O'REILLY®

4th Edition

Information Architecture

FOR THE WEB AND BEYOND

Louis Rosenfeld,
Peter Morville & Jorge Arango

Information Architecture: For the Web and Beyond
by Louis Rosenfeld, Peter Morville, and Jorge Arango

Printed in Canada.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://safaribooksonline.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Acquisitions Editor: Mary Treseler
Editor: Angela Rufino
Production Editor: Matthew Hacker
Copyeditor: Jasmine Kwityn
Proofreader: Rachel Head

Indexer: Judith McConville
Interior Designer: David Futato
Cover Designer: Ellie Volckhausen
Illustrator: Rebecca Demarest

February 1998: First Edition
August 2002: Second Edition
December 2006: Third Edition
September 2015: Fourth Edition

Revision History for the Fourth Edition
2015-09-01: First Release

See *http://oreilly.com/catalog/errata.csp?isbn=9781491911686* for release details.

978-1-491-91168-6

[TI]

Table of Contents

Part II. Basic Principles of Information Architecture

PART I

Introducing Information Architecture

Information is more abundant today than ever before. With smartphones, activity monitors, smart watches, tablets, and new Internet-enabled appliances of every kind, we also have many more ways of interacting with it than before. This abundance and pervasiveness makes our lives better in many ways, but it also introduces new challenges. With so much information available in so many places, it can sometimes be difficult to cut through the noise to *find* the information you need and *understand* it once you have found it.

Information architecture (IA) is a design discipline that is focused on making information findable and understandable. Because of this, it is uniquely well suited to address these challenges. IA allows us to think about problems through two important perspectives: that information products and services are perceived by people as *places made of information*, and that these information environments can be *organized for optimum findability and understandability*.

This first part of the book explains what IA is, what problems it solves, and how it can help you create more effective products and services. Part II and Part III will then show you how.

Let's get started!

CHAPTER 1

The Problems That Information Architecture Addresses

And it really doesn't matter
If I'm wrong, I'm right
Where I belong I'm right
Where I belong
—"Fixing a Hole,"
Lennon–McCartney

In this chapter, we'll cover:

- How information broke free from its containers
- The challenges of information overload and contextual proliferation
- How information architecture can help people deal with these challenges

Marla was in the mood for The Beatles. She walked over to the shelf where she kept her LP records and looked through her collection. Fortunately, Marla was very organized: her record collection was neatly sorted alphabetically by the artist's name. Alice Cooper, Aretha Franklin, Badfinger... and there, next to her Beach Boys albums, were The Beatles. She pulled the *Sgt. Pepper's Lonely Hearts Club Band* vinyl disc out of its sleeve and put it on the turntable, and relaxed as the music started.

For most of our history, the information we have interacted with has existed in a one-to-one relationship with the artifacts that contain it.

Marla had only one *Sgt. Pepper's* album, and if she wanted to listen to it, she needed to know exactly where it was on the shelf. If she was traveling and didn't bring her record with her, she couldn't listen to it. Because the information (the music) was physically embedded in containers (vinyl discs), and she only had one copy of each, she had to define "one right way" to organize her records. Should they be ordered alphabetically based on the artists' first names, as shown in Figure 1-1, or their last names? What about albums in which the composer mattered more than the performer, as in her copy of Holst's *The Planets*? Then there were compilation albums, containing music by many artists. Should they be listed under "Various Artists"? And when she bought a new album, she needed to remember to store it in the right place in the collection. It all got very complicated very quickly. Perhaps she shouldn't bother with organizing them at all... but then she wouldn't be able to find them easily when she was in the mood for a particular artist.

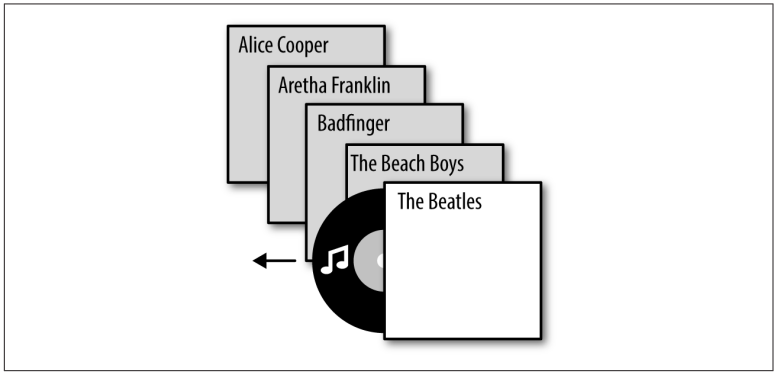

Figure 1-1. Marla's music is embedded in physical objects—vinyl records—so she must choose how to organize them on the shelf

Now meet Marla's son, Mario. Instead of vinyl discs, Mario's record collection consisted of compact discs (CDs). Because the music in the discs was stored digitally, he could now randomize the order in which the songs were played. He'd been promised that the music would also sound better, and the discs would last longer than the previous technology. It was great! However, even though the music was stored digitally, his plastic discs were not that different from his mother's collection: the music was still tied to the individual physical discs that contained it. He still had to choose whether to organize the discs by the artist's name or the album's name; he couldn't do both.

But then, in 2001, Mario got an iMac. The colorful computer's advertising campaign invited him to "Rip, Mix, Burn" his music—in other words, liberate it from the plastic discs that contained it and get it into his computer ("Rip"). Once there, it would sound just as good as the CDs, but now he could explore it any way he pleased: he could browse his collection by artist, genre, album title, song title, year produced, and more. He could search it. He could save backup copies. He could make playlists that combined the music from various albums ("Mix") and record songs onto blank discs ("Burn") to share with friends (much to the chagrin of the people who'd produced the music).

As shown in Figure 1-2, Mario was no longer limited to the one-to-one relationship between information (the music) and containers (the discs) that his mother had to deal with. He was no longer constrained to deciding between sorting the albums alphabetically by artist name or album name; he could now do both simultaneously. He could make multiple perfect copies of his songs, and bring them with him on his laptop when he traveled. Mario stopped thinking of his music as something tied to its container. It had dematerialized.

