

# Statistics Finland's file-based px-cube publishing process

Juhani Kauppo

27th of October 2022

# Aims of the publishing process

- Ensure that no part of the publication (the analysis, graphs, tables or the aggregated px-data) is made public prematurely
- Ensure that everything is published in the correct order (graphs/tables depend on the aggregated px-data) and in its entirety
- Minimize the delay between the planned and actual time of publishing

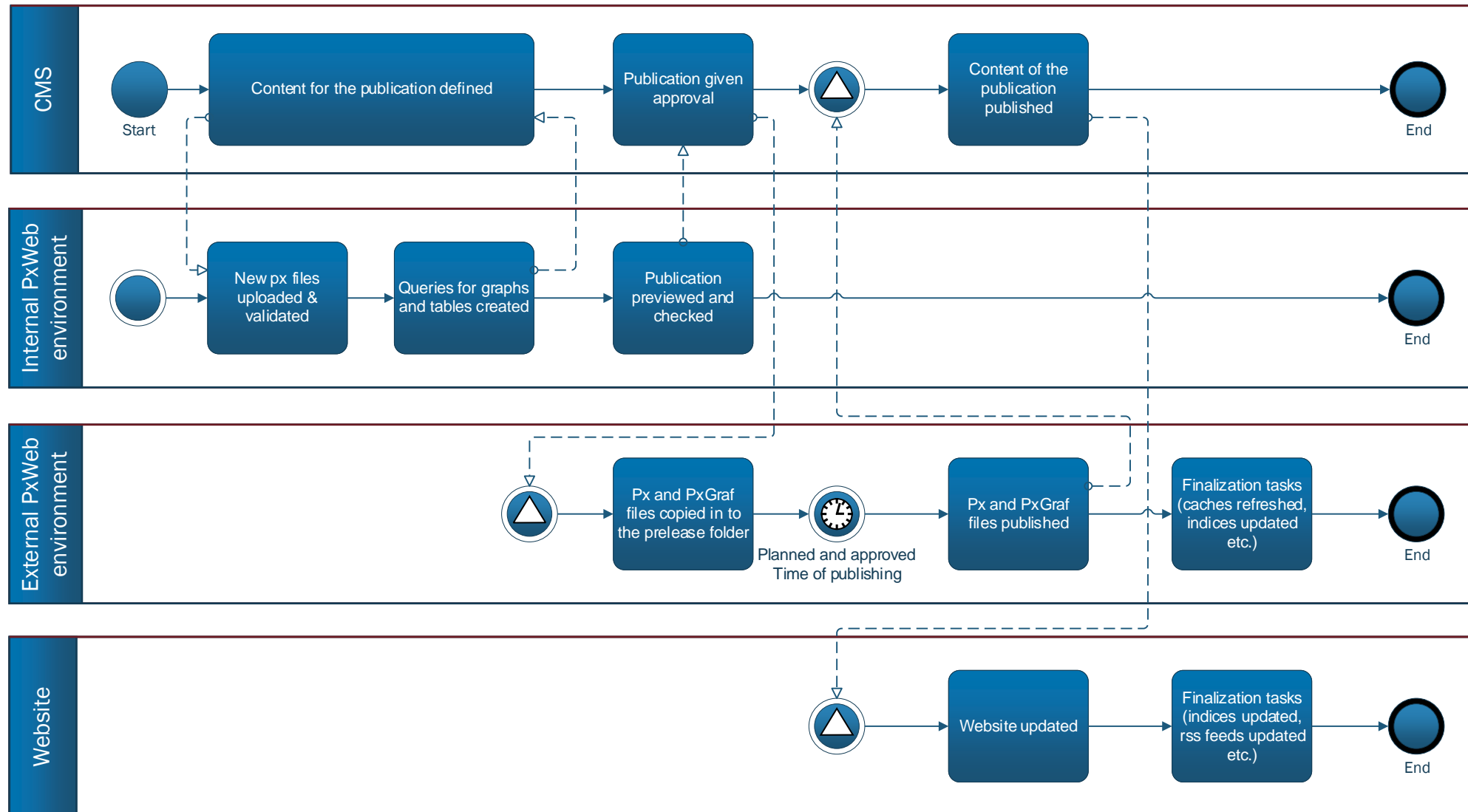
# Main steps in the process

1. Creating the content for the publication
2. Preview and validation
3. Confirming the time and date & approval for publishing
4. Pre-release of the px-content
5. Actual release of the px-content
6. Release of the rest of the publication
7. Finalizing tasks

