

Tvheadend - Feature #1368

The default language should be used when grabbing XMLTV EPG

2012-10-28 11:22 - Sébastien Aubry

Status:	Need feedback	Start date:	2012-10-28
Priority:	Normal	Due date:	
Assignee:	Adam Sutton	% Done:	0%
Category:	EPG - Grabbers	Estimated time:	0.00 hour
Target version:			
Description			
Hi, Since I recently switched to Tvheadend 3.2 using the awesome PPA , I noticed that the XMLTV EPG grabbing works differently. My XMLTV EPG, tv_grab_fr, looks like this:			
<pre><programme start="20121028224000 +0100" stop="20121028232500 +0100" channel="C1.telepoche.com"> <title>Les Experts : Manhattan</title> <title lang="en">CSI: NY</title> <sub-title>La cité des rêves</sub-title></pre>			
In Tvheadend 2.x, the <title> element was used, meaning that the title language in the Tvheadend web interface (FR) was consistent with the grabber language (tv_grab_fr). Since Tvheadend 3.x, the <title lang="en"> (or the last <title> element?) is used, meaning that I only see "CSI: NY" in the web interface. In the Configuration > General tab, I selected "French" as my default language.			

History

#1 - 2012-10-30 13:58 - Adam Sutton

- Category set to EPG - Grabbers
- Status changed from New to Need feedback
- Assignee set to Adam Sutton
- Target version set to 3.4

Your interpretation is incorrect, though I can see how you came to this conclusion. And possibly there needs to be some workaround put in place.

TVH actually parses and stores ALL language variants provided in the XMLTV file (as of 3.2) and these are stored according to the language string. However if NO language is provided it will be stored under the lang code "und" (undefined), since that's exactly what it is.

The client (web, xbmc, showtime, etc...) can then request which language it wants (or you can configure default in TVH should none be requested). However its unlikely a client will be trying to request "und" and it will almost certainly be returned AFTER "eng" in the event that it has to fallback to picking a language at random.

Can you provide me a copy of your full XMLTV output just in case there is a default value in there somewhere, otherwise I'll try and think of some way of working around this issue. Possibly by assuming that no lang spec = TVH default (if configured).

Adam

#2 - 2012-10-30 14:59 - Sébastien Aubry

Hi Andy,
I will send you the full output tonight, but I am pretty sure that these were the only two <title> tags.

In this case (and probably in some other cases), if no language is provided, then the default grabber language should be applied ("fr" here). Shouldn't that be the case for all the "tv_grab_<language_code>" default XMLTV grabbers ?

Another approach would be to have an "other" or "default" lang code in TvHeadend, which would be used when the client (web for me) requests an unlisted language.

Of course if needed we will update the tv_grab_fr script to force this "lang" attribute.

#3 - 2012-10-30 15:00 - Sébastien Aubry

Sorry for my mistake in your first name. I thought Adam but typed Andy. It would be so nice to allow the edition of the posts during a few minutes!

#4 - 2012-10-30 15:46 - Adam Sutton

I agree that something better could probably be done to work around the XMLTV scripts not providing language codes for the default languages. And in your specific case sure you can assume that the first bit after tv_grab_** is the language code.

But what about tv_grab_uk../tv_grab_nz.. UK/NZ are not language codes they are country codes, sure I can assume they mean something (en in this case probably?) but really if a script is providing multi-lang support it should either explicitly define the default language somewhere or should provide a language for each string.

And those 2 are just the first 2 that come to mind immediately, I'm sure there are others.

However it's possible that I could assume that if no language code is provided TVH should use its own default language code (where configured). That would overcome the problem, but would require users to correctly set the config etc..

Adam

#5 - 2012-10-30 16:46 - Sébastien Aubry

You are right: if no language code is provided, TVH should use its own default language code, if there is not any other <title> attribute for this language.

But let us wait and check tonight if there is any other place in the XMLTV file where the default language is indicated. I can also act on the source and ask them to update the grabber so that it always fills the language attribute.

#6 - 2012-10-30 17:00 - Sébastien Aubry

I just looked at the [XMLTV DTD](#) and I could not find any default language indication. Another possibility, perhaps the cleanest, would be to configure this default language in the "EPG Grabber" tab.

We could also use the [XmltvWorldDomination](#) page to get a list of the default language values for the default XMLTV grabbers, which could be a constant in TvHeadend.

