

## SEMANTICS AND CONTEXT

# Achieving Business Value by Integrating Tasks, Topics and Content

Duane Degler

In User Assistance terms, "the shortest distance between two points is a relevant keyword." When assistance is needed, the most direct path returns users to their task as quickly as possible with the knowledge needed to be successful. This requires us to design and write with an understanding of the user's context, task, and need. We then reduce seeking time by carefully defining the 'glue' between applications and supporting information. How do we do that? What techniques and technologies are now available to help us reach this goal in new ways? This talk provides 'big picture' ideas for UA practices: understanding the user's context, identifying relevant topic keywords, and integrating applications and content using techniques from the Semantic Web and Topic Maps.

There are four main sections below:

1. The Role of Keywords, and How to Create Them
2. Topics as the "Glue" between Types of Content
3. Topics as the "Glue" between Content and Applications
4. Getting Started: Think Globally, Act Locally

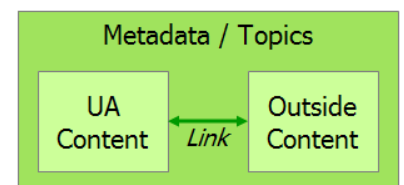
## 1. The Role of Keywords, and How to Create Them

### Why do we need more than content and hyperlinks?

- User assistance (UA) content does not exist in isolation...
- What we're going to talk about today are the implications for finding UA content, as well as integrating it with other content and applications
- Support "content" has grown more complex and dispersed in recent years, to include not just direct/embedded support but also:
  - affiliated content (such as online training libraries and historical databases/repositories)
  - user-generated content (such as power user forums, how-to videos)
  - public content (commentary about your product from outside your "neighborhood")
- How do you help users traverse this expanded space? How do you help them "target" what they need, rather than relying on them to find their way to relevant information?

### The building blocks of "about-ness": content and meta-content

- The foundation of support is your content, along with linked relationships to additional content
- Surrounding that is **metadata**, particularly **topics** – they describe sets of related items of content in terms of what they are about
- You will sometimes hear the term *self-describing* information for the role that metadata plays
- Describing "about-ness" is part of the motivation behind standards like SCORM, Dublin Core, DITA, Topic Maps, and the Semantic Web



*This is a summary of ideas presented at the WritersUA (User Assistance) conference, Portland, OR, March 18, 2008.*

## Metadata and Topics

What are they, and why are they important?

### Metadata

I use this term for the data that describes the nature of an item of content – the “about-ness.” In my work, it has helped me to break this into three categories:

- **“Inward-looking”**: data about the content or application itself, such as the fields in a properties dialog box (title, author, dates, status, etc.)
- **“Outward looking”**: the situations, tasks, context, and real world subjects that are discussed within the content, or to which the content item is relevant
- **“Cross-referring”**: the relationships that an item of content participates in, to allow the management of linked items

### Topics

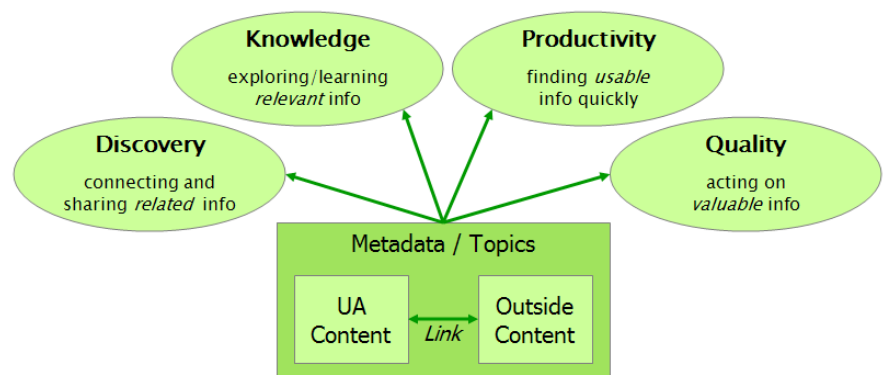
- Keywords are terms/phrases that are important descriptions of your content; their use in navigation and their connections to other keywords allows them to be used as “topics” in a system
- Topics can be used to create collections of content items based on “about-ness” that are *identifiable within a computer*
- The computer should be able to process and follow connections (relationships) between topics
- For more on the often-confusing vocabulary, here is a brief tour: [www.ipgems.com/present/swuidemo/swui\\_terms\\_200706.pdf](http://www.ipgems.com/present/swuidemo/swui_terms_200706.pdf) [1]

## Topics are an *enabler*...

This will be a constant theme throughout this talk: topics have to exist for a reason... to enable something that is not possible without them... in other words, to make things better!

