EUROPEAN LANGUAGE GRID

META NET
META FORUM 2021

EUROPEAN LANGUAGE GRID

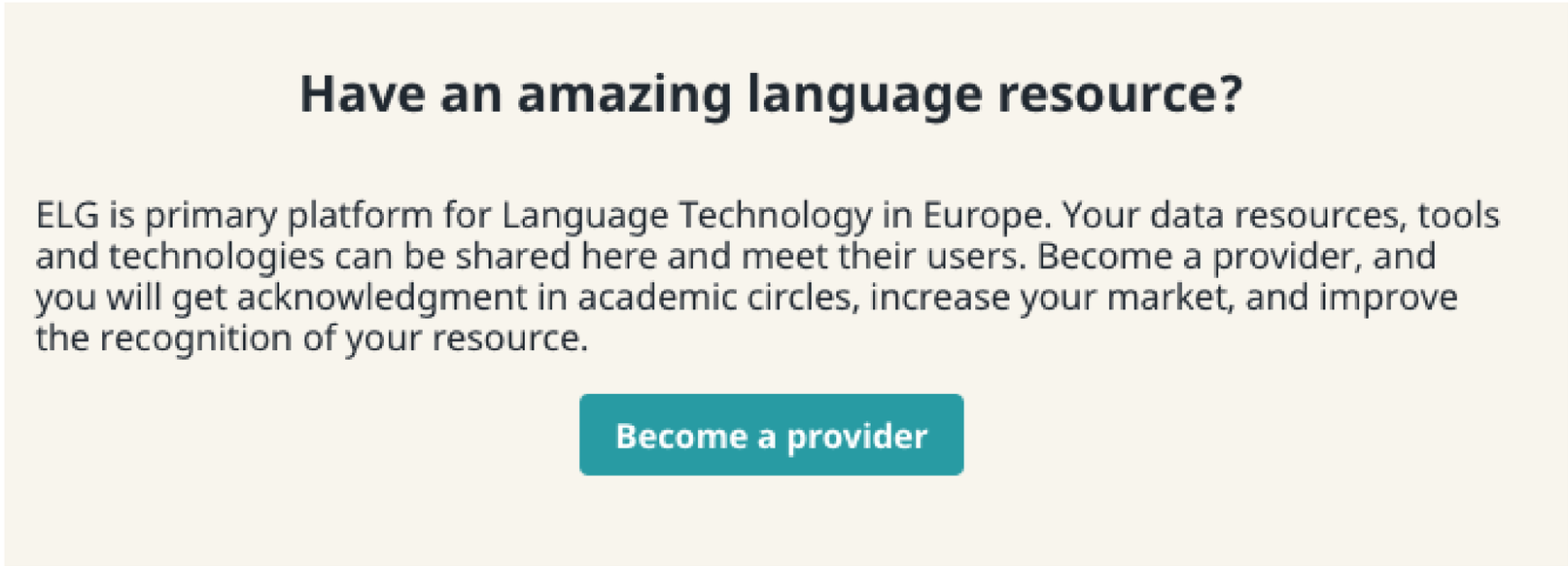
Registering services in the ELG catalogue

Ian Roberts (University of Sheffield)

15/16/17-11-2021 META-FORUM 2021 – Using the European Language Grid (virtual conference)
http://www.european-language-grid.eu

