

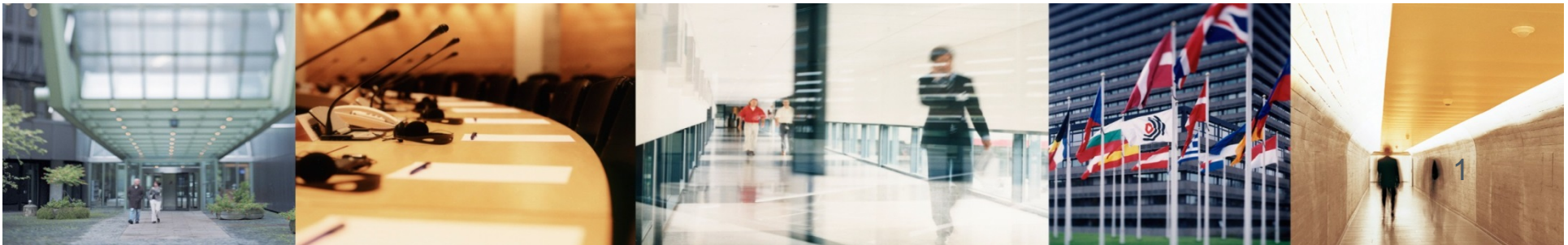


Europäisches
Patentamt
European
Patent Office
Office européen
des brevets

The EPO RAW DATA AREA REST interface

- automating your access to EPO raw data -

Patent Data Services, Vienna, September 2010



REST interface is for automated data retrieval

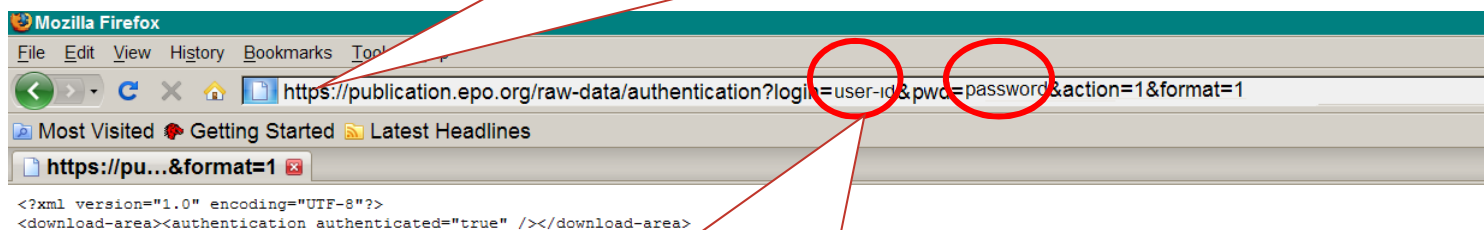
- In the following slides you will see examples of the use of the EPO RAW DATA Area REST interface via a browser.
- These examples follow a manual procedure and are for **instruction purposes only**.
- In practice it is expected that the use of the REST interface will be nearly exclusively limited to automated procedures (dedicated software, scripts etc).
- While going through these manual examples, the logic of the corresponding automated procedures will become apparent to users.

Log-in

Insert following link into your browser's address bar:

`https://publication.epo.org/raw-data/authentication?login=YOUR-USER-ID&pwd=YOUR-PASSWORD&action=1&format=1`