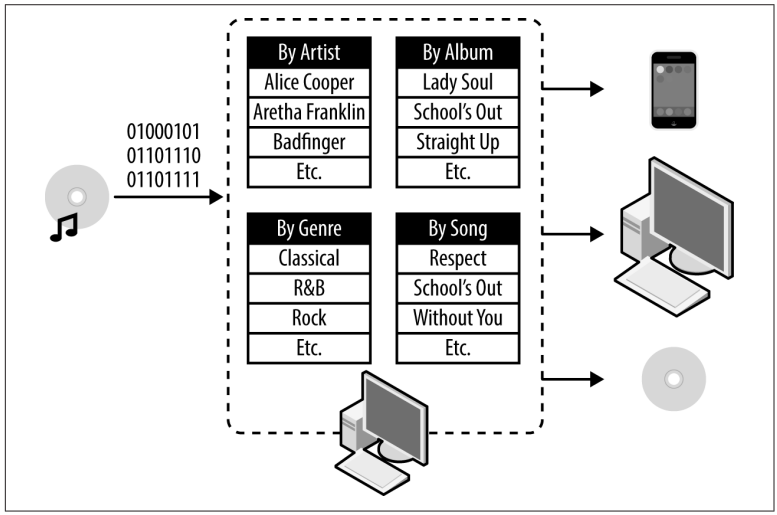

Figure 1-2. Being digital, Mario's music collection can be organized in more than one way and can live in multiple devices simultaneously

Hello, iTunes

The tool that Mario used to do all of this, iTunes, is shown in Figure 1-3. Digital music had been around for a long time before iTunes, but this was the first time that many people encountered it in the mainstream. Originally a third-party application called SoundJam, iTunes was acquired by Apple in 2000 to become the default music player included with Macintosh computers. In its initial release, iTunes served a clear purpose: it allowed Mario to create and manage a music library for use in his own computer ("Rip, Mix, Burn"). He spent a long weekend importing his collection of 40 CDs into his Mac and organizing his music, and put the discs away for good. From now on, his music would be all-digital.

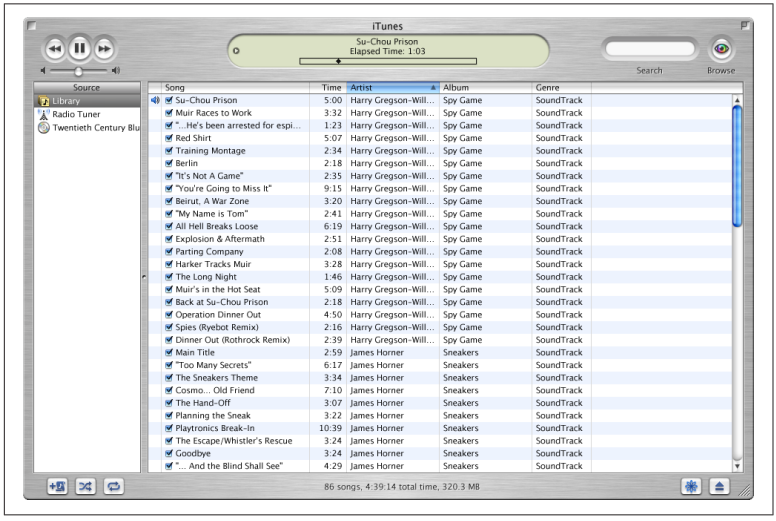

Figure 1-3. iTunes 1.0 browsing by Artist and Album (image: http://bit.ly/et_tu_itune)

The first version of iTunes had a few distinct modes—for example, there was a "ripping" mode that showed progress when the user was extracting music from a CD into the computer—but its focus was clearly on allowing people like Mario to find and play music from their own collections. As a result of this reduced feature set, it had a very simple user interface and information structures. Mario loved it, and playing music became one of his favorite uses for his Mac.

However, iTunes started to become more complex over time. Each new release of the app introduced amazing new features: smart play-

lists, podcast subscriptions, Internet radio station streaming, support for audiobooks, streamed music sharing, and more. When Apple released the iPod, Mario rushed to get one. iTunes was now about more than just managing music on his Mac: it was also about managing the library on his portable music player. In 2003, Apple introduced the iTunes Music Store. Now Mario could enter a separate mode within iTunes that allowed him to purchase music, using a categorization scheme that was different from the one he used to organize his own library. By 2005, the iTunes Music Store had more than 2 million songs available, a far cry from the 40 albums that Mario had in his collection to begin with. But Apple didn't stop there: soon it started selling TV shows and then movies through the (now renamed) iTunes Store. TV shows, movies, and music were presented as distinct categories within the store, and each "department" had its own categorization scheme: rock, alternative, pop, hip-hop/rap, etc. for music; kids & family, comedy, action & adventure, etc. for movies; and so on.

iTunes was not just where Mario listened to and organized his music anymore. Now it was where he went to:

- Buy, rent, and watch movies
- Buy, rent, and watch TV shows
- Preview and buy music
- Buy applications for his iPod
- Search for and listen to podcasts
- Browse and subscribe to "iTunes U" university courses
- Listen to streaming radio stations
- Listen to audiobooks
- Browse and listen to music shared by others in his household

Each of these functions introduced new content types with particular categorization schemes. iTunes still had a search box, as it had on day one, but search results were now much more difficult to parse, because they included different (and incompatible) media types. Was the result for "Dazed and Confused" referring to the movie, the movie soundtrack, the Led Zeppelin song, or one of its myriad covers?

Later, when Mario bought his first iPhone, he was surprised to discover that the functionality that he was used to having in iTunes on the Mac (music, movies, TV shows, podcasts, etc.) had now been "unbundled" into various apps, as shown in Figure 1-4. On the iPhone, iTunes is not where you play music; for that there is an app called (appropriately) "Music." However, there are no "Movies" or "TV Shows" apps; there is one app ("Videos") that plays both. This is not where Mario can see the videos he has shot himself, though; for that he has to go to the "Photos" app. There is also an app on the phone where Mario can buy movies, music, and TV shows, called "iTunes Store"—the only reference to iTunes on the phone—and another where he can buy iPhone apps, called "App Store." All of these apps offer functionality that is available within iTunes on the Mac, and all of them have different content organization structures. Later on, Apple introduced a service called iTunes Match, which allowed Mario to upload his music collection to Apple's "cloud"; now he also had to keep track of which songs were actually on his phone and his Mac, and which were on Apple's servers.

Mario bought Apple products in part because of the company's reputation for excellent design. He'd heard that Apple "controls the hardware and the software," and was thus able to provide a unified, coherent experience across all of its products. Yet managing his media across his Mac and his iPhone was neither unified nor coherent. Also, over time, Mario became a consumer *and* an organizer of an information ecosystem; he had to deal with the information structures designed into the system by Apple and his own organization schemes for his personal music collection, which were now transcending many device form factors and contexts. Mario couldn't quite put his finger on it, but he could tell that something big was amiss with the design of these products, even though he found them visually appealing.

