

Automated Publishing at O'Reilly Media, Inc.

Andrew Savikas
andrews@oreilly.com

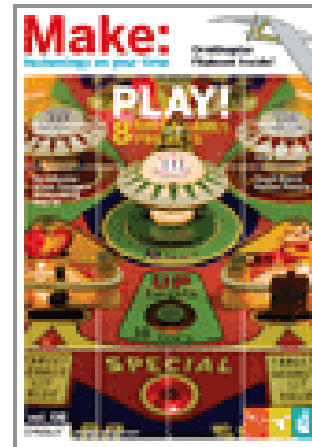
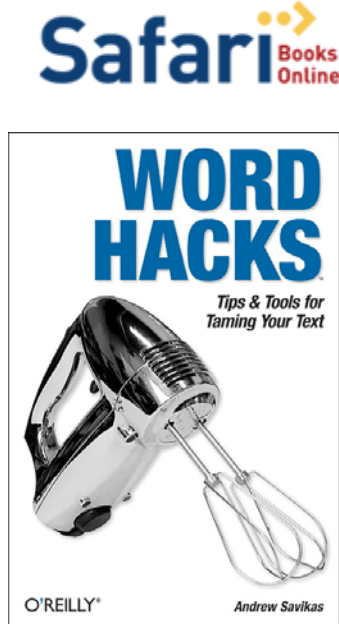