So let’s look outside the box we just saw, and consider how topics touch the “real world”

- **Productivity**: finding *usable* information quickly
- **Quality**: acting on *valuable* information to achieve successful outcomes
- **Knowledge**: exploring and learning *relevant* information
- **Discovery**: connecting and sharing *related* information, supporting innovation and problem-solving



These outcomes seem pretty relevant to business leaders, right?

So if we can focus our topic and metadata efforts on these outcomes, then we can successfully decide where and how to use keywords in our applications. We can also encourage our business leaders to encourage the development of applications that are based on metadata and keywords, because they can produce real benefits.

## How to identify quality keywords? There are many interesting ways!

This part of the talk gives you some tips for identifying and locating good keywords for your own content. You will find more suggestions at this URL: [www.ipgems.com/swui/keywords](http://www.ipgems.com/swui/keywords). [2]

### Term and Concept extraction

You can use automated tools to scan the text in a document or set of documents, perform various forms of frequency and linguistic analysis, and then provide you with a list of candidate terms. For example:

- **Term Extractor:** from the Linguistic Computing Laboratory of the Sapienza University of Rome, Italy. [icl2.uniroma1.it/termextractor](http://icl2.uniroma1.it/termextractor) (beta)
- **Yahoo Term Extraction API:** part of the Yahoo developer library. [developer.yahoo.com/search/content/V1/termExtraction.html](http://developer.yahoo.com/search/content/V1/termExtraction.html)
- **Gnosis:** a FireFox plug-in recently released by ClearForest. [addons.mozilla.org/en-US/firefox/addon/3999](http://addons.mozilla.org/en-US/firefox/addon/3999)
- **Cloud generators:** such as TagCrowd ([www.tagcrowd.com/](http://www.tagcrowd.com/)), Tag the Net ([www.tagthe.net/](http://www.tagthe.net/)), and the Keyword Density Checker ([www.webconfs.com/keyword-density-checker.php](http://www.webconfs.com/keyword-density-checker.php)).

### Existing terminology sets

Many organizations are releasing their “common” terminology, so you can adopt rather than invent! These are just a few of many examples.

- **Open Directory project:** a high-level categorization project. [www.dmoz.org](http://www.dmoz.org)
- **GeoNames:** over 8 million geographical names. [www.geonames.org](http://www.geonames.org)
- **SWOOGLE:** a semantic web search engine. [swoogle.umbc.edu](http://swoogle.umbc.edu)
- **Linking Open Data Project:** a W3C-supported initiative, compiling a reference set of large RDF data sets. [esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData](http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData)
- **MeSH (Medical Subject Headings):** National Institutes of Health, part of the larger Unified Medical Language System (UMLS). [www.nlm.nih.gov/mesh/meshhome.html](http://www.nlm.nih.gov/mesh/meshhome.html)

### Examples from “the world at large”

A large number of sites are now using keywords. Find some that relate to your subject area.

- **Amazon:** subject classification (near the bottom of the page), statistically improbable phrases, capitalized phrases, and even user tags. [www.amazon.com](http://www.amazon.com)
- **Del.icio.us:** see how different terms are related. [del.icio.us](http://del.icio.us)
- **Wikipedia:** browse the extensive categories table of contents, look up specific subjects identify highlighted and linked terms, use the Wiktionary. [www.wikipedia.org](http://www.wikipedia.org), [www.wiktionary.org](http://www.wiktionary.org)

### Manual and internal identification

Some of my favorite techniques:

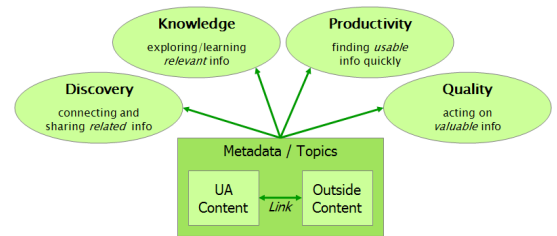
- Have authors and users pull out a highlighter pen and mark the key “about-ness” concepts
- Card sort exercises
- Search term review, looking at the common searches for your content
- Even database controlled value lists can be a valuable source of terms – after all, users encounter them in application drop-down boxes every day!