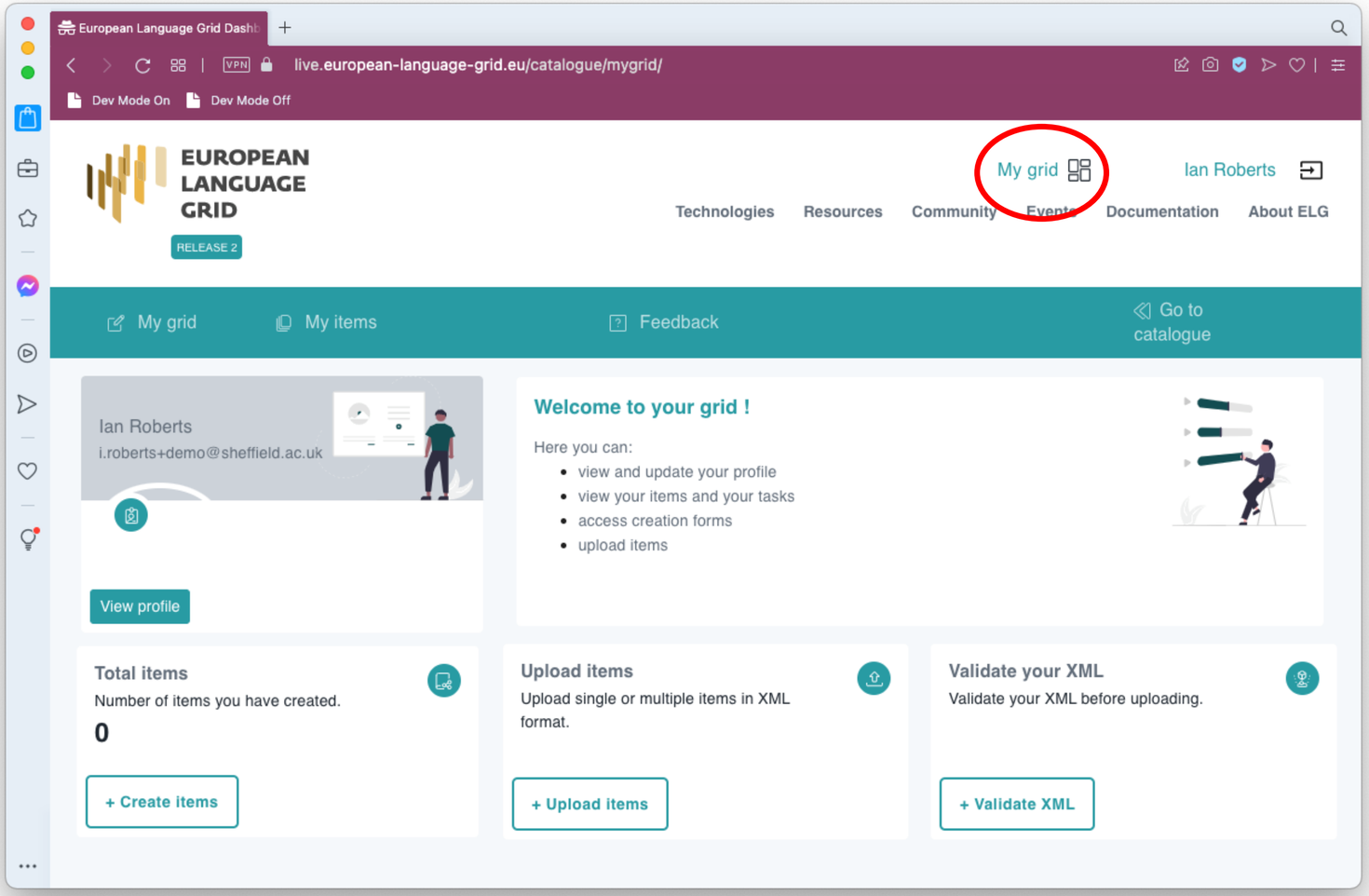
Become a provider

- If you haven’t already, you will need to register as a “provider” on the ELG
  - Sign up & sign in
  - Click “become a provider” on the https://live.european-language-grid.eu homepage
  - Wait for your request to be approved