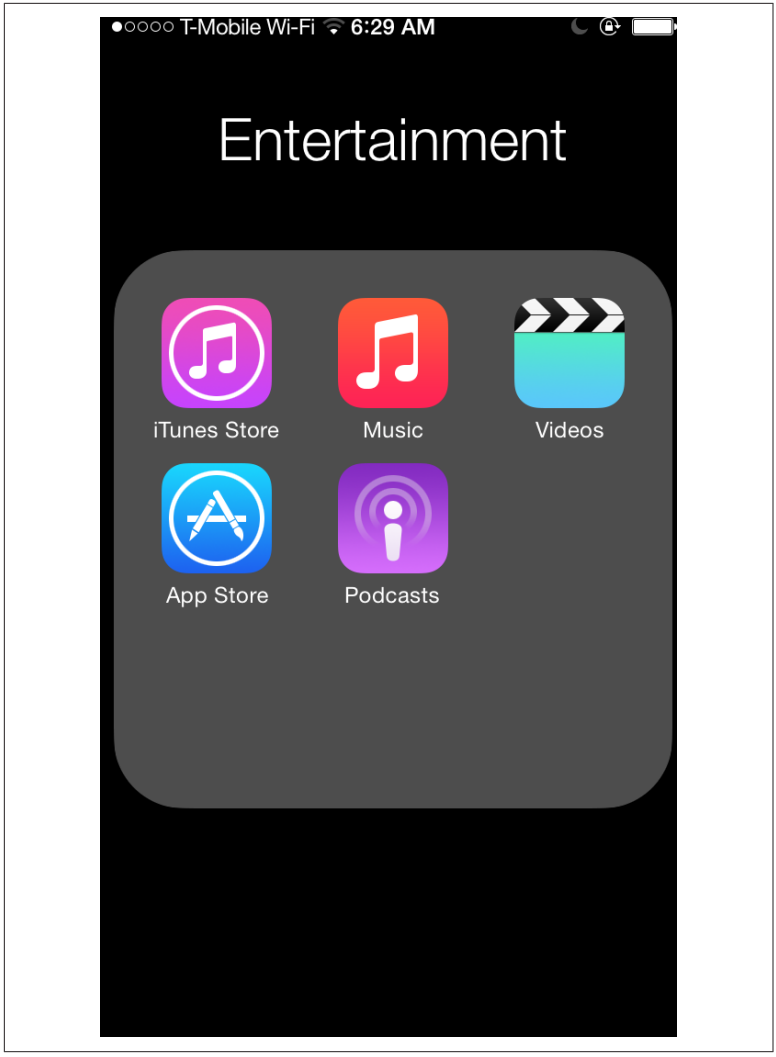

Figure 1-4. iOS's unbundled iTunes apps

The Problems Information Architecture Addresses

Mario was experiencing two problems:

- The tool he used to manage and navigate his simple library of 40 or so music albums had changed into one that dealt with hundreds of millions of different data objects of various types (songs, movies, TV shows, apps, podcasts, radio streams, university lectures, and more), each with different organization schemes, business rules (e.g., restrictions on which device he is allowed to play back his rented movie on within the next 24 hours), and ways of interacting with the information (e.g., viewing, subscribing, playing, transcoding, etc.).
- The functions provided by this tool were no longer constrained to Mario's computer; they are now available across multiple devices, including his iPhone, iPod, Apple TV, CarPlay, and Apple Watch. Each of these devices brings with it different constraints and possibilities that define what they can (and cannot) do with these information structures (e.g., "Siri, play 'With a Little Help from My Friends'"), and Mario doesn't experience them as a consistent, coherent interaction model.

Let's look at these challenges in a bit more detail.

Information Overload

People have been complaining about having to deal with too much information for centuries. As far back as Ecclesiastes (composed in the 3rd or 4th century BCE), we read that "of making many books there is no end." However, the information technology revolution that started around 70 years ago has greatly increased the information available to us. The phrase "information overload" was popularized by futurist Alvin Toffler in the 1970s.[1] Toffler called out the increased rate and pace of information production, and the resulting reduction in the signal-to-noise ratio, as problems that we'd have to deal with in the future. (As you can see from Mario's example, this

1 Alvin Toffler, *Future Shock* (New York: Random House, 1970).

future is now!) The career of Richard Saul Wurman—originator of the term "information architect"—is based on using design to address information overload. His book *Information Anxiety*[2] is considered a classic in the field.

In the 19th and 20th centuries, electronic media such as the telegraph, telephone, radio, and television allowed more information to reach more people over greater distances than ever before. However, the process really sped up in the second half of the 20th century with the appearance of digital computers and their eventual connection into what became the Internet. Suddenly, massive amounts of information could be shared with anyone in the world. The Internet—and the World Wide Web, especially—were conceptualized as two-way, interactive media. For example, you could not only receive email, but also send it. Sir Tim Berners-Lee meant for the Web to be a read/write medium; the first web browser, called WorldWideWeb (with no spaces), gave as much prominence to editing web pages as it did to browsing them. Compared to previous information media, publishing on the Web was fast, cheap, and efficient. As a result, the amount of information being published today in information environments like Facebook, Twitter, and WordPress dwarfs anything that has ever come before.

It's important to note that while every advance in information technologies has increased the overall amount of information available and has made it possible for more people to publish and have access to information, the resulting glut has also led to the creation of new technologies to help people organize, find, and make better use of information. For example, the invention of the movable type printing press in the 15th century made more books and pamphlets available more cheaply to more people. This, in turn, led to the creation of technologies such as encyclopedias, alphabetic indexes, and public libraries, which allowed people to better manage and make sense of the new information sources.[3]

It should not be surprising, then, that some of the great success stories of the early Web, such as Google and Yahoo!, were companies

2 Richard Saul Wurman, *Information Anxiety* (New York: Bantam, 1989).

3 For more on this topic, see Ann Blair's Boston Globe article "Information overload, the early years".

founded to help users find information online.[4] Still, there is much more information out there than we can manage, and the findability techniques that were effective in the late 1990s (e.g., Yahoo!'s curated hierarchical directory) are ineffective today.

With the rise of app-centric Internet-connected mobile devices such as smartphones, it has become fashionable for pundits to postulate the demise of the World Wide Web. However, instead of making the Web irrelevant, these devices have given more people access to the information available on the Internet. For many applications, the data sources that feed apps tend to be indistinguishable from (if not identical to) those that power the Web. If anything, the mobile revolution has increased access to the information available in the world.

So, back to Mario. Instead of the 400 or so songs in his record collection, he can now peruse the iTunes Store's collection of 37 million songs. Not that he can flip through it like he could with his CDs (or even at his local Tower Records[5]); here, he's going to need a bit of help to find what he's looking for.

