# Specifications and development of an ELI Pillar IV RSS generator

*“In the discussion with the ELI Taskforce members it has become clear that not all Member States will be able to provide the envisaged combination of the sitemap for the initial harvesting of ELI data and the RSS feed for the subsequent harvesting of updates. Proposal is to provide a central ELI RSS generator that will allow to generate an RSS feed with the updates on the basis of a submitted sitemap. Aim is to keep the central Pillar IV harvesting solution as simple as possible. Task is to write the specifications of the tool and to develop the tool itself.”*

## Scope

We propose to specify and develop a sitemap to RSS (or Atom[[1]](#footnote-2)) feed generator that can be deployed locally by ELI providers.

The provided description mentions a “central RSS generator”, implying some central service hosted by the Office of Publications (similar to the ELI validator). We propose instead to develop a tool that can be deployed locally by ELI providers on their server to periodically generate an RSS (or Atom) feed from a local sitemap file. This has the following advantages :

1. more simple to develop (command-line tool vs. web application)
2. no maintenance cost for OP for a web application
3. no dependency for the ELI providers to a remote service hosted by OP

Our proposal goes one step further : instead of only generating the RSS/Atom feed from the sitemap, we even propose - as an option - to generate both the sitemap and the RSS/Atom feed from an input flat tabular file containing the URI and date of modification of each ELI. This would handle the split of sitemap in multiple files.

This will allow to use the tool in 2 modes :

1. ELI providers can take care of their sitemap generation, and the tool can handle the generation of the RSS/Atom feed
2. Or ELI providers can simply provide a tabular list of their ELI + modification date, and the tool can handle the generation of the sitemap and the RSS/Atom feed

The tool will be developed as an open-source tool, with code hosted on a public Git repository.

The tool will be based on open-source libraries for Sitemap parsing, sitemap generation, and RSS/Atom feed generation.

## Deliverables

1. Specifications for the tool functionalities
2. Tool developed as a command-line tool in Java that can :
   1. take as an input a sitemap and sitemap index, and generate the corresponding RSS/Atom feed
   2. be run periodically using “crontab” to regenerate the RSS/Atom feed on a daily basis (this implies that the sitemap is updated on the same pace)
3. Documentation on how to deploy and setup the tool
4. Optionally, an additional functionality that takes as an input a tabular list of ELI + date of modification, and outputs both the sitemap (including sitemap index) and the RSS/Atom feed

## Tasks

|  |  |
| --- | --- |
| **Tasks** | **Number of man.days** |
| Write the tool specifications, including meetings | 3 |
| Develop the tool, with the functionality to convert a sitemap to RSS/Atom | 5 |
| Write the documentation for the tool deployment by ELI providers | 1 |
| Optionally, add the functionality to convert a flat tabular file to both sitemap and RSS/Atom | 4 |

## Planning

The work can start after the ELI Pillar IV Protocol has been formally validated by the ELI Taskforce. The expected start date is 1st of February 2022.

The expected duration is 3 months (until 30th April 2022); the duration is subject to Eur-Lex team availability.

1. At the time of writing, it has not been validated whether the ELI Pillar IV Protocol will use RSS or Atom. [↑](#footnote-ref-2)