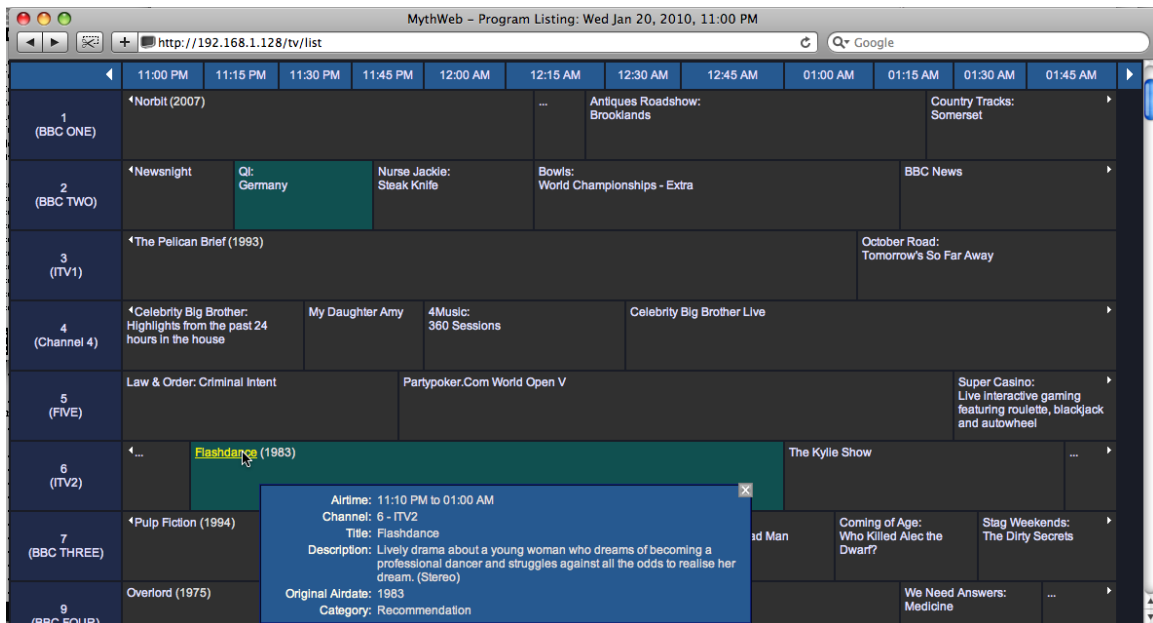


Figure 2: Recommendations appearing in MythTV web frontend (MythWeb)



Figure 3: Jabber bot communication with MythTV and Delicious

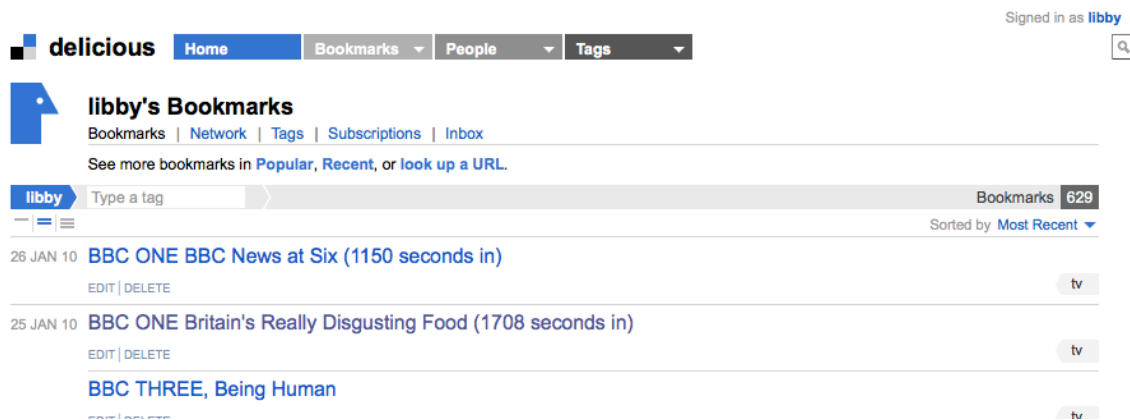


Figure 4: TV bookmarks appearing in Delicious