More Ways to Access Information

While the information explosion has been happening for a long time, the second problem Mario faces is newer: the relentless miniaturization of electronics, combined with widespread adoption of wireless communications technologies, has resulted in a proliferation of small, inexpensive Internet-connected devices that are transforming the way that we interact with information and with one another.

As we mentioned earlier, there was a time when information existed in a tightly coupled relationship with the artifacts that conveyed that information. Recall Marla's record collection. The music in her copy of *Sgt. Pepper's* was set into a singular vinyl disc that sat on her shelf. Marla's copy was a reproduction: many more people had similar vinyl discs with that particular music on it. However, this particular container (the disc) and the information (the music) were irrevocably tied together after being manufactured.

---

4 Google's stated mission is to "organize the world's information and make it universally accessible and useful."

5 R.I.P.

computer keyboard in a quiet office than one tapped into a five-inch glass screen in a football stadium or spoken into a car's Bluetooth audio system while driving at 50 miles per hour. Increasingly, organizations have to consider how users will access their information in these and many other wildly different contexts. They will obviously want these experiences to be consistent and coherent regardless of where and how the information is being accessed.

So, Mario is not only faced with finding new music to listen to from a collection of over 37 million songs; he's having to do so using multiple devices—notebook computer, smartphone, TV set-top box, and more—that provide very different ways of interacting with the information, and in a wide variety of different contexts. Mario is going to need a lot of help from the people who design these products and services.

Enter Information Architecture

Part of the reason Mario is confused is that while most software applications are designed to solve very specific problems, the successful ones tend to outgrow their problem-set boundaries to encompass more and more functionality over time. As a result, they lose clarity and simplicity. As we saw, while iTunes started its life as a tool to enable the digitization and management of music collections in personal computers, it grew to become a *media platform* that encompasses the original music ripping, playing, and organizing functionalities plus other media types (movies, podcasts, audiobooks, university courses, other software applications), other modes of access (buying, renting, streaming, subscribing, sharing), and various device/interaction paradigms (Microsoft Windows computers, iPods, iPads, Apple Watches, Apple TVs). In other words, iTunes went from being a tool to being an *ecosystem*.

Given the information and device class proliferation we mentioned earlier, this is a situation many organizations are already struggling with. What is needed is a systematic, comprehensive, holistic approach to structuring information in a way that makes it easy to find and understand—regardless of the context, channel, or medium the user employs to access it. In other words, someone needs to step out of the product development trenches and look at the broader picture in the abstract, to understand how it all fits together so that information can be easier to find and to understand. Information

architecture can be used as a lens to help teams and individuals gain this perspective.

Places Made of Information

As we've said before, the experience of using digital products and services is expanding to encompass multiple devices in different places and times. It's important to recognize that we interact with these products and services through the use of language: labels, menus, descriptions, visual elements, content, and their relationships with one another create an environment that differentiates these experiences and facilitates understanding (or not!). For example, the language employed by a recipe app on a mobile phone is bound to be different from that employed by an auto insurance company's website. These differences in language help define them as distinct "places" that people can visit to accomplish certain specific tasks: they create a frame for the information they convey, allowing us to understand it relative to concepts we already know.

In his book *Understanding Context*, information architect Andrew Hinton argues that we make sense of these experiences much like we do physical places: by picking up on particular words and images that define what can and can't be done in the environment—be it an idyllic open field in the English countryside or a web search engine. Digital experiences are new (and very real) types of places made of information; the design challenge lies in making them be coherent across multiple contexts. As Andrew says, "Information architecture is a discipline well-suited for attending to these challenges. It has been working with them in one way or another for decades."[6]

6 Andrew Hinton, *Understanding Context* (Sebastopol, CA: O'Reilly, 2014), 252.

Coherence Across Channels

How does information architecture achieve this coherence? To begin with, it does so by asking designers to think about these challenges in the abstract. Where other design disciplines are focused on specific instances of an artifact—the label on a bottle of detergent, the look and feel of an app's user interface—information architecture asks designers to define semantic structures that can be instantiated in multiple ways depending on the needs of different channels. A navigation structure that works well in a desktop web page should function differently when presented on a five-inch touchscreen, but the user's experience with both should be coherent (Figure 1-7).

In their landmark book *Pervasive Information Architecture*, Andrea Resmini and Luca Rosati argue for consistency as a critical component of what they call a *pervasive information architecture*—that is, one that is experienced across multiple channels and contexts. As they explain it:

> Consistency is the capability of a pervasive information architecture to serve the contexts it is designed for (internal consistency), and to preserve this logic across different media, environments, and uses (external consistency)...Consistency needs to be designed with the context it is addressing clear in mind, and in respect to the several media and environments that the service or process will span.[7]

In other words, when an organization serves its users via multiple channels, the users' experiences across those channels should be consistent and familiar. For example, a person using a bank's mobile app should experience consistent semantic structures when using the bank's website or calling the bank's phone-based service. While the capabilities and limitations of each channel are different, the semantic structures employed in each of them should be familiar and consistent. In order for this to happen, they must be abstracted from actual implementations.

7 Andrea Resmini and Luca Rosati, *Pervasive Information Architecture: Designing Cross-Channel User Experiences* (Burlington, MA: Morgan Kaufmann, 2011), 90.