#7 - 2012-10-30 22:45 - Sébastien Aubry

Here is the beginning of my XMLTV file. I don't post the end of the file because it only consists in <programme> elements.

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE tv SYSTEM "xmltv.dtd">

<tv source-info-url="http://telepoche.guidetele.com/" source-data-url="http://telepoche.guidetele.com/" generator-info-name="XMLTV" generator-info-url="http://xmltv.org/">
  <channel id="C1.telepoche.com">
    <display-name>TF1</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/tf1.gif" />
  </channel>
  <channel id="C2.telepoche.com">
    <display-name>France 2</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/france2.gif" />
  </channel>
  <channel id="C3.telepoche.com">
    <display-name>France 3</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/france3.gif" />
  </channel>
  <channel id="C4.telepoche.com">
    <display-name>Canal +</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/canalplus.gif" />
  </channel>
  <channel id="C5.telepoche.com">
    <display-name>Arte</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/arte.gif" />
  </channel>
  <channel id="C6.telepoche.com">
    <display-name>M6</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/m6_logo.gif" />
  </channel>
  <channel id="C7.telepoche.com">
    <display-name>France 5</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/france5.gif" />
  </channel>
  <channel id="C122.telepoche.com">
    <display-name>i-télé</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/i_tele.gif" />
  </channel>
  <channel id="C167.telepoche.com">
    <display-name>Direct 8</display-name>
  </channel>
</tv>
```

```

    <icon src="http://telepoche.guidetele.com/medias/chaines/new-logo.png" />
</channel>
<channel id="C168.telepoche.com">
    <display-name>NT1</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/nt1.gif" />
</channel>
<channel id="C169.telepoche.com">
    <display-name>NRJ 12</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/nrj_12.gif" />
</channel>
<channel id="C170.telepoche.com">
    <display-name>Public Sénat - LCP AN</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/parlement.gif" />
</channel>
<channel id="C193.telepoche.com">
    <display-name>Direct Star</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/direct_star.jpg" />
</channel>
<channel id="C194.telepoche.com">
    <display-name>BFM TV</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/bfm_tv.jpg" />
</channel>
<channel id="C195.telepoche.com">
    <display-name>Gulli</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/gulli.png" />
</channel>
<channel id="C28.telepoche.com">
    <display-name>France 4</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/france4.gif" />
</channel>
<channel id="C38.telepoche.com">
    <display-name>W9</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/w9.png" />
</channel>
<channel id="C49.telepoche.com">
    <display-name>France Ô</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/franceo.gif" />
</channel>
<channel id="C9.telepoche.com">
    <display-name>TMC</display-name>
    <icon src="http://telepoche.guidetele.com/medias/chaines/new_tmc.png" />
</channel>
<programme start="20121028035000 +0100" stop="20121028044500 +0100" channel="C1.telepoche.com">
    <title>Patrimoine immatériel : Chef d'oeuvre de l'humanité</title>
    <sub-title>Jouer le kutiyattam avec le maître</sub-title>
    <credits>
        <director>Jérôme Segur</director>
    </credits>
    <date>2010</date>
    <category lang="fr">série/feuilleton</category>
    <category lang="fr">documentaire</category>
    <country lang="fr">France</country>
    <country>FR</country>
</programme>
<programme start="20121028044500 +0100" stop="20121028050000 +0100" channel="C1.telepoche.com">
    <title>Musique</title>
    <category lang="fr">musique</category>
    <category lang="fr">variétés</category>
    <country lang="fr">France</country>
    <country>FR</country>
    <video>
        <colour>yes</colour>
    </video>
</programme>
<programme start="20121028050000 +0100" stop="20121028052500 +0100" channel="C1.telepoche.com">
    <title>Histoires naturelles</title>
    <sub-title>Fées et gens en Brocéliande</sub-title>
    <category lang="fr">documentaire</category>
    <category lang="fr">découverte</category>
</programme>
[...]
```

#8 - 2012-10-31 11:08 - Adam Sutton

Yeah looks like a bad mix of using lang="" and not. While it may be possible to get the upstream devs to update the script and that will solve the problem, I think that it might also be worth my adding the mod I suggested.

I will look into doing this.

Adam

#9 - 2013-04-04 23:38 - Adam Sutton

- *Target version deleted (3.4)*