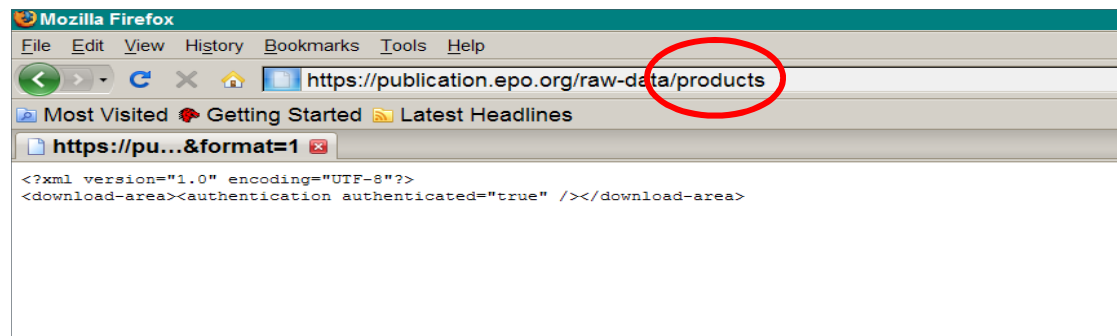
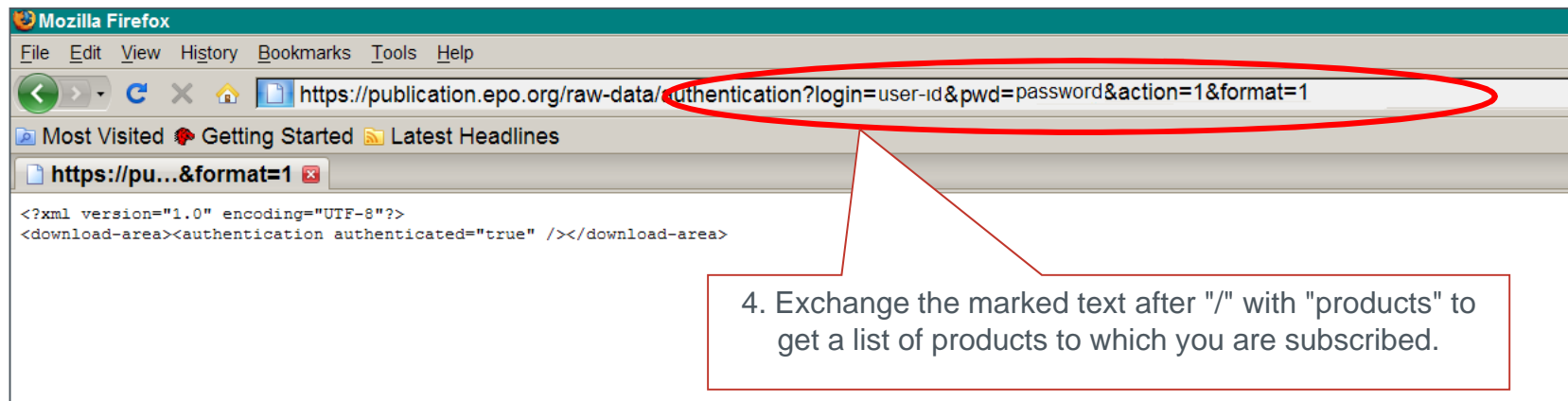
1. Copy the above link into your browser and exchange text in red with your user ID and your password



2. Exchange "YOUR-USER-ID" and "YOUR-PASSWORD" with your user-id and password as provided by Patent Data Services

3. Check that this text appears.
If yes, you are logged in

Now you should be logged in, i.e. A session Cookie is set, and you will be allowed to go to the next URL in the REST interface.



The list of all products that you are subscribed to will be displayed:

Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://publication.epo.org/raw-data/products