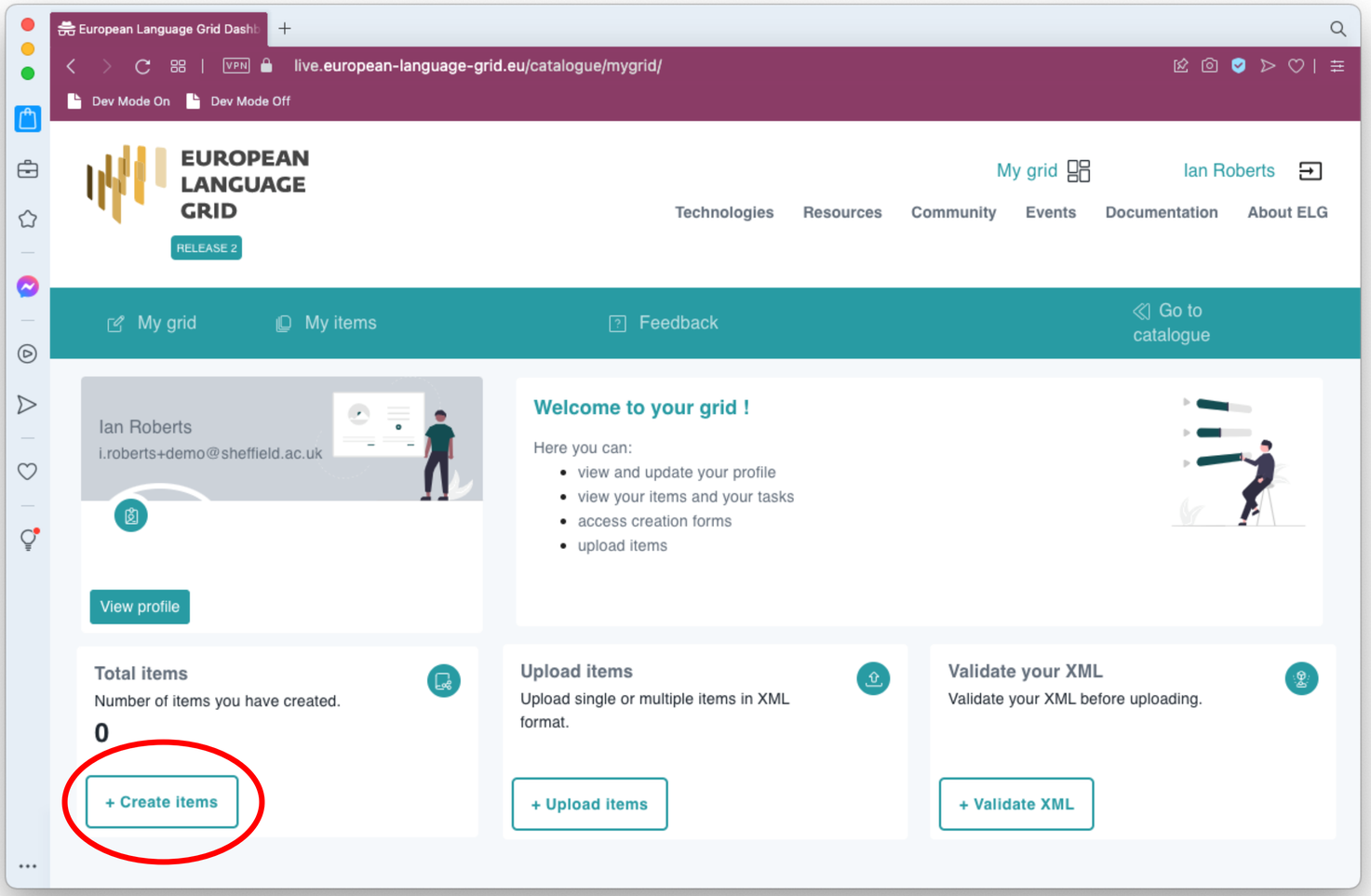
Become a provider

- Once approved as a provider, you will see “My Grid” link



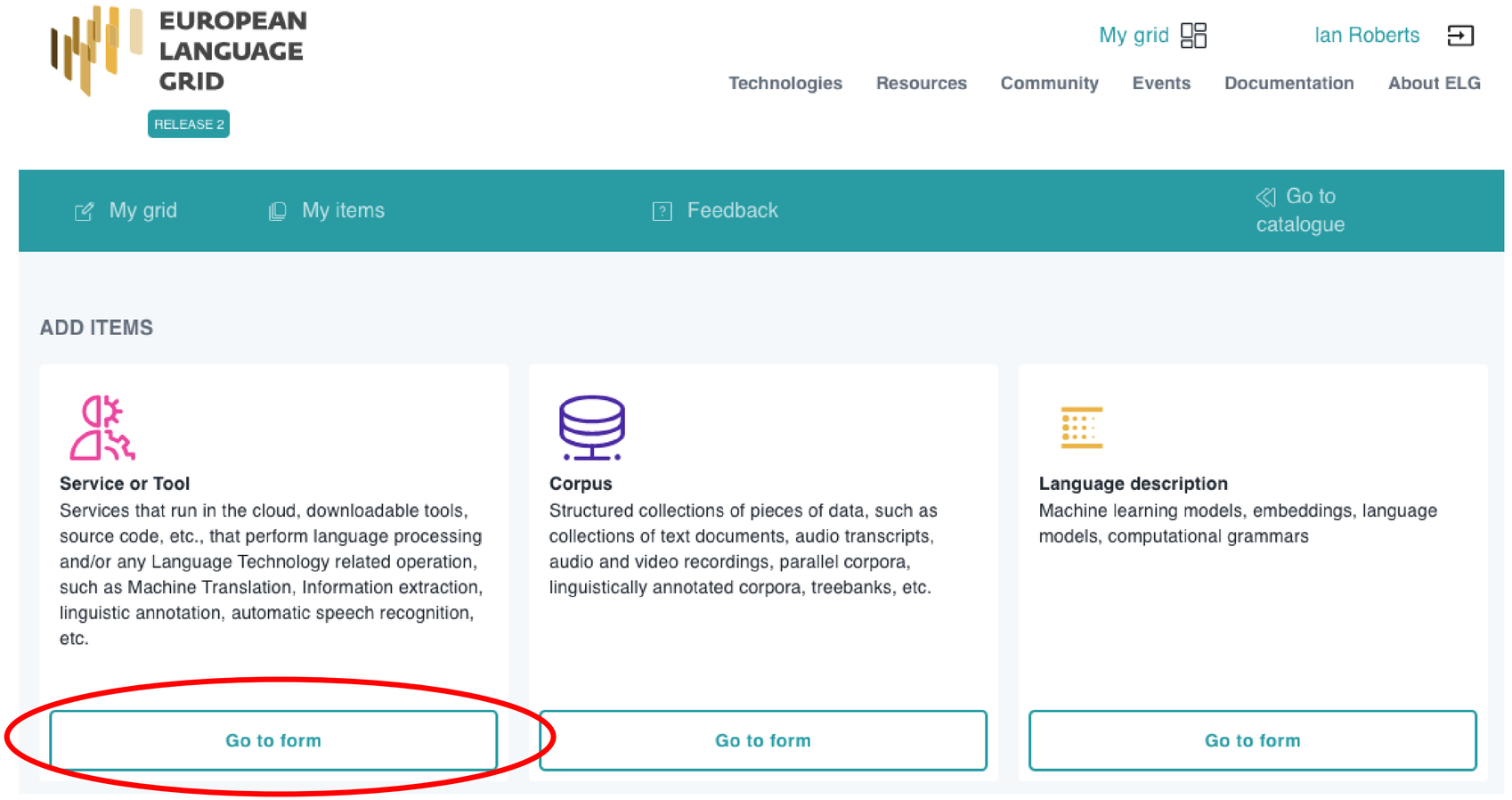
Registering your service

- Simplest approach for a small number of items is to use the metadata editor



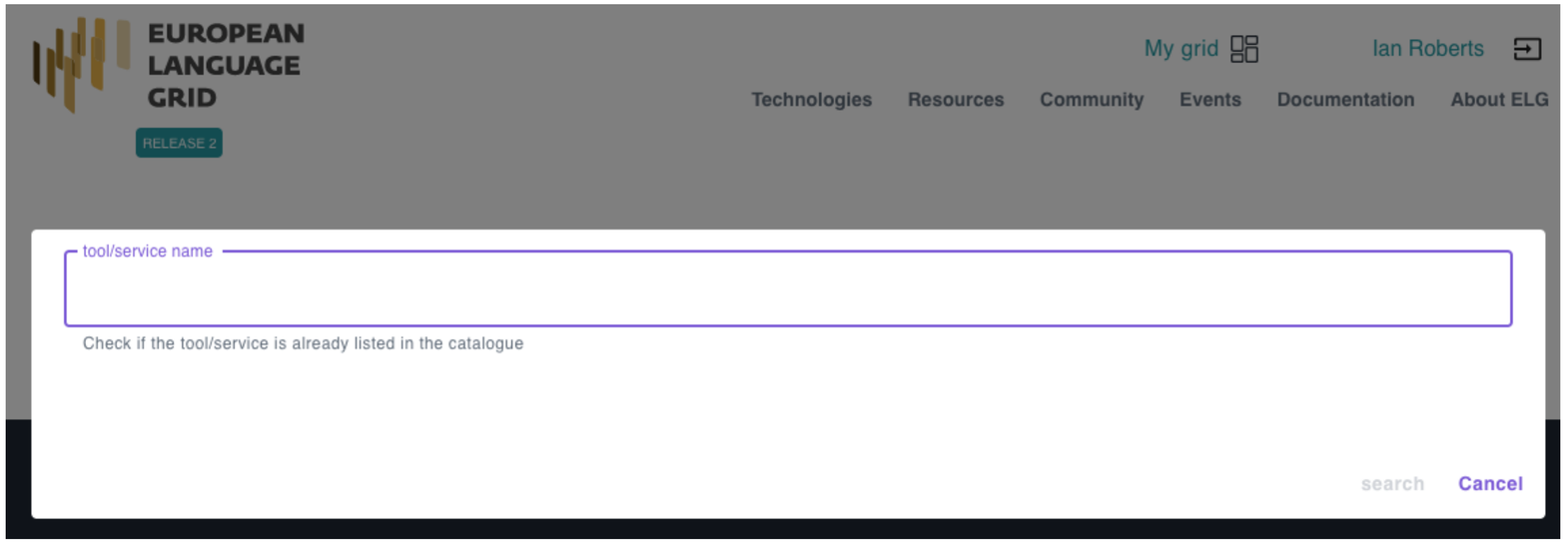
