The Search Engines and us: auto-indexing of web content versus the Europeana approach
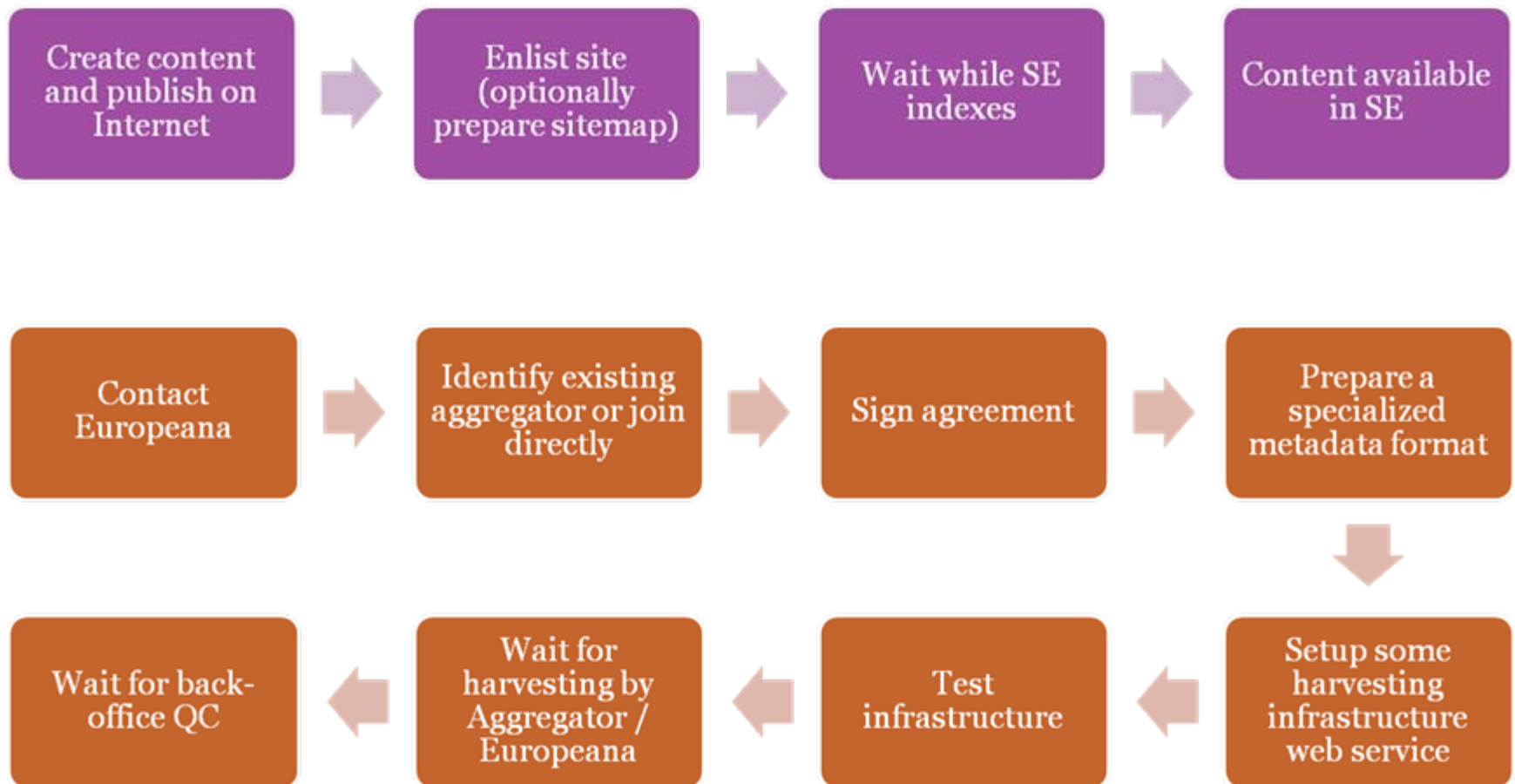
Panel Discussion

Content Ingestion Process 1/2

Create content and publish on Internet → Enlist site (optionally prepare sitemap) → Wait while SE indexes → Content available in SE

Contact Europeana → Identify existing aggregator or join directly → Sign agreement → Prepare a specialized metadata format → Setup some harvesting infrastructure web service → Test infrastructure → Wait for harvesting by Aggregator / Europeana → Wait for back-office QC

Content Ingestion Process 2/2

- Getting content into Europeana is a more comprehensive task than getting it into any search engine.
- Is there room for improvement to this process – or must it have as many semi-manual steps as it has today?
- Why is there a need for an editorial review of the content beyond the content providers' professional assessment of what is good or not?

Does manually crafted metadata, like EDM, deliver value beyond indexing?

- In Europeana we are spending a lot of time creating structured metadata in a specific target format, first the ESE and now the EDM
- The search providers index our content and "guess" what, where, when and who metadata –making a pretty good job out of it. They also allow us to expose deep content through the simple "site-maps" protocol and embed specific metadata using micro-formats such as RDFa or Schema.org
- What is it that Europeana can do which the search engines can't that justifies the extra effort
- ...and can we do <u>it</u> without even more content work – semantic integration is not a technical fix?

What is the most important function of aggregators and aggregations?

- To source as much content as possible into Europeana?
- To improve quality through enrichment as close to the source as possible?
- To provide services on the way between local content resources and Europeana?

- How about business models and licensing for aggregations?

What is the most important function of Europeana today?

- An engine to stimulate growth of digital cultural heritage across Europe?
- As a single point-of-access to European cultural heritage?
- As a provider of Linked Data and APIs allowing European cultural heritage to be interconnected from multiple points of access?
- Is Europeana changing its operating paradigm?

What is the Europeana portal to its content providers?

- We often tell ourselves that the "Europeana portal" is not a competitor to the search engines and yet it both looks like and quacks like a search engine
- In the panel members' views, what is the Europeana portal now – and what should it aspire to be in the future.