Most Visited Getting Started Latest Headlines

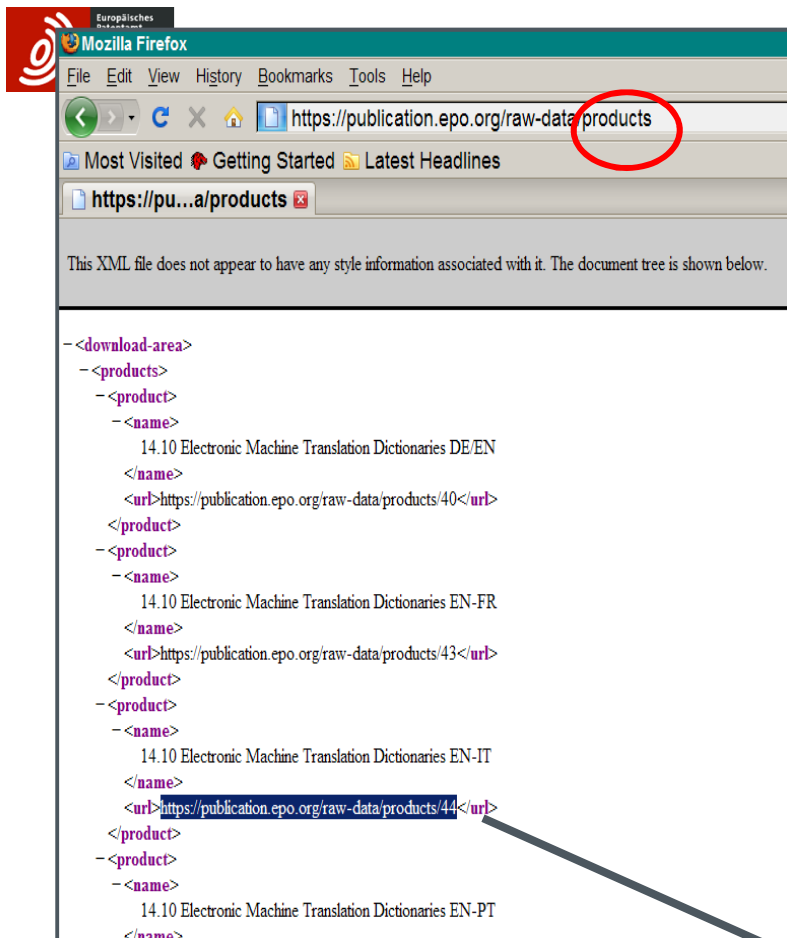
https://pu...a/products

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```

- <download-area>
- <products>
- <product>
- <name>
14.10 Electronic Machine Translation Dictionaries DE/EN
</name>
<url>https://publication.epo.org/raw-data/products/40</url>
</product>
- <product>
- <name>
14.10 Electronic Machine Translation Dictionaries EN-FR
</name>
<url>https://publication.epo.org/raw-data/products/43</url>
</product>
- <product>
- <name>
14.10 Electronic Machine Translation Dictionaries EN-IT
</name>
<url>https://publication.epo.org/raw-data/products/44</url>
</product>
- <product>
- <name>
14.10 Electronic Machine Translation Dictionaries EN-PT
</name>
<url>https://publication.epo.org/raw-data/products/45</url>
</product>
- <product>
- <name>
14.10 Electronic Machine Translation Dictionaries ES/EN
</name>
<url>https://publication.epo.org/raw-data/products/41</url>
</product>
- <product>
<name>14.11 INPADOC worldwide legal status data</name>
<url>https://publication.epo.org/raw-data/products/31</url>
</product>
- <product>
<name>14.12 EP-A & EP-B combined</name>
<url>https://publication.epo.org/raw-data/products/29</url>
</product>
- <product>
<name>14.18 WO Full Text database TEST FILES NEW FORMAT</name>
<url>https://publication.epo.org/raw-data/products/39</url>
</product>
- <product>

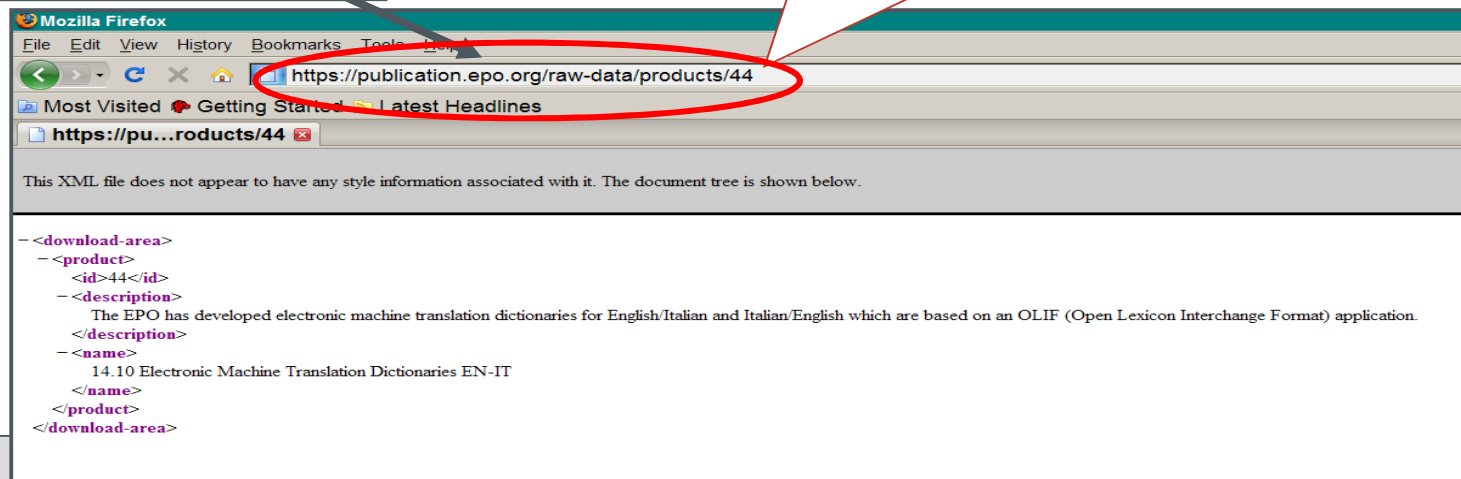
```



This is normally done programmatically on the client side.

If you go to the URL of a given product, a description will be returned.

5. Choose the product you are interested in and copy the URL in the browser, e.g.:



Here is a list of URLs giving the location of the edition(s) of the given product

6. Add "/editions" here