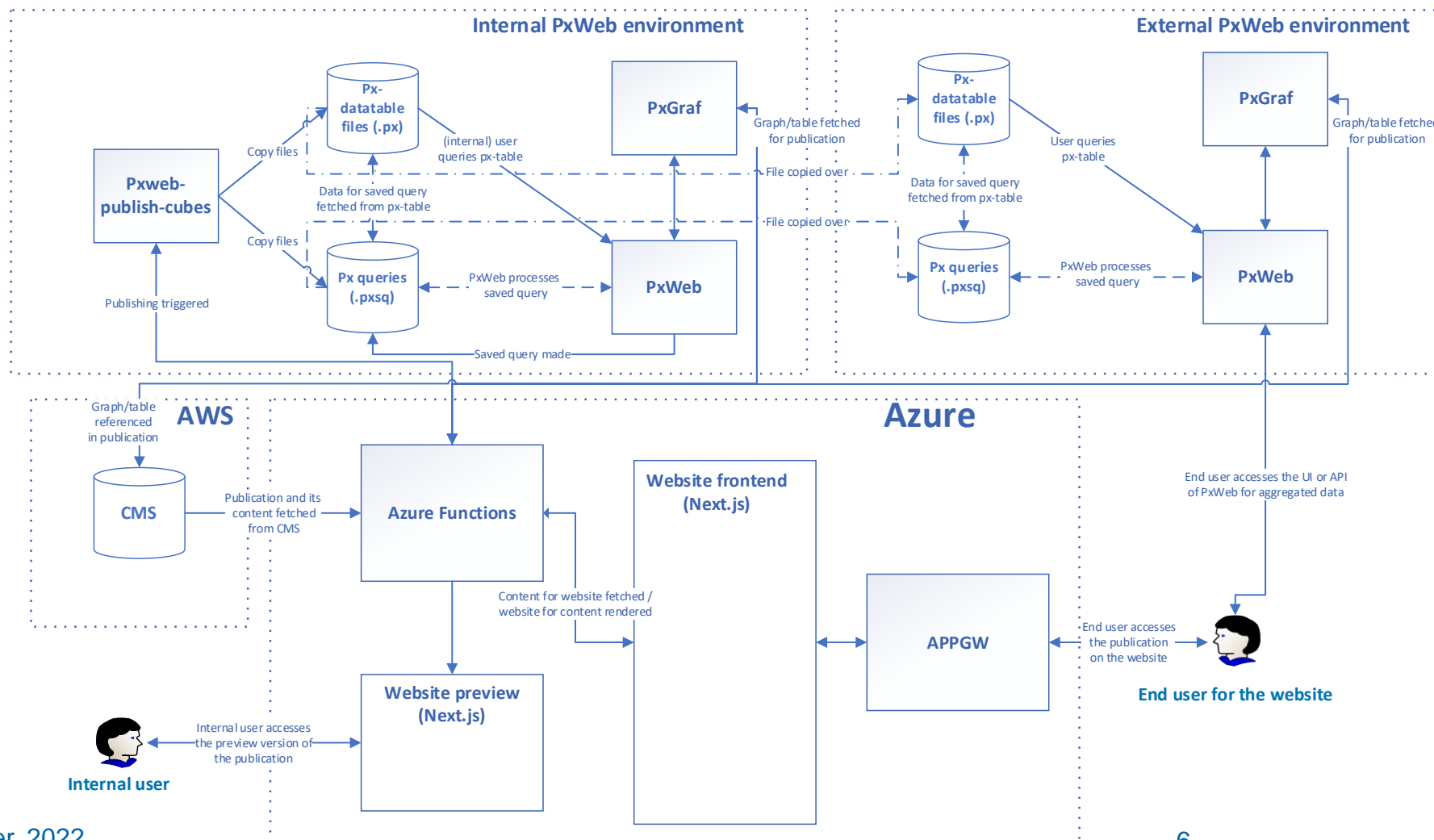
# Important systems and components

- Content management system
- Preview environment
- Website frontend
- Internal PxWeb server + PxGraf solution
- External PxWeb server + PxGraf solution
- Publishing process management service
- "Publish cubes" app