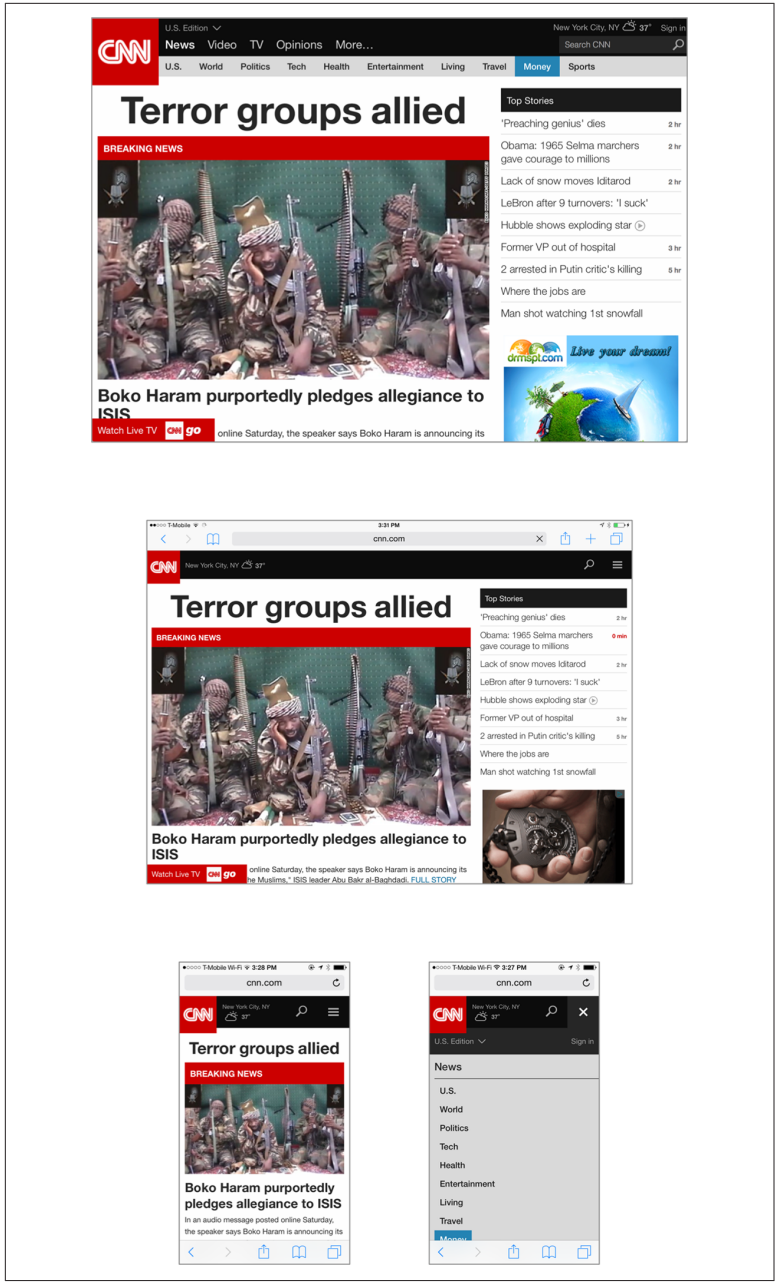

Figure 1-7. CNN's website uses a responsive layout that adapts page elements to fit different screen sizes, while offering a coherent experience

Systems Thinking

Because of this emphasis on abstracting solutions to complex challenges, information architecture also requires that the designer think systemically about the problems at hand. Where other design disciplines focus on the design of particular artifacts, information architecture is concerned with defining the semantic systems that the individual artifacts—apps, websites, voice interfaces, etc.—will be working within. Peter's book *Intertwingled* is an impassioned plea for systems thinking in the design of complex information environments. He calls out the dangers of low-level thinking when trying to design these new types of products and services:

> In the era of ecosystems, seeing the big picture is more important than ever, and less likely. It's not simply that we're forced into little boxes by organizational silos and professional specialization. We like it in there. We feel safe. But we're not. This is no time to stick to your knitting. We must go from boxes to arrows. Tomorrow belongs to those who connect.[8]

You can't design products and services that work effectively and coherently across various interaction channels if you don't understand how they influence and interact with one another and with various other systems that affect them. As mentioned earlier, each interaction channel brings to the mix different limitations and possibilities that should inform the whole. A high-level, comprehensive understanding of the ecosystem can help ensure that its constituent elements work together to present coherent experiences to users. As a discipline, information architecture is ideally suited to this task.

That said, the focus of information architecture is not only on high-level, abstract models: the design of products and services that are findable and understandable requires the creation of many low-level artifacts as well. Traditionally, many people think of website navigation structures when they think of information architecture, and this view isn't entirely off: navigation menus and their ilk are certainly within the remit of what information architecture produces. It's just that you can't get there without having explored the more abstract territory first. Effective information environments strike a balance between structural coherence (high-level invariance) and

8 Peter Morville, *Intertwingled: Information Changes Everything* (Ann Arbor, MI: Semantic Studios, 2014), 5.

suppleness (low-level flexibility), so well-designed information architectures consider both.

Having a systems-level view that is informed by (and that informs) day-to-day design activities is also a good way of ensuring that you are solving the right problems. In his book *Introduction to General Systems Thinking*, computer scientist Gerald Weinberg uses the following story to illustrate what he calls *fallacies of absolute thought*:

> A minister was walking by a construction project and saw two men laying bricks. "What are you doing?" he asked the first.
>
> "I'm laying bricks," he answered gruffly.
>
> "And you?" he asked the other. "I'm building a cathedral," came the happy reply.
>
> The minister was agreeably impressed with this man's idealism and sense of participation in God's Grand Plan. He composed a sermon on the subject, and returned the next day to speak to the inspired bricklayer. Only the first man was at work.
>
> "Where's your friend?" asked the minister.
>
> "He got fired."
>
> "How terrible. Why?"
>
> "He thought we were building a cathedral, but we're building a garage."[9]

So ask yourself: am I designing a cathedral or a garage? The difference between the two is important, and it's often hard to tell them apart when your focus is on laying bricks. Sometimes—as in the case of iTunes—designers start working on a garage, and before they know what's happening, they've grafted an apse, choir, and stained-glass windows onto it, making it hard to understand and use. Information architecture can help ensure that you're working on the plans for a great garage (the best in the world!)—or a cathedral, if such is the problem you're trying to solve. In the rest of the book, we'll show you how.

9 Gerald Weinberg, *An Introduction to General Systems Thinking* (New York: Dorset House, 2001) 61.

Recap

Let's recap what we've learned thus far:

- Historically, information has shown a tendency to dematerialize, going from having a one-to-one relationship with its containers to being completely detached from its containers (as is the case with our digital information).
- This has had two important effects in our time: information is more abundant than ever before, and we have more ways of interacting with it than ever before.
- Information architecture is focused on making information findable and understandable. Because of this, it is uniquely well suited to address these issues.
- It does this by asking the designer to think about problems through two important perspectives: that our products and services are perceived as places made of information, and that they function as ecosystems that can be designed for maximum effectiveness.
- That said, information architecture doesn't operate solely at the level of abstractions: for it to be effective, it needs to be defined at various levels.

In Chapter 2, we will give you a deeper overview of the discipline of IA, and will have a shot at defining the damned thing.[10]

10 "Defining the damned thing"—or DTDT, as it is often shortened on Twitter and mailing lists—is an ongoing source of contention in the IA community, to the merriment of some and annoyance of others. When you make a living labeling things, squabbles about conceptual boundaries are an occupational hazard.

CHAPTER 2

Defining Information Architecture

We say nothing essential about the cathedral when we speak of its stones.
—Antoine de Saint-Exupéry

In this chapter, we'll cover:

- A working definition (or four!) of information architecture
- Why it's so hard to point to something and say, "that's a great IA"!
- A model for effective IA design

If you're new to information architecture, at this point you may be wondering what this is all about. This chapter has answers for you! And if you have been working in one of the UX design disciplines for a while, you may be thinking, "But isn't information architecture about making sitemaps, wireframes, and website navigation menus?" Well, yes—these are important elements of information architecture design. But there is much more to this story! In this chapter, we'll give you a broader picture of what information architecture is—and isn't.