## 2. Topics as the “Glue” between Types of Content

### So you have self-described content... now what?

Now we need to have the discussion on:

- Where to look for opportunities to leverage topics
- The impact that this will have on users and the business
- The case for making the effort!



### Extending the concept of “support”

We’re talking about using topics to help create a *consistent* user experience

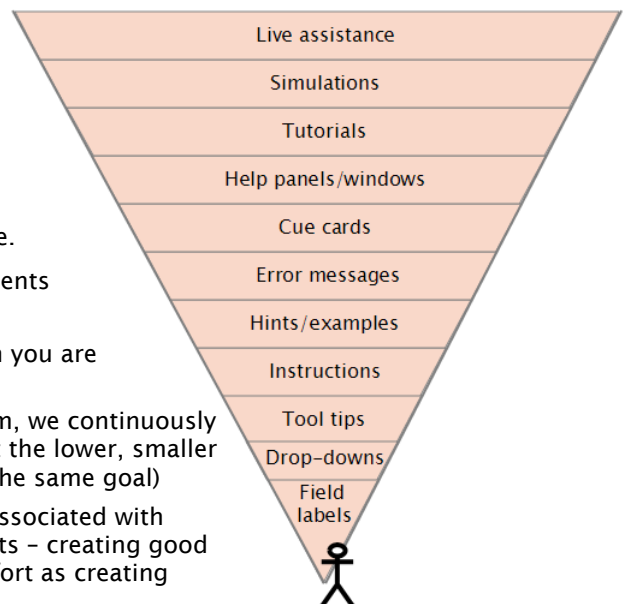
Years ago on a project, I inherited a development team, a documentation team, and a training team. Rather than segmenting responsibilities, we integrated their work to help encourage a seamless whole in the product.

At that time, I sketched up a quick diagram that I have used regularly since then to help people understand that “support” doesn’t just reside in a training program or help file.

Users experience a “continuum of support” to help them achieve their goals – every word they see becomes a message that is part of an application’s overall support. So it all has to work together from labels to live assistance.

There are three reasons for the “funnel” where some elements are larger than others:

- In “distance from task” terms, the lower down you are the closer you are to the task
- Our priority focus flows downward – as a team, we continuously assess whether the user’s need can be met at the lower, smaller level of support (and we’re all committed to the same goal)
- In project terms, there is less time and cost associated with creating smaller, more direct support elements – creating good field labels is probably not as intensive an effort as creating product simulations



[3]

### Extending the role of “context-sensitive help”

What does “contextual” help really mean? In technical terms, is it just a Context ID for a field or page that can be linked to a general description in a help file or web page?

A user’s actual context is much richer... and harder to capture. It includes things like their current task and goals, their experience and knowledge, the time available to them to find information, their awareness of their need for support, the specific situation they are working on (for example, the customer’s situation when processing an insurance claim), and the tools available to them. [5]

### Identifying the task questions at field and page level

When users seek support, they are asking a question which the information needs to predict and answer as clearly as possible. What I have found is that user questions are rarely directly about the technology they are using. More often, the questions are about the task they are working on, or the data that they are managing in a system. For example:

- Field-level questions include “What should I put here, based on what I just heard from a customer?” or “How will the data I put in here affect the outcome I need to achieve?” or “What will the system route me to next, based on what I just entered?”
- Page-level questions include “Am I processing this order correctly according to my purchasing procedures?” or “How do I reconcile the shipping policy and the refund policy for two different items that have to be processed together?”

### Automatic reminders and updates

Another role for context sensitivity is in handling new information. We know that information and instructions change, and we want to make sure people are using the most current information. However, we also don’t want people to have to read through a constant stream of messages about updates to content – they’re more likely to begin to “tune them out” and then miss something which they need to perform effectively.

So a goal here is to tie update announcements to context relevance – this way, people see the updates that relate to them, and to the tasks they are performing. By providing updates at point of need, rather than days/weeks separated from the time they are concentrating on that subject, is a good way to help the information “stick” and ensure the correct actions are taken.