```

- <download-area>
- <editions>
  - <edition>
    <version>EN/IT</version>
    - <url>
      https://publication.epo.org/raw-data/products/44/editions490
    </url>
  </edition>
  - <edition>
    <version>IT/EN</version>
    - <url>
      https://publication.epo.org/raw-data/products/44/editions491
    </url>
  </edition>
</editions>
</download-area>
  
```

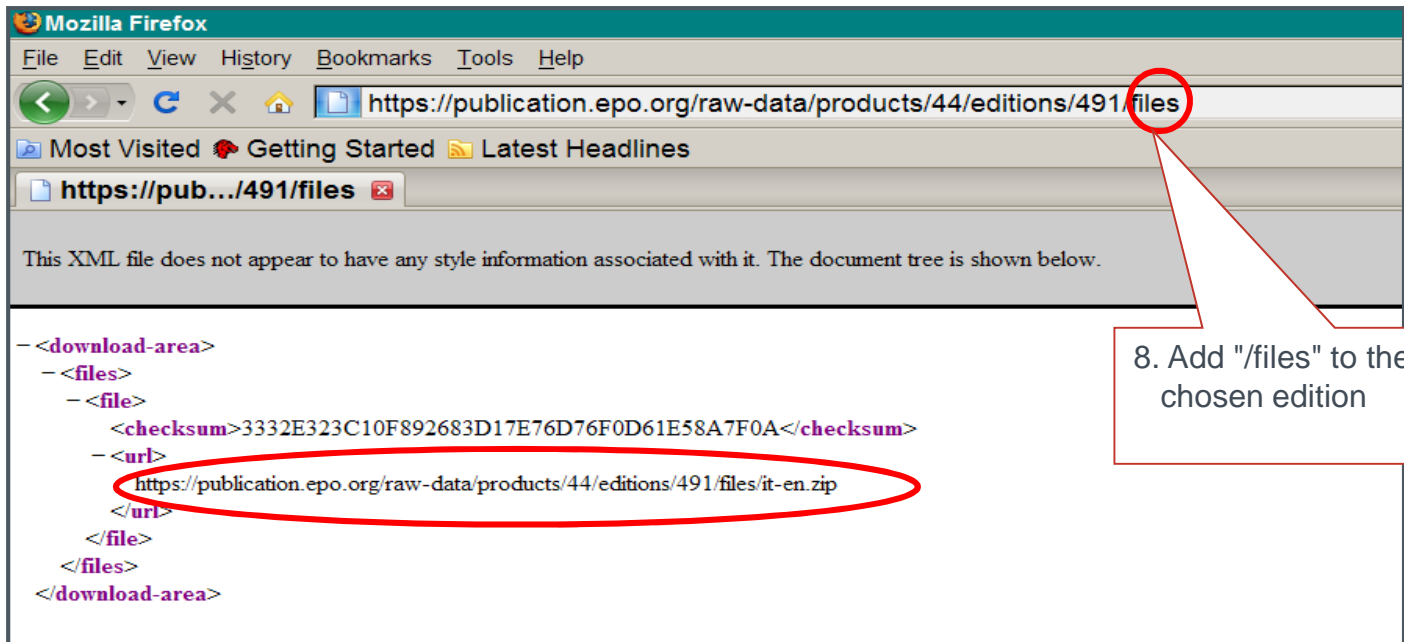
7. Add the number of the edition, e.g. "/491" here

7. Add the number of the edition, e.g. "/491" here