Definitions

Let's start by clarifying what we mean by information architecture:

1. The structural design of shared information environments
2. The synthesis of organization, labeling, search, and navigation systems within digital, physical, and cross-channel ecosystems
3. The art and science of shaping information products and experiences to support usability, findability, and understanding
4. An emerging discipline and community of practice focused on bringing principles of design and architecture to the digital landscape

Were you expecting a single definition? Something short and sweet? A few words that succinctly capture the essence and expanse of the field of information architecture? Keep dreaming!

The reason we can't serve up a single, all-powerful, all-purpose definition is a clue to understanding why it's so hard to design good digital products and services. We're talking about the challenges inherent in language and representation. No document fully and accurately represents the intended meaning of its author. No label or definition totally captures the meaning of a document. And no two readers experience or understand a particular document or definition or label in quite the same way. The relationship between words and meaning is tricky at best.[1] And here's the paradox of defining information architecture: by defining and clarifying semantic concepts, IA makes them more understandable and findable, but at a cost, because definitions are so imperfect and limiting at the same time. The definition of IA itself is a great illustration of this paradox.

We'll now descend from our philosophical soapbox and get down to basics. Let's expand on our definitions to explore some basic concepts of information architecture:

1 For a humorous perspective on the trickiness of the English language, see Bill Bryson's *The Mother Tongue: English and How It Got That Way* (New York: William Morrow, 1990).

Information

We use the term "information" to distinguish information architecture from data and knowledge management. Data is facts and figures. Relational databases are highly structured and produce specific answers to specific questions. Knowledge is the stuff in people's heads. Knowledge managers develop tools, processes, and incentives to encourage people to share that stuff. Information exists in the messy middle. With information systems, there's often no single "right" answer to a given question. We're concerned with information of all shapes and sizes: websites, documents, software applications, images, and more. We're also concerned with metadata: terms used to describe and represent content objects such as documents, people, processes, and organizations.

Structuring, organizing, and labeling

Structuring involves determining the appropriate levels of granularity[2] for the information "atoms" in your product or service, and deciding how to relate them to one another. Organizing involves grouping those components into meaningful and distinctive categories, creating the right contexts for users to understand the environment they are in and what they're looking at. Labeling means figuring out what to call those categories and the navigation structure elements that lead to them.

Finding and managing

Findability is a critical success factor for overall usability. If users can't find what they need through some combination of browsing, searching, and asking, then the system fails. But designing for the needs of users isn't enough. The organizations and people who manage information are important, too. An information architecture must balance the needs of users with the goals of the business. Efficient content management and clear policies and procedures are essential.

Art and science

Disciplines such as usability engineering and methodologies such as ethnography bring the rigor of the scientific method to the analysis of users' needs and information-seeking behaviors.

2 Granularity refers to the relative size or coarseness of information chunks. Varying levels of granularity might include journal issue, article, paragraph, and sentence.

We're increasingly able to study patterns of usage and subsequently make improvements to our websites. But the practice of information architecture will never be reduced to numbers; there's too much ambiguity and complexity. Information architects must rely on experience, intuition, and creativity. We must be willing to take risks and trust our intuition. This is the "art" of information architecture.

Just Because You Can't See It, Doesn't Mean It Isn't There

One of the challenges people have with information architecture is that they can't easily point to it. How many times have you heard someone say, "Boy, that website's information architecture is really terrific!" or, "I can't find anything in this app! Its information architecture sucks!" Our bet is, not many. But the fact that you can't readily *see* the information architecture in things doesn't mean it's not there. As de Saint-Exupéry said, sometimes what is essential is invisible to the eye.

To illustrate, think of the game of chess. Perhaps the image that comes to your mind is of a chessboard like the one shown in Figure 2-1, with beautifully sculpted wooden pieces, and a goblet of brandy sitting near a flickering fireplace. That beautiful chessboard is a common instantiation of the game we call chess. However, chess is more than that. You could argue that what makes chess "chess" is a set of information structures that relate to one another according to predefined rules.

To begin with, chess has a taxonomy of pieces that represent army units: pawns, rooks, bishops, knights, kings, and queens. In play, there are two sets ("armies") of such pieces: "black" and "white." These armies face each other in a field that consists of an eight-by-eight grid of alternating light- and dark-colored squares. This field—the chessboard—creates a context (a "place") for the battle to take place.

Figure 2-1. A chess board with pieces in the opening position (image: http://bit.ly/opening_chess_position)

The different types of pieces can move and interact in different ways in this board; there are lots of rules that determine how the armies can interact. Differences in the pieces' range, scope, and numbers determine their relative worth to each army (Table 2-1).

Table 2-1. The different types of chess pieces, including their relative values and starting amounts

Name	Amount per army	Relative value
Pawn	8	1
Knight	2	3
Bishop	2	3
Rook	2	5
Queen	1	9
King	1	—

(The king is invaluable: its capture ends the game.)

So think back to the beautiful wooden chess set. If chess can be reduced to these basic information structures, perhaps you're suspecting that the wooden pieces and board are somewhat superfluous and that you should be able to play chess with many different types of sets. You'd be correct: in fact, chess can be played in many

different ways that do not involve carved wood—or any types of physical pieces—at all. For example, you may have heard of correspondence chess, which is played via postal mail using pen and paper (Figure 2-2).

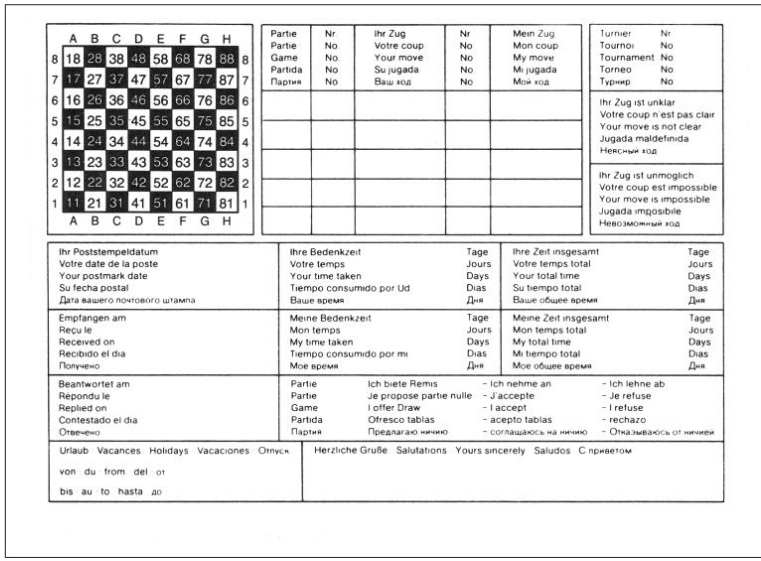

Figure 2-2. Correspondence chess postcard (image: Schach Niggemann GFDL, http://www.gnu.org/copyleft/fdl.html, or CC-BY-SA–3.0, http://creativecommons.org/licenses/by-sa/3.0/, via Wikimedia Commons)

Or perhaps you're more familiar with chess as a video game, an example of which is shown in Figure 2-3. This variant is played on a computing device with the board and pieces rendered as pixels on a screen, with the game mechanics adjusted to conform to the device's user interface particularities.