## Relating topics to each other

What are “related topics”? What if we take the concept further?

- Using documents that are closely related by topic, automatically produce link lists
- Create categories of related information, based on the nature of the relationship between two items of content (e.g. overviews, alternatives, next steps, compliance standards)
- Signpost additional sources for information on a topic, without interrupting the user
- Define glossary terms in context, for situations where a term may have slightly different meanings depending on the situation

Some examples:

- Procedural content pages (see below)
- IRS TaxMap
- MeSH browser
- Recommender systems

## A procedural content page example

Below is an example of a page that illustrates the way topics can be used to:

- Create links between related information
- Embed useful related information into documents dynamically, requiring less maintenance on the part of authors
- Adapt the document’s presentation according to user context

The screenshot shows a web page titled "Create Multi-Trip Itinerary" with the following content:

- Header:** Previous Page | Index | Home
- Navigation:** Recent Pages, Go, Previous Step, List of Steps, Next Step
- Main Content:**
  - Multi-Trip Itinerary** (Last updated: 12.18.2007)
  - Text: "Your travel schedule may include more than one city stop, requiring you to con travel to is part of the same business reason. If you are preparing a multi-trip it to more than one project, then refer to RAG charge code tables to be sure that correctly. This will save time and rework."
  - Text: "A Multi-Trip Itinerary includes more than one destination apart from your home location, which requires either expensed travel (air, train, or rental car) or overnight stays in different locations."
  - Steps:**
    - Select "Multi-Trip" radio button on the booking form.
    - Select your starting airport and then all destinations, in the order you will visit them. (more >>)
    - For each trip between two destinations, identify the dates and times of travel, then provide a description of the activity in each location, or the project reference. You describe the details of your trip in this step. Have as much detail as possible available when you begin. Where possible, copy wording from the project justification form. You can edit this information later, if required. (<< less)
    - Select "Look up Travel Options" to continue.
  - For more details:**
    - Add a layover when en-route
    - Change part of a trip
    - More than one airline on an itinerary
  - Legal Basis:**
    - CFR 28-439 Creating travel plans based on government project funds: reporting and considerations
    - Guidance note 86A-2006: More than one person traveling for project purposes
- Right Side Widgets:**
  - Remember...** Authorization approval levels updated for recharging to projects (2.23.2008) Learn more...
  - Glossary for this page** Print
  - See Also:**
    - Traveling out of country
    - Set airline seat and meal preferences
    - Completing expense form
  - What others said:**
    - Work-around for single night layover
    - Clarification of policy when charging to more than one project
  - Comment

Prompts, updates and related links can appear based on topic matching

Steps can show or hide based on user context, like removing steps or adding additional instructions for novice users

Sets of links can be typed, to go beyond “see also” or “related subjects,” so users know what they’re getting

Glossary items can be managed as topics, making it easier to add glossaries to a page

## Using topics to help people navigate information and sets

Topics are often used for navigation and to filter large lists down to more relevant sets of information. The phrasing of topics is an important consideration when they are used in this way, because users will be more successful the more they are working with clear and consistent terminology.

- Faceted navigation and browsing:
  - eCommerce faceted categories
  - mSpace ([mspace.fm](http://mspace.fm))
  - Exhibit ([simile.mit.edu/exhibit](http://simile.mit.edu/exhibit))
  - Flamenco ([bailando.sims.berkeley.edu/flamenco-interface.html](http://bailando.sims.berkeley.edu/flamenco-interface.html))
- Showing parts of the taxonomy as “breadcrumbs”

### 3. Topics as the “Glue” between Content and Applications

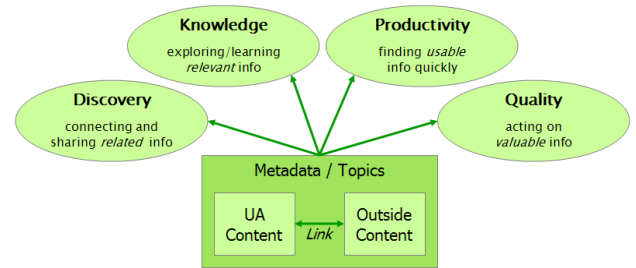
#### Responding to goals and motivations

What matters to our **users**? Information must be:

- Relevant
- Usable
- Related
- Quick

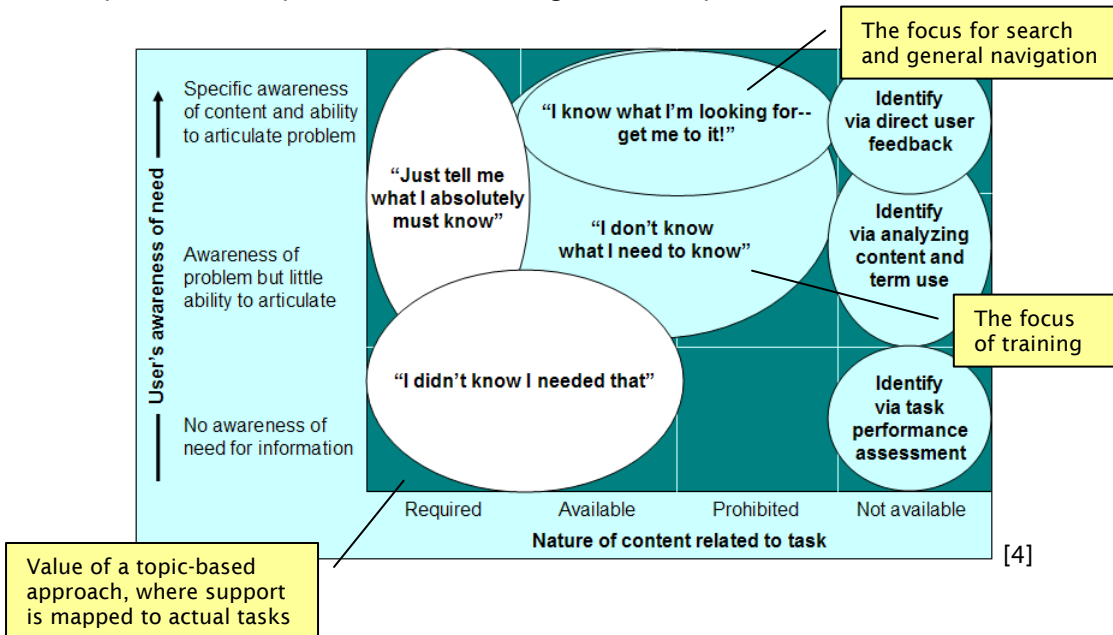
What matters to our **business leaders**? Users must achieve:

- Quality
- Productivity
- Consistency
- Innovation



#### What people don't know... can affect quality

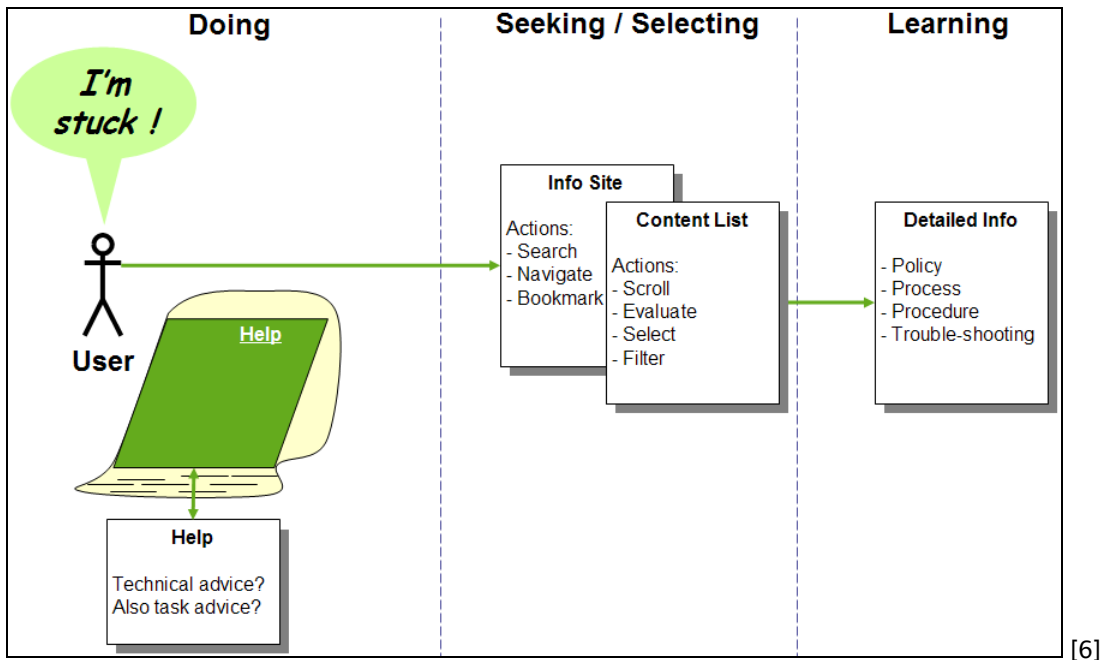
- Topics play a key role in creating *proactive* information resources
- Often, it's not your most experienced person who needs support, so relying on them seeking relevant information may not be the most efficient or effective method
- In addition, for experienced people who are engaged in unfamiliar tasks, the situation is amplified – they are more likely to believe they understand something, even if they don't