Registering your service

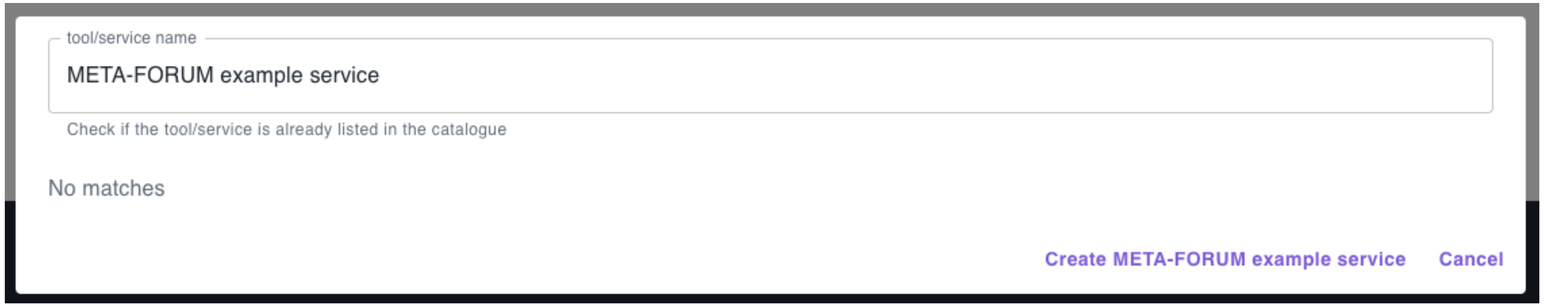
- Simplest approach for a small number of items is to use the metadata editor



Step 1: check for duplicates

- Start by entering a name for your service, and do an initial check that it hasn't already been registered (by you or someone else!)



