

Matrix 2: Linking to Entries vs Embedding Entries in Content Files

Considerations	Entries link to unique IDs (i.e., creating separate index file)	Entries embedded in content (i.e., some form of embedded indexing)
What's inserted into the content files (either by the publisher or the indexer)?	HTML anchors or unique IDs. IDs must be able to be converted by compiler into HTML anchors or other anchor format required for links to work	Index entries are inserted directly into content as fields, XML elements, or using the program's own unique marking system.
Single sourcing capabilities: Can you use one set of files to get indexes in all kinds of output (PDF, HTML, XML, eBook?)	<p>Yes.</p> <p>NOTE: If content is later edited, unique IDs may be lost.</p> <p>NOTE: Most workflows will rely on an outside compiler to generate different outputs, and to generate the index and create links to unique IDs. If the publishing software doesn't natively provide good output, publisher may have to tweak manually.</p>	<p>Yes.</p> <p>NOTE: InDesign Creative Cloud (CC) outputs EPUB with active indexes; it also outputs active indexes in HTML, but only if the book is one long file. Older versions do not output indexes for HTML files or eBooks. Indesign and Frame can output active PDF indexes.</p> <p>NOTE: Word can output a print index. Word does not export active indexes in PDF. If the publishing process requires importing Word into InDesign, note that Word's index markers don't always import correctly. New IndexLinker tool may help, but has not yet been tested.</p> <p>NOTE: XML can support any type of output but requires an XSL style sheet or other automated process to do so.</p>
Granularity: When you click on a link in the index, how close do you get to the referenced content in the text?	<p>Links can only go to where unique IDs are located (e.g., if they are at the paragraph level, links go to paragraph level).</p> <p>NOTE: Display of content will vary dependent on device size and settings. The user may have to scroll a bit to get to the referenced content (e.g., a "page" may consist of several "screens")</p> <p>NOTE: Indexers should consider which ID to use (e.g., choose the ID closest to the term or choose the ID for the nearest section title to provide context).</p>	<p>Links can go to any level.</p> <p>NOTE: Different workflows will dictate what is displayed as locators. Indexes must be recompiled following changes to content.</p> <p>NOTE: Indexers should consider where to embed markers (e.g., choose a location closest to the term or choose a location in the nearest section title to provide context). The publisher may have guidelines for where they want index markers placed.</p>
File management: Who has the files?	<p>Indexer has index files. Publisher has content files (indexer has copy). Publisher needs to notify indexer when/if content changes.</p> <p>NOTE: Publisher can either run scripts to insert anchors and hand off copies of files to indexer, or indexer can run scripts and return a set of files to publisher. Publisher can continue tweaking files once anchors/unique IDs are in, as long as they don't delete anchors/unique IDs.</p>	<p>Indexer has copy of the content files.</p> <p>NOTE: May be done in three ways: (1) Indexer works directly in content files; (2) Indexer works outside the files using unique IDs as interim locators, client or indexer later runs macro to merge index entries into the files; (3) Indexer works in copy of content files which will later be merged via software workflow with the "master copy" held by the publisher</p> <p>NOTE: Unless publisher has an accurate process for merging changed files, they should not make changes to content files once the indexer begins work. If publisher does make changes to content files, the workflow should include error checking.</p>
What if pagination changes and content moves around?	This is only a problem if using page numbers as locators.	Index must be recompiled after each content change and the index reviewed for accuracy or errors.
What about translating/localizing content and index into other languages?	<p>Translation of entries: Index is translated as a complete document OR publisher may choose to request a complete re-index of the new translation.</p> <p>Locators: If using page numbers as locators, pagination may change due to some languages being "wordier" than others and/or different editions being paginated differently.</p>	<p>Translation of entries: The embedded entries must be translated along with the text.</p> <p>Locators: Not an issue, since locators will be generated when the index is regenerated</p>
Chunking content: Can you create smaller books from selected chapters and still have a workable index?	Yes. Potential issues include orphaned subheads, missing cross-references, and broken links.	Yes. Potential issues include orphaned subheads, missing cross-references, and broken links.

Matrix 2: Linking to Entries vs Embedding Entries in Content Files

Considerations	Entries link to unique IDs (i.e., creating separate index file)	Entries embedded in content (i.e., some form of embedded indexing)
Updates and revisions: What about next year/month when we revise it?	If original entries can be sorted consecutively (that is, in order of appearance in the content), indexer can review the index for changed paragraphs and page numbers. NOTE: If later editions are anticipated, consider planning for this capability.	Depends on software used. The ability to view index entries in context is helpful, as is a tool that highlights differences in versions of files (e.g., ExamDiff or TextWrangler). Content merging tools for InDesign (Em WordsFlow and Em Docsflow) retain indexing in InDesign when authors change linked stories.
Can legacy books with an index be published as ebooks with an active index?	Yes. (1) Obtain content files (e.g., PDFs, eBook), including index; (2) If necessary, insert unique anchors into the content files using a tool like LevTech's EPUB Index Manager™, TExtract, or the InDesign scripts from Ole Kvern; (3) sort existing index into locator order; (4) enter new locators that point to the anchors rather than the page numbers; (5) output the index in the appropriate format to include in the eBook. NOTE: Consider where unique anchors should be inserted: page breaks, section level, paragraph level. The presence/absence of unique IDs dictates where entries will be able to link to.	Yes. If legacy book is in a layout/desktop publishing software, and if the software has a later version that supports active ebook export, convert files to the later version.
Last updated 11/13/2013		