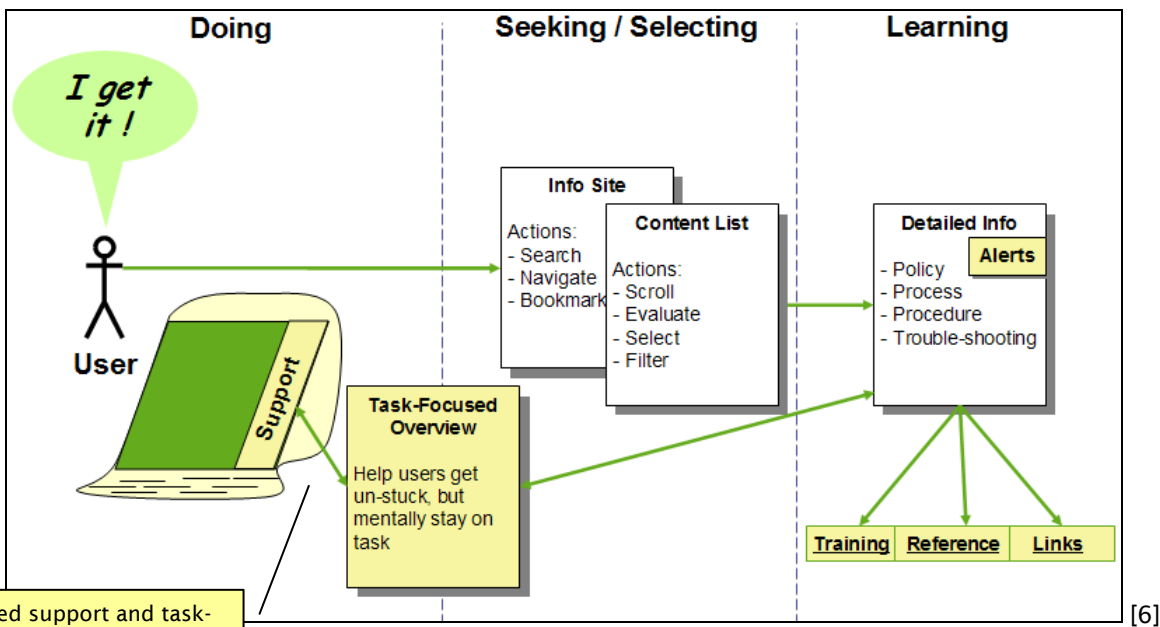


### Distance from task to information affects productivity

If someone has to seek information, they are mentally being drawn away from their focus on the task.