Figure 2-3. Deep Green chess, played on an iPhone with a touchscreen interface

Chess can also be played in a computer terminal console, with the most minimally symbolic user interface imaginable (Figure 2-4).

Figure 2-4. GNU Chess, played with a command-line interface

And of course, there are also countless variations of physical chess sets, ranging from our beautiful wooden set, to cheap "travel" sets with minimally rendered graphics on magnetic pieces (Figure 2-5), to the "Jewel Royale Chess Set" that is valued at almost $10 million.

These incarnations of chess are all physically very different from one another, yet they are all still chess. Why? Because they make possible and express the underlying information structures and rules of chess. Expressing and supporting these information structures is what *makes* all of these incarnations chess; their physical form and interaction mechanisms are merely matters of interaction or industrial design. In many ways, this abstract idea of chess is more "real" —but less tangible—than the physical (or virtual) chess sets that we interact with, because it is what makes chess different from other games.

Figure 2-5. Intense game of chess unfolding on a cheap magnetic travel set (image: http://bit.ly/magnetic_chess; cropped)

It's worth noting that nobody set out to explicitly create this "information architecture" of chess—the game, its piece types and rules, its lore, etc. have evolved over centuries. This is also true of the ways we've organized other information structures that afford understanding over time: it's only in retrospect that we can point to them and say, "that's a damned good information architecture!"

Toward a Damned Good Information Architecture

Users. Content. Context. You'll hear these three words again and again throughout this book. They form the basis of our model for practicing effective information architecture design. Underlying this model is a recognition that you can't design useful information architectures in a vacuum. An architect can't huddle in a dark room with a bunch of content, organize it, and emerge with a grand solution. It simply won't hold up against the light of day.

Websites, intranets, apps, and other information environments are not lifeless, static constructs. Rather, there is a dynamic, organic nature to both the information systems and the broader contexts in which they exist. This is not the old world of yellowing cards in a library card catalog. We're talking complex, adaptive systems with emergent qualities. We're talking rich streams of information flowing within and beyond the borders of departments, business units, institutions, and countries. We're talking messiness and mistakes, trial and error, survival of the fittest.

We use the concept of an "information ecology"[3] composed of users, content, and context to address the complex dependencies that exist in these information environments. And we draw upon our trusty Venn diagram (see Figure 2-6) to help people visualize and understand these relationships. The three circles illustrate the interdependent nature of users, content, and context within a complex, adaptive information ecology.

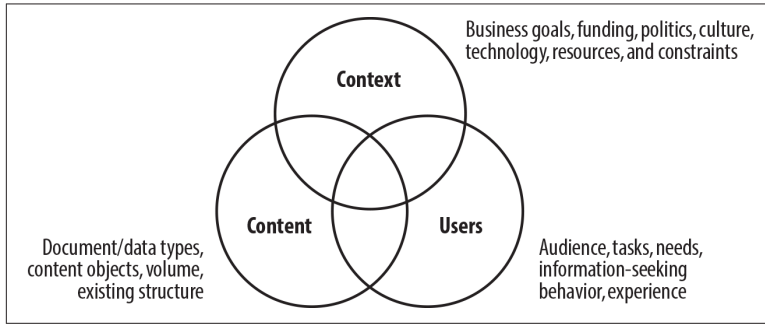

Figure 2-6. The infamous three circles of information architecture

In short, we need to understand the business goals behind the project and the resources available for design and implementation. We need to be aware of the nature and volume of content that exists today and how that might change a year from now, and we must learn about the needs and information-seeking behaviors of our major audiences.

3 For more about information ecologies, read *Information Ecology* by Thomas Davenport and Lawrence Prusak (Oxford: Oxford University Press, 1997) and *Information Ecologies* by Bonnie Nardi and Vicki O'Day (Cambridge, MA: MIT Press, 1999). Nardi and O'Day define an information ecology as "a system of people, practices, values, and technologies in a particular local environment."

Good information architecture design is informed by all three areas, and all three are moving targets. Users can vary in their attitude, demographics, psychographics, tasks and information needs, information-seeking behaviors, and more. Content can vary in quality, currency, authority, popularity, strategic value, cost, and more. And organizational context can vary based on mission, vision, goals, organizational politics, organizational culture, degree of centralization or autonomy, and more. The particular mix of variables differs from one information environment to another, and within the same environment it varies over time.

Even so, this is an oversimplified view of reality. Is it still useful? Absolutely. We've been using this model for 20 years. It's held up well in all sorts of environments, from global websites of Fortune 100 corporations to standalone intranet applications within small nonprofits. More importantly, we find these three circles incredibly helpful whenever we're confronted by a difficult question. After mouthing the trusty phrase "It depends"—as all smart practitioners of information architecture do—we develop our answer by deconstructing the question into three parts that coincide with our three circles. When asked what are the most important qualities that we should bring to the table, the answer becomes quite simple: some knowledge of users and their needs (which might come from exposure to human–computer interaction and a variety of other fields), content (think technical communication and journalism), and context (read a book on organizational psychology).

The three circles help with other tough questions, too, such as:

- What research and evaluation methods should we be familiar with?
- What kinds of people should be part of the team that designs the information architecture?
- What kinds of books and blogs should we read to keep up with the field and its practice?
- What should go into the IA strategy that we propose to a new prospect?

The answer to each starts with a balance among the three areas: users, content, and context.

Should technology have its own circle? Maybe. But we find that technology usually gets too much attention. Also, we increasingly find that much of what falls under the rubric of technology can be expressed within the "context" circle. After all, what technology brings to the table are new possibilities and constraints that give shape to the final product, and this is squarely within the realm of the context we're designing for.

