
Visualization of International Help Data - Backend Application Data Extractor

Lukáš Dolniak

2017-05-31

Since the IATI organisation provides only very unreliable alpha version of its API, our team decided that it would be more convenient to create custom built application to download and extract needed data. My task was to create .Net based backend data analyzer. First of all I created a static utility class that contains almost all the methods to fetch and parse the data.

In this class I developed a method `DownloadWebpages` that sends request to IATI servers with search query for UNDP data. Then it saves it's response to the html file on a local drive in created directory. `ParseWebpages` method uses `HtmlAgilityPack` library to extract data from html files with Xpath expressions to navigate through the document. The method finds specific anchor tag elements with href links that point to Xml data locations in previously downloaded webpages. These are saved in list collection and are passed to `ParseXml` method that saves them in newly created directory.

In order to create usable Xml document for frontend application we needed to obtain data about population ISO number country codes. These are provided in csv files in `csvFolder` directory. Methods `CountriesCsvToXml` and `CountriesCodesCsvToXml` transform these files to Xml using .Net LINQ expressions. These Xml files are then connected into one file with `XmlConnector` method.

Last method that I implemented in `DataExtractor` class was `FinalXmlToJson`. It transforms the final Xml data into Json file with specific form needed for frontend to work effectively.

I also created form for user interface with four buttons. After clicking on any of them it calls its own `BackGround` worker object to perform specific tasks without the interface getting stuck. After processing every file (whether it is webpage downloaded or parsed Xml) its name is written to list box window.