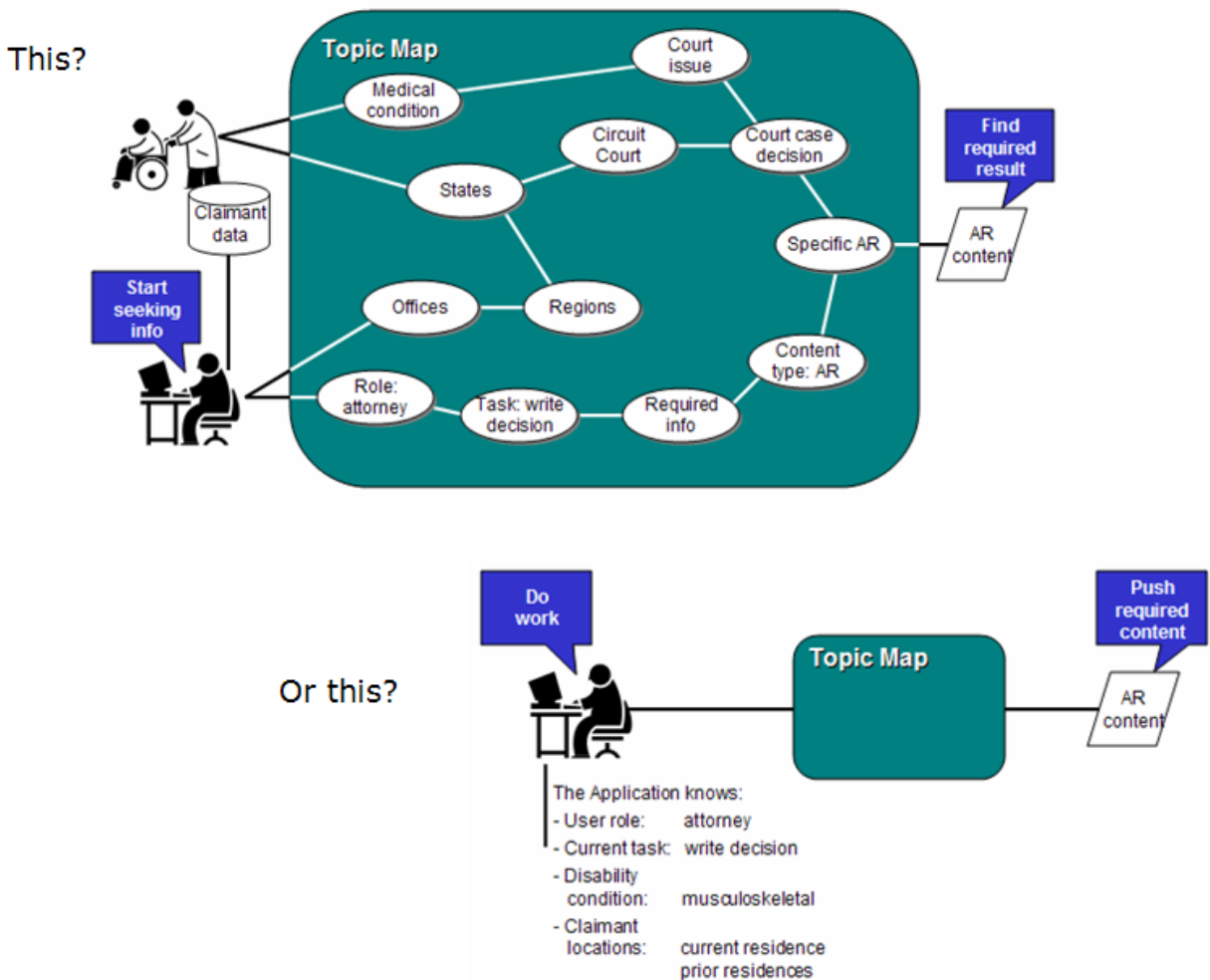
The more we can map support information directly to the *task* and the *context*, then the easier it should be for the user to remain *on task*. At the same time, topic-linked information should provide a richer environment for increasing knowledge.



To embed support and task-focused information, you need the user's context, which is more than just what page/field they are on

**We already have all the context information... It's in the transactional apps!**

- What if our interactive applications passed *more* than just a context ID? What if they passed *context data*?
- What do we mean by context data? Well, for example, anything about the user that is typically held in a profile or "header" (location, account types, claim numbers, dates), data about the process or transactions being performed, and all the current selections of values (drop-down list values, options, etc.)... *topics* that must be associated with the content.
- In the illustration below, court cases that affect certain insurance claims require information about the claim, the claimant's condition, and the location (to identify court jurisdiction). Why shouldn't this information be available to the support system, so new court judgments that the user may not be aware of can be provided proactively, allowing the correct procedures to be followed?



[4]

## 4. In Summary: Think Globally, Act Locally

---

### Craft your own vision

- Keep the “big picture” goals in mind while working on your immediate projects
- You’ve seen some of my ideas... now think about your own vision
- Get the vision into sound bites... so others can begin to recite the vision, too (beware, it can take about 6-12 months for this to happen)
- Have the persuasive evidence in mind at all times:
  - Objective – performance, usage, and satisfaction data
  - Subjective – stories and anecdotes that bring the true user situation to life

### Make it easy for others to adopt

Focusing on topics and the relationships between content requires work. There is one lesson I’ve learned from a number of projects: if the new approach that is required is more difficult than the “old way” then people will stick with what they know.