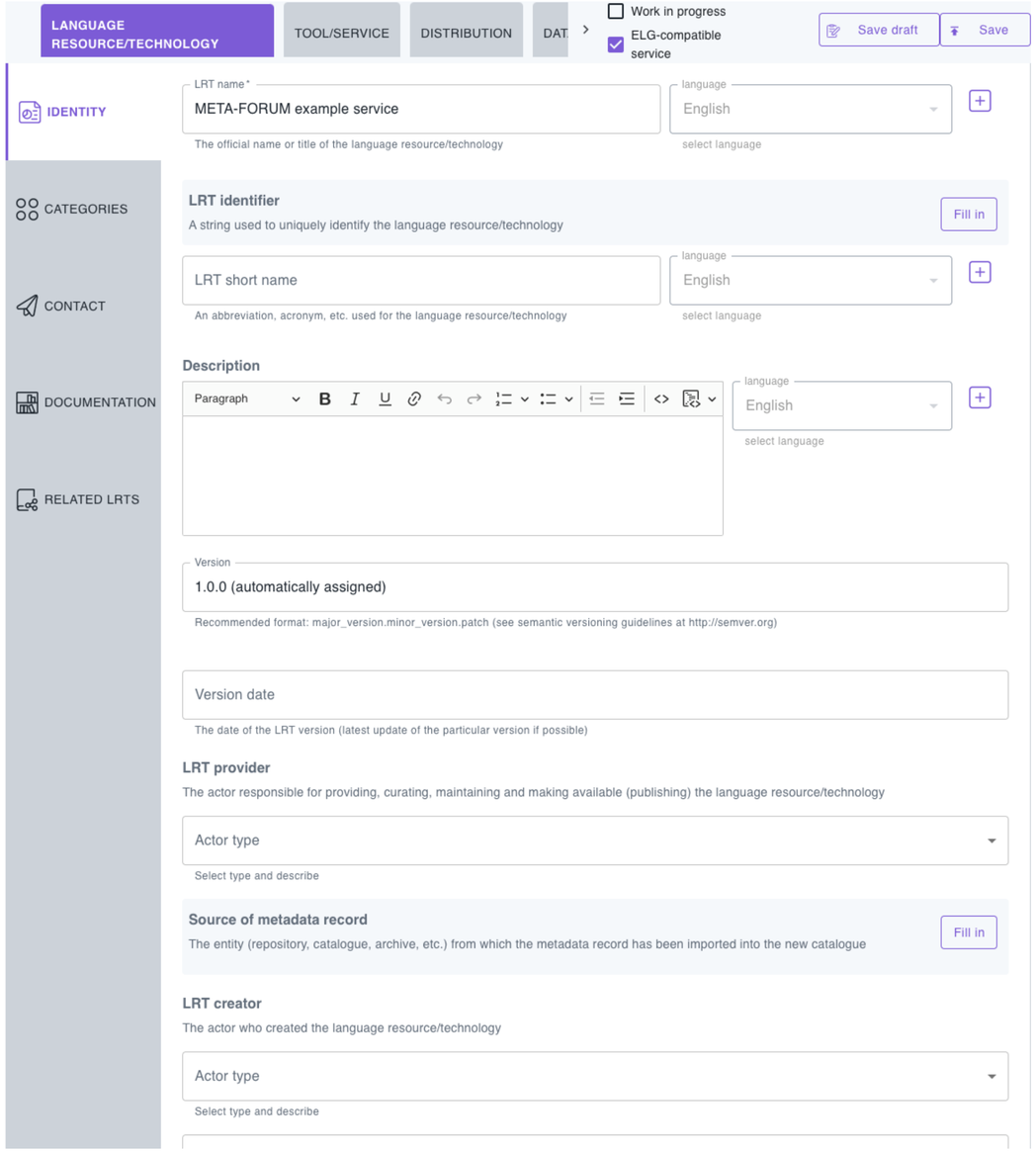


Step 2: confirm this is an ELG-compatible service

- ELG can list both tools/services that are ELG-compatible (hosted in the ELG cluster) and those that are not (e.g. links to software you can download and run locally)
- We are submitting an ELG-compatible service so we must select “yes”

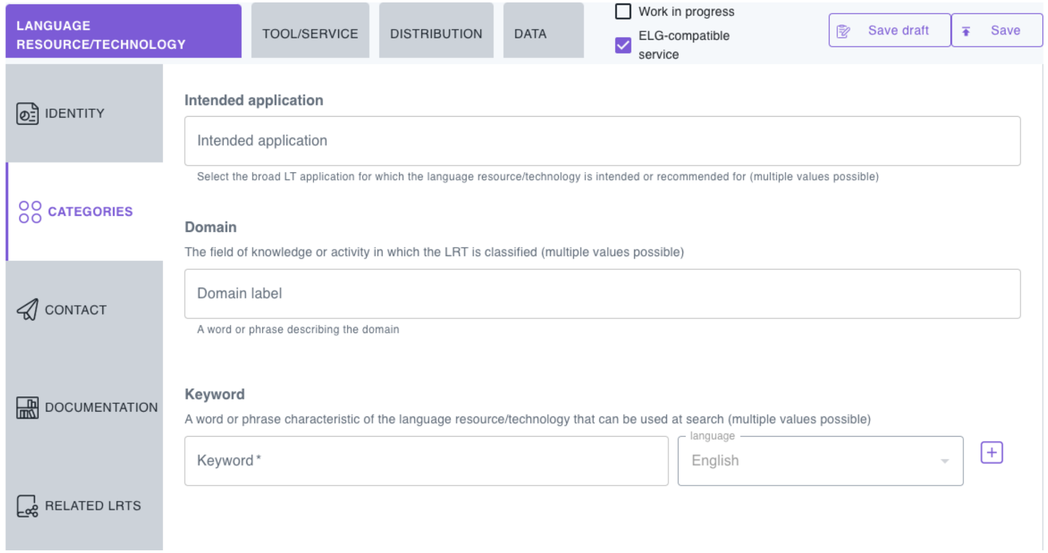
Step 3: Basic information common to all resources

- First tab “Language resource/technology” is for general metadata
  - Identity: name, description, version, logo, links to provider and funding project
    - (key required elements in red)



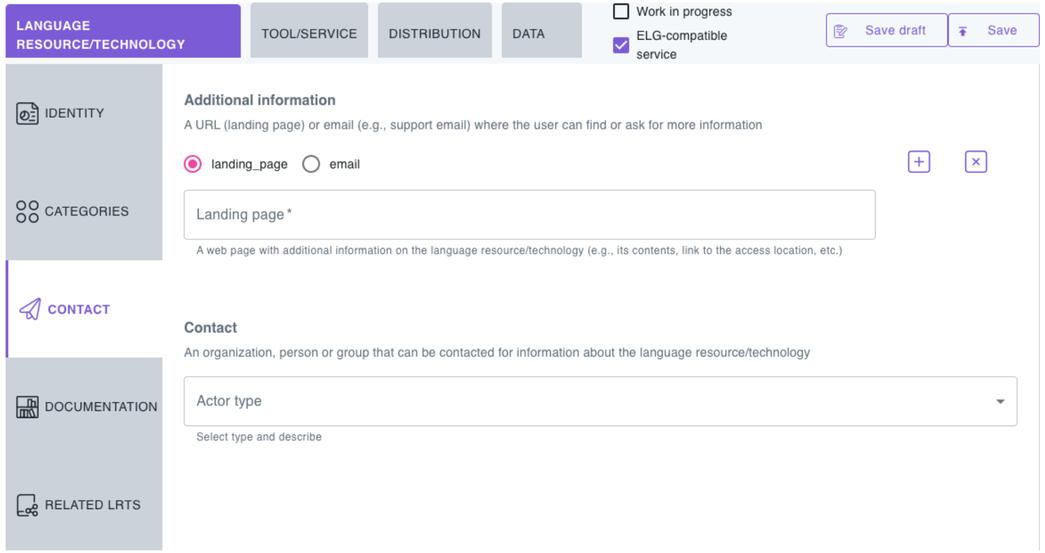
Step 3: Basic information common to all resources

- First tab “Language resource/technology” is for general metadata
  - Identity: name, description, version, logo, links to provider and funding project
  - Categories: intended application, keywords