Incidentally, we think it's important to have a good sense of humor about this stuff. Perhaps you've already figured this out. The work we do involves high levels of abstraction, ambiguity, and occasionally absurdity, and to some degree we're all still making it up as we go along.

If there's one thing that many years of information architecture consulting has taught us, it's that every situation is unique. We don't just mean that websites are different from intranets or that extranets should vary by industry. We mean that, like fingerprints and snowflakes, every information ecology is unique. The Toyota intranet is vastly different from that of Ford or GM. Fidelity, Vanguard, Schwab, and E*TRADE have each created unique online financial-service experiences. Despite all the copycatting, benchmarking, and definitions of industry best practices that have surged throughout the business world in recent years, each of these information systems has emerged as quite distinctive.

That's where our model comes in handy. It's an excellent tool for learning about the specific needs and opportunities presented by a particular project. Let's take a look at how each of our three circles contributes to the emergence of a totally unique information ecology.

Context

All digital design projects exist within a particular business or organizational context. Whether explicit or implicit, each organization has a mission, goals, strategy, staff, processes and procedures, physical and technology infrastructure, budget, and culture. This collective mix of capabilities, aspirations, and resources is unique to each organization.

Because of this, information architectures must be uniquely matched to their contexts. The vocabulary and structure of your websites and your apps is a major component of the evolving con-

versation between your business and your customers and employees. It influences how they think about your products and services. It tells them what to expect from you in the future. It invites or limits interaction between customers and employees. Your information architecture provides perhaps the most tangible snapshot of your organization's mission, vision, values, strategy, and culture. Do you really want that snapshot to look like that of your competitor?

The key to success is understanding and alignment. First, you need to understand the business context. What makes it unique? Where is the business today, and where does it want to be tomorrow? In many cases, you're dealing with tacit knowledge. It's not written down anywhere; it's in people's heads and has never been put into words. We'll discuss a variety of methods for extracting and organizing this understanding of context. Then, you need to find ways to align the information architecture with the goals, strategy, and culture of the business. We'll discuss the approaches and tools that enable this custom configuration.

As mentioned in Chapter 1, you also need to understand the contextual differences imposed by the channels that the user will be using to interact with your organization. Will they be experiencing your services primarily via apps on mobile phones, or via a website in a desktop-based browser? Both platforms have things they can do well, and things they can't. For example, smaller screens mean less space, which in turn implies shorter labels and navigation menus. Devices with small screens are also used at different times and places than those with larger screens. If your service will be used via more than one channel, you need to consider how these channels will overlap and interact with one another. All of these factors form part of the context that will shape your information architecture.

Content

We define "content" very broadly to include the documents, applications, services, schemas, and metadata that people need to use or find in your systems. To employ a technical term, it's the "stuff" that makes up your sites and apps. Many digital systems are heavily textual; among other things, the Web is a wonderful communication tool, and communication is built upon words and sentences trying to convey meaning. Of course, we also recognize it as a tool for tasks and transactions; a flexible technology platform that supports buying and selling, calculating and configuring, sorting and simulating.

But even the most task-oriented ecommerce website has "content" that customers must be able to find.

As you survey content across a variety of digital systems, the following facets will bubble to the surface as distinguishing factors of each information ecology:

Ownership
: Who creates and owns the content? Is ownership centralized within a content authoring group or distributed among functional departments? How much content is licensed from external information vendors? How much is produced by the users themselves? The answers to these questions play a huge role in influencing the level of control you have over all the other dimensions.

Format
: Websites and intranets have become the unifying means of access to all digital formats within many organizations. Databases, product catalogs, discussion archives, technical reports in MS Word, annual reports in PDF, office supply purchasing applications, and video clips of the CEO are just a few of the types of documents, databases, and applications you'll find on a given site.

Structure
: All documents are not created equal. An important memo may be fewer than 100 words. A technical manual may be more than 1,000 pages. Some information systems are built around the document paradigm, with the fully integrated document as the smallest discrete unit. Other systems take a content component or digital asset approach, leveraging some form of structural markup (e.g., XML or JSON) to allow management and access at a finer level of granularity.

Metadata
: To what extent has metadata that describes the content and objects within your system already been created? Have documents been tagged manually or automatically? What's the level of quality and consistency? Is there a controlled vocabulary in place, or have users been allowed to tag the content? These factors determine the extent to which you're starting from scratch

with respect to both information retrieval and content management.

Volume
: How much content are we talking about? A hundred applications? A thousand pages? A million documents? How big is the system?

Dynamism
: What is the rate of growth or turnover? How much new content will be added next year? And how quickly will it go stale?

All of these dimensions make for a unique mix of content and applications, which in turn suggests the need for a customized information architecture.

Users

The most important thing to know about users is that when we are talking about "users" we are talking about *people*. These are human beings with desires, needs, concerns, and foibles—just like you and us. We use the word "users" as shorthand to mean "the people who will use your information environment."

When we worked on the first corporate website for Borders Books & Music, back in the mid-1990s before Amazon became a household name, we learned a lot about how customer research and analysis was applied to the design and architecture of physical bookstores.

Borders had a clear understanding of how the demographics, aesthetic preferences, and purchasing behaviors of its customers differed from those of its main competitor, Barnes & Noble. It was no mistake that the physical layout and the selection of books differed significantly between these two stores, even within the same town. They were different by design. And that difference was built upon an understanding of their unique customer or market segments.

Differences in customer preferences and behaviors within the physical world translate into different information needs and information-seeking behaviors in the context of websites and apps. For example, senior executives may need to find a few good documents on a particular topic very quickly. Research analysts may need to find all the relevant documents and may be willing to spend several hours on the hunt. Managers may have a high level of industry knowledge but low navigation and searching proficiency. Teenagers

may be new to the subject area but skilled in handling a search engine.

Do you know who's using your system? Do you know how they're using it? And perhaps most importantly, do you know what information they want from your systems? These are not questions you can answer in brainstorming meetings or focus groups. As our friend and fellow information architect Chris Farnum likes to say, you need to get out there in the real world and study your "users in the mist."

Recap

Let's recap what we've learned in this chapter:

- There's more than one way to define information architecture, and that's OK.
- Information architecture is not something you can easily point to; it is mostly abstract and exists below the surface, in the deep semantic structures of products and services. This is OK, too!
- Our model for practicing effective information architecture design considers three things: users, context, and content.
- The particular mix of variables changes not just from one information environment to another, but also for a single information environment over time.

As we mentioned in the introduction to Part I, IA is focused on making information environments *findable* and *understandable*. These are related, but different, objectives. In the next chapter, we'll look more closely at designing for findability. Onward!