```

- <download-area>
- <edition>
  <id>491</id>
  <expiryDateTime>None</expiryDateTime>
  <publicationDateTime>23/03/2010 14:00:00</publicationDateTime>
  <version>IT/EN</version>
</edition>
</download-area>
  
```

Following the URL takes you to a list of URLs giving the location of file(s) associated with the selected edition. - e.g.:



The screenshot shows a Mozilla Firefox browser window. The address bar contains the URL `https://publication.epo.org/raw-data/products/44/editions/491/files`, with the word `files` circled in red. Below the address bar, a message states: "This XML file does not appear to have any style information associated with it. The document tree is shown below." The XML document tree is displayed, showing a root element `<download-area>` containing a `<files>` element. Inside `<files>`, there is a `<file>` element with attributes `<checksum>` and `<url>`. The `<url>` attribute value is `https://publication.epo.org/raw-data/products/44/editions/491/files/it-en.zip`, which is circled in red. A red callout box points to the `files` part of the URL in the address bar, containing the text: "8. Add '/files' to the URL for the chosen edition".

Each file also has a SHA-1 hash value given here too.

It is recommended that users make use of this SHA-1 value, e.g. for identifying data files that have been added / modified since the user's last visit to the Raw Data area via the REST interface.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
-<download-area>
- <files>
- <file>
  <checksum>3332E323C10F892683D17E76D76F0D61E58A7F0A</checksum>
  <url>
https://publication.epo.org/raw-data/products/44/editions/491/files/it-en.zip
  </url>
</file>
</files>
</download-area>
```

- Fetch "https://publication.ep..." in new tab
- Fetch "https://publication.ep..." in current tab
- Fetch "https://publication.ep..." in new window
- Copy**
- Select All
- Search Google for "https://publica..."
- View Selection Source
- Translate "https://publication..."

9. Copy the link into your browser to download the file

Mozilla Firefox

File Edit View History Bookmarks Tools Help

<https://publication.epo.org/raw-data/products/44/editions/491/files/it-en.zip>

Most Visited Getting Started Latest Headlines

<https://pub.../491/files>

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
-<download-area>
- <files>
- <file>
  <checksum>3332E323C10F892683D17E76D76F0D61E58A7F0A</checksum>
  <url>
https://publication.epo.org/raw-data/products/44/editions/491/files/it-en.zip
  </url>
</file>
</files>
</download-area>
```

Opening it-en.zip

You have chosen to open

it-en.zip

which is a: WinZip File
from: <https://publication.epo.org>

What should Firefox do with this file?

☐ Open with WinZip (default)

☒ **Save File**

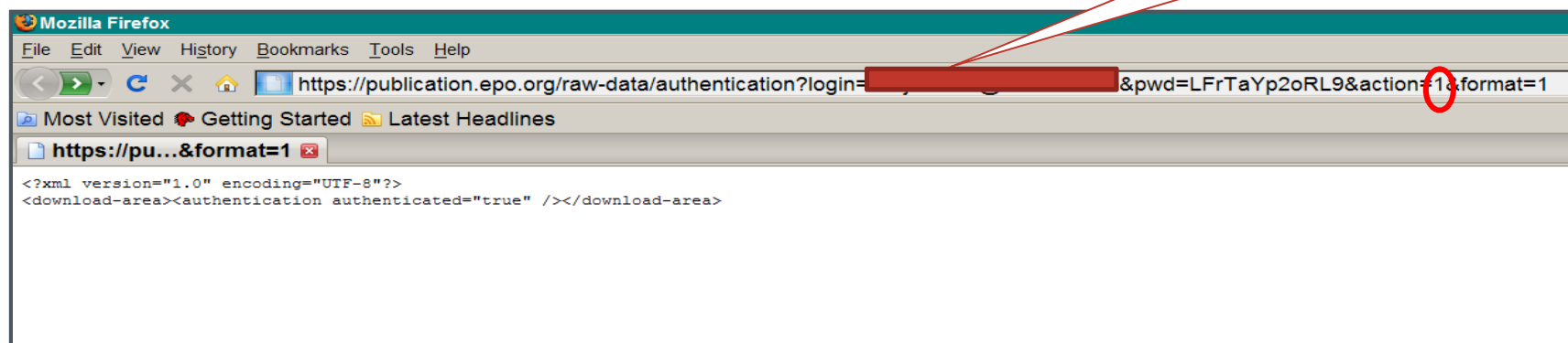
☐ Do this automatically for files like this from now on.

OK Cancel

Log-out

This should be done at the end of the process described below:

1. Copy the link from slide 2 (1 step) into your browser



2. Change the action parameter value from "1" to "0"