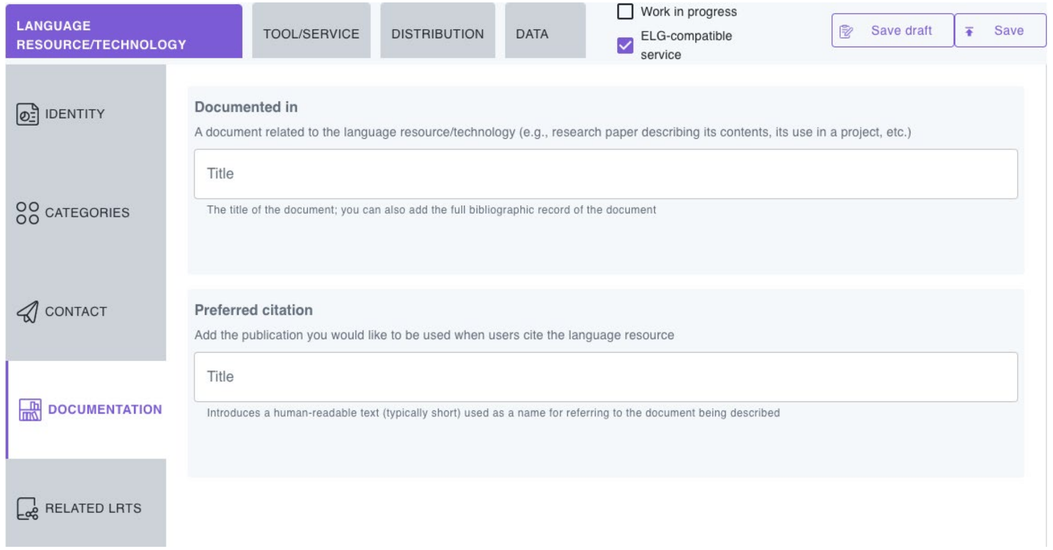
Step 3: Basic information common to all resources

- First tab “Language resource/technology” is for general metadata
  - Identity: name, description, version, logo, links to provider and funding project
  - Categories: intended application, keywords
  - Contact: contact details for support



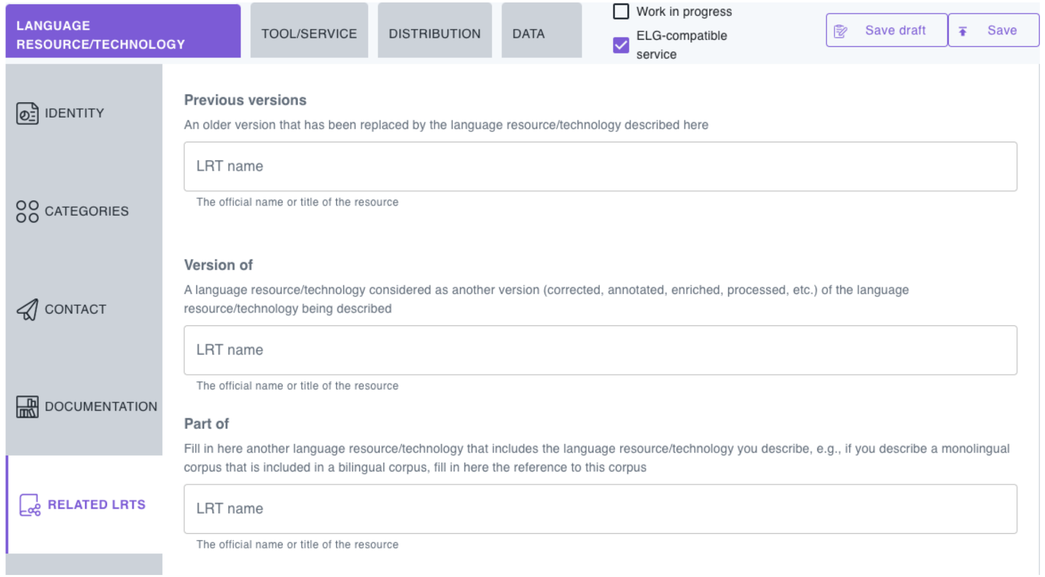
Step 3: Basic information common to all resources

- First tab “Language resource/technology” is for general metadata
  - Identity: name, description, version, logo, links to provider and funding project
  - Categories: intended application, keywords
  - Contact: contact details for support
  - Documentation: links to documentation and citation that users should reference



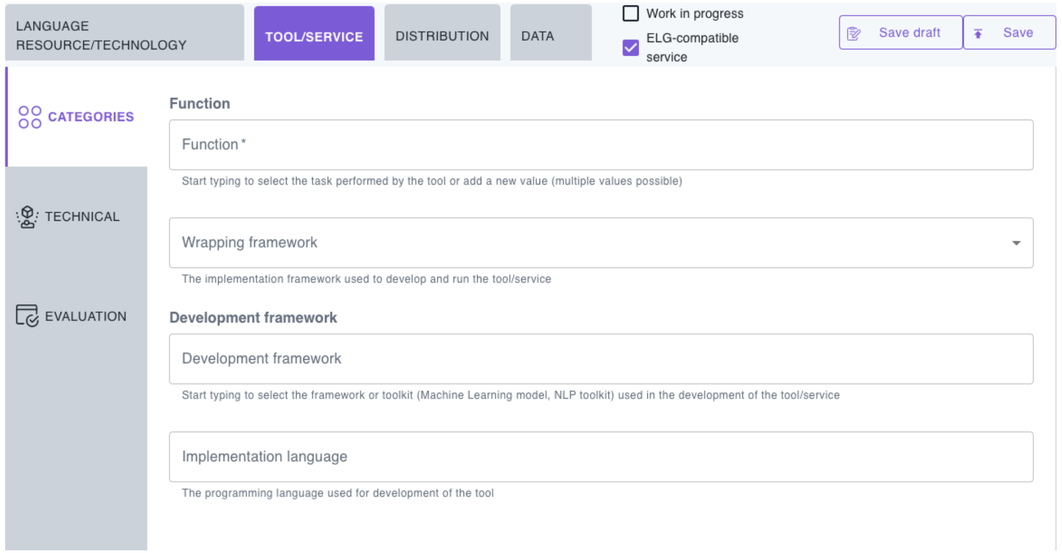
Step 3: Basic information common to all resources