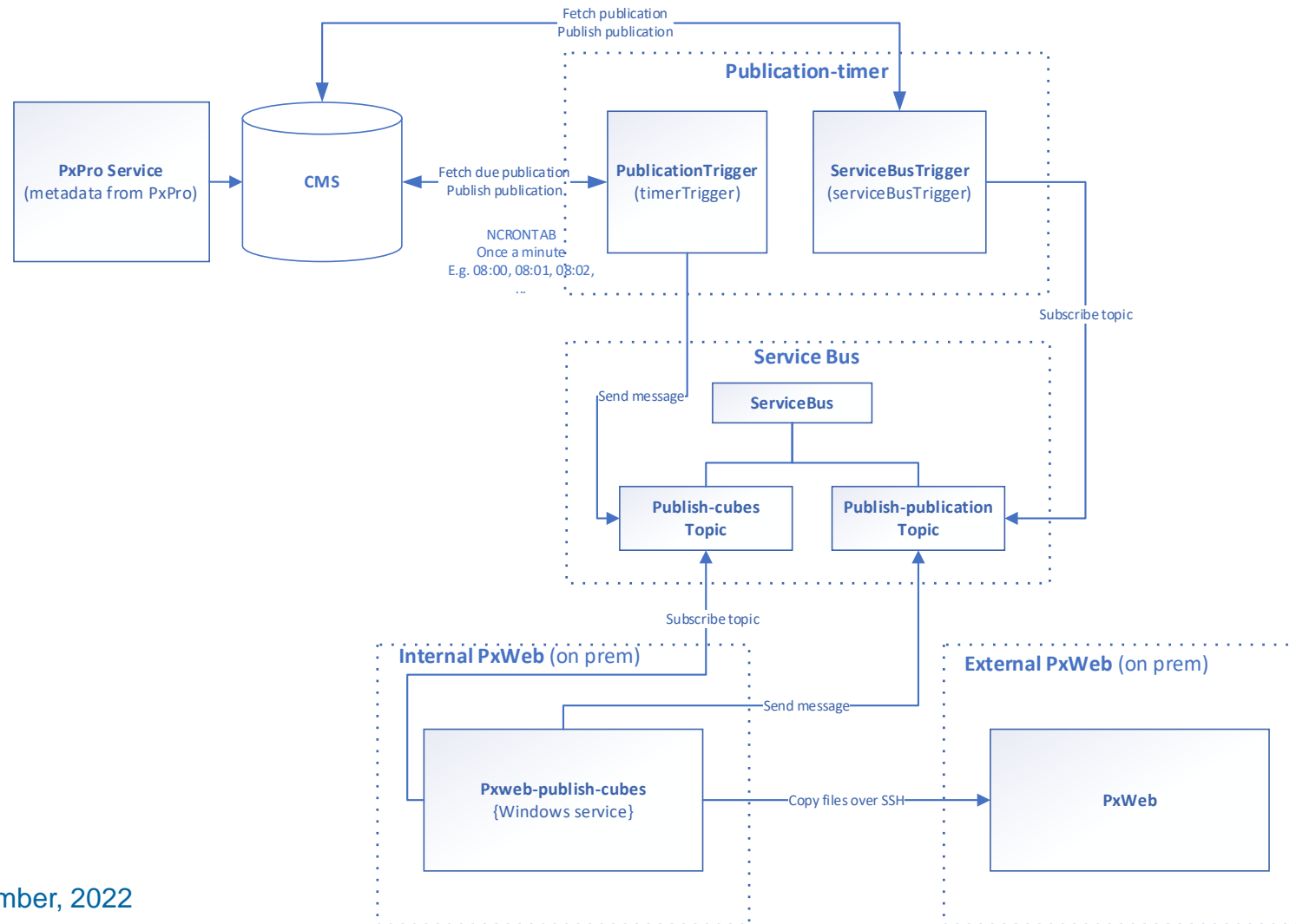
# The publishing process



# Technical architecture (Environments and components)



# Technical architecture (Process management)



# Pros of using px-files

- Updating PxWeb is relatively simple
  - Just copying files
- Archiving the content from each update/release is similarly simple
- Mass export of data is simple
  - e.g. a user can just download the whole file



# Cons of using px-files

- Harder to limit the size of the update
  - The whole .px file is replaced every time
  - Files can get large over time and copying and uploading is not instant
- Problems with locked files
- Indexing is probably less efficient