You not only have to solve the technical and substantive challenges with using topics, but also work to solve a pain point among your collaborators. That is what will produce success.

For example, when I was working on an intranet content management system, I realized that it was important to manage software Help content from the same repository that managed the web-based procedures and training content. That would allow the cross-referencing to be relevant to the users. However, I didn’t get the Help authoring tools prioritized highly enough to win that user group... they continued to author Help using a number of different systems and build new authoring tools in-house.

The missed opportunity? Accesses from transactional applications could have been up to 50% of the overall context access traffic (reduced time seeking), and be more meaningful to business users (more relevant to business performance)... but only if it was the “least effort” solution for the authors.

### Take small steps now, rather than waiting to leap

Grab every opportunity you can to do something. For example:

- Build a relationship with the technology and business people
  - Share your vision, and begin to map it to the problems they describe
  - Review their controlled vocabulary lists and data models, and even help them refine and document them (nobody ever has enough time for this, so your help may be welcome)
  - Get involved in tool/vendor evaluations, and ask how tools handle topic management
- Start capturing keywords in whatever tools you currently use – even if they aren’t used immediately, some of the hard, slow work is already done
  - For example, when developing an organizational taxonomy for navigation and search purposes, we started capturing additional data about the relationships between the terms, so that in the future if we needed to change a term we could assess what other terms were affected – even though we didn’t have a change system that could consume that data at the time
- Think about the chunking and modularity of your content – is each chunk “about” one key topic? Will you be able to make sense of individual items if they are delivered in a different way in the future?

---

## Further details on these ideas

- [1] Degler, Duane (2006). From *Next Generation Design: Interacting with the Semantic Web*. UPA 2006 conference proceedings. Broomfield, CO. June 2006. Online: [http://www.ipgems.com/present/swuidemo/swui\\_terms\\_200706.pdf](http://www.ipgems.com/present/swuidemo/swui_terms_200706.pdf)
- [2] Degler, Duane (2007). *Coming to Terms with Keywords*. UserFocus 2007 conference proceedings. Washington, D.C. October 2007. Online: [http://www.designforcontext.com/pubs/dd\\_keywords\\_userfocus2007.pdf](http://www.designforcontext.com/pubs/dd_keywords_userfocus2007.pdf)
- [3] Degler, Duane (2008). *User support starts small, but is consistent*. Blog posting: [www.oddapproach.com/200802/consistent\\_support](http://www.oddapproach.com/200802/consistent_support)
- [4] Degler, Duane and Battle, Lisa (2003). *Can Topic Maps Provide Context for Enterprise-Wide Applications?* Extreme Markup Conference, Montreal, Canada, August 2003. Online: <http://www.idealliance.org/papers/extreme03/html/2003/Degler01/EML2003Degler01.html>
- [5] Degler, Duane and Battle, Lisa (2000). *Knowledge Management in Pursuit of Performance: the Challenge of Context*. Performance Improvement (EPSS Special Edition). ISPI, 39(6), July 2000. Online: [www.ipgems.com/writing/kmcontext.htm](http://www.ipgems.com/writing/kmcontext.htm).
- [6] Degler, Duane (1999). *The Relationship Between Distance Learning, CBT and EPSS - Common Checklists*. Performance Support 1999, part of the Online Learning Conference, Los Angeles, CA. Online: [www.ipgems.com/present/dd\\_epss-elearning\\_ps1999.pdf](http://www.ipgems.com/present/dd_epss-elearning_ps1999.pdf).

© Degler, D. (2008). *Achieving Business Value by Integrating Tasks, Topics and Content*. WritersUA conference proceedings. Portland, OR. March 18, 2008. Online:

[http://www.designforcontext.com/pubs/dd\\_buinessttopics\\_writersua2008.pdf](http://www.designforcontext.com/pubs/dd_buinessttopics_writersua2008.pdf)

Contact the author at [duane@designforcontext.com](mailto:duane@designforcontext.com). This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License](http://creativecommons.org/licenses/by-nd/3.0/).