- First tab “Language resource/technology” is for general metadata
  - Identity: name, description, version, logo, links to provider and funding project
  - Categories: intended application, keywords
  - Contact: contact details for support
  - Documentation: links to documentation and citation that users should reference
  - Related LRTs: other items that are related to this one in some way



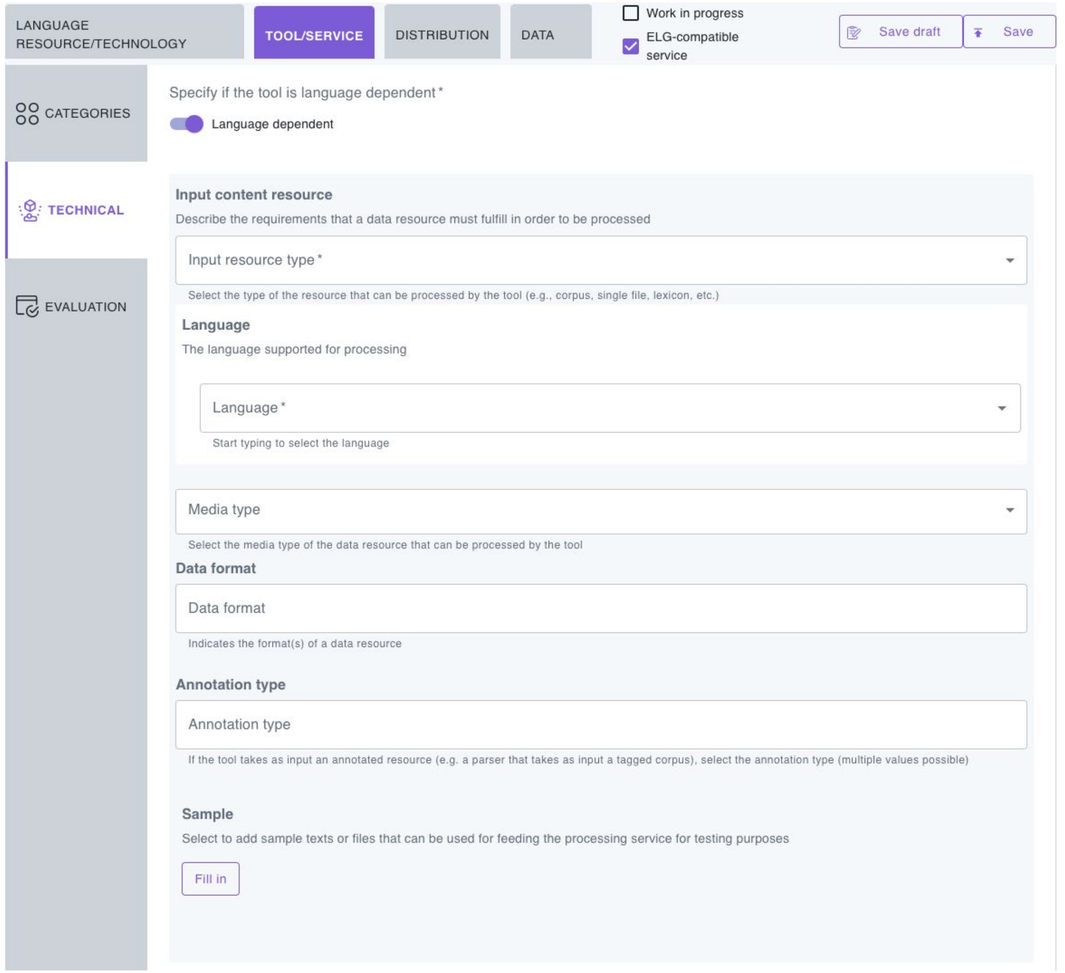
Step 4: Information specific to services

- Tool/Service tab is for all services (ELG-compatible or not)
  - Categories: service-specific categorization info, especially function of the service



Step 4: Information specific to services

- Tool/Service tab is for all services (ELG-compatible or not)
  - Categories: service-specific categorization info, especially function of the service
  - Technical: details about input, output and parameters

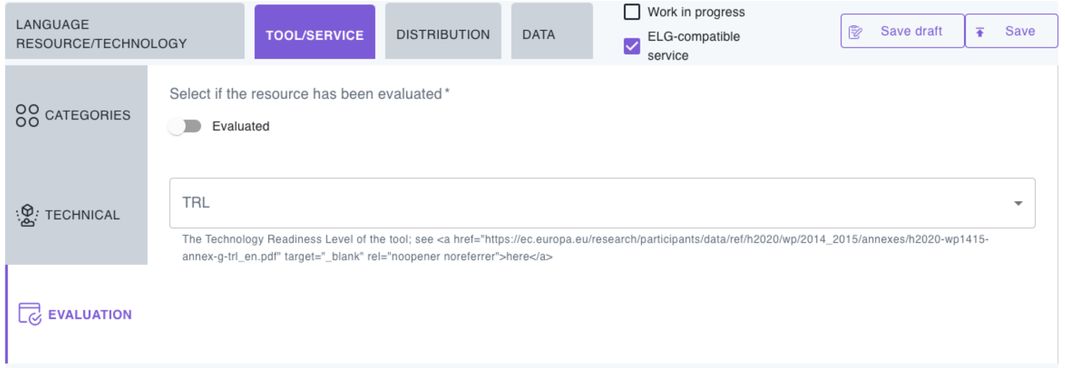


Step 4: More on the “technical” tab

- If your service processes specific languages, set “language-dependent”
- Specify the language(s) on input and output resources – BCP47 tags allowing variant, script, etc.
- For text/audio analysis services (where the output is annotations rather than language), use output resource type “file”, data format “JSON”, the same language(s) as input, and select relevant annotation types
- For MT specify source languages on input resource and target languages on output resource
- If possible, supply at least one known-good “sample” for each input language
  - Sample “tag” should be a meaningful description of this sample – it will be shown in the picklist
  - Either inline “sample text” *or* upload a file as “samples location”, not both
- Parameter declarations in metadata are for the benefit of “try out” GUI – you must still verify them in code and return errors for invalid values
  - For technical reasons, the parameter names “version”, “split” and “audioOnly” are reserved

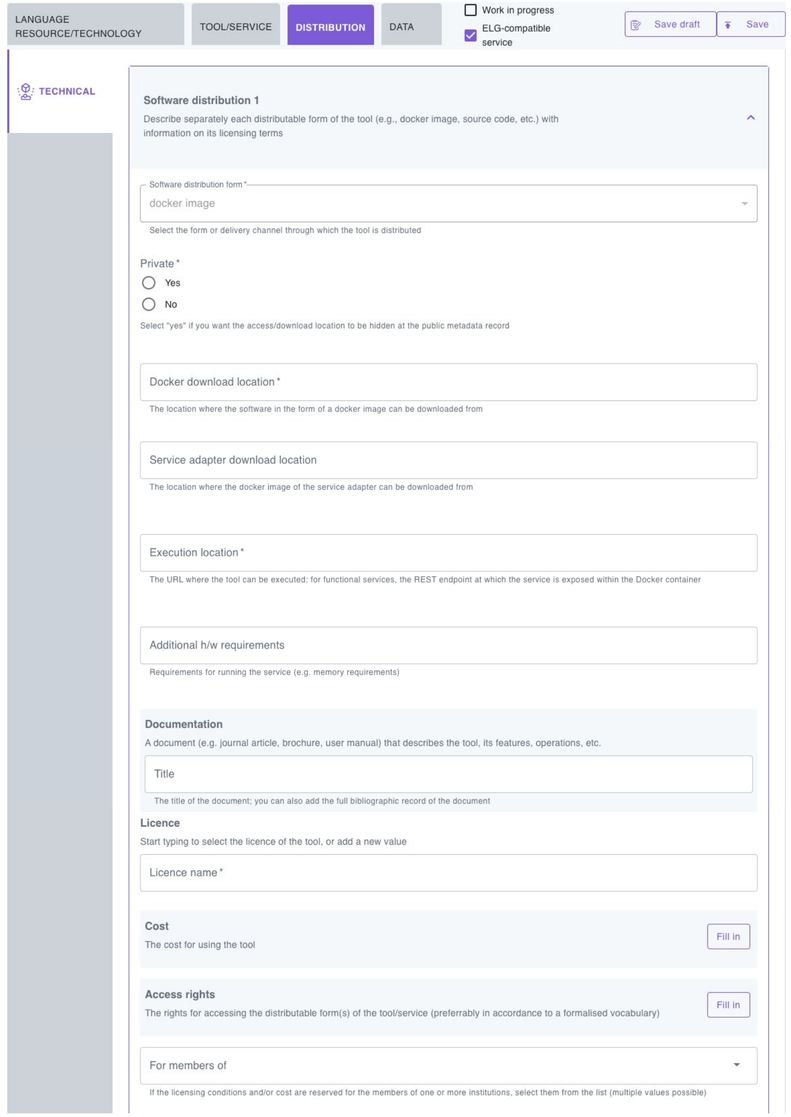
Step 4: Information specific to services

- Tool/Service tab is for all services (ELG-compatible or not)
  - Categories: service-specific categorization info, especially function of the service
  - Technical: details about input, output and parameters
  - Evaluation: TRL from 1 to 9



Step 5: Where can ELG find the code?

- Distribution tab describes Docker image
  - Distribution form: always “docker image”
  - Private: should we advertise the “docker pull” image location publicly?
  - Docker download location: image reference `[registry/]name:tag` from where we can pull your image – this is *not* a URL. Please avoid “latest” tag
  - Service adapter download location: same for the adapter, if this is a main+adapter integration type
  - Execution location: URL of the ELG API endpoint *within the running container*, expressed as http://localhost[:port]/[path]
  - Additional h/w requirements: e.g. if your service needs lots of memory, GPU*, etc.
  - Licence: use existing SPDX options if possible. For closed-source services you can add your own licence terms



Notes

- If possible, always try to re-use existing values rather than creating your own
  - Languages, SPDX licences, annotation types, etc.
- Similarly, please create metadata records for your organization, funding projects, etc. rather than just specifying free text names
  - This helps promote your organization/project and makes it easier to find your records – ELG creates back-links from a project to all the services/resources it funded
  - Create organization first, then project linked to organization, then services linked to project
- If you have many similar records, it may be easier to
  - Create the first one through the editor
  - Export as XML, and use that as a template for others
  - Use “upload items” in “My Grid” to import them
  - We recommend going through the manual cycle at least once before switching to XML

And finally…

- Yes, the form is complex, because it needs to cover many different scenarios
- You will make mistakes!
- … but that’s OK, we are here to help
- If anything is unclear, ask us
  - details under the “feedback” link at the top of “My Grid”
  - or fill in what you can, submit the record for publication, your validator will be in touch with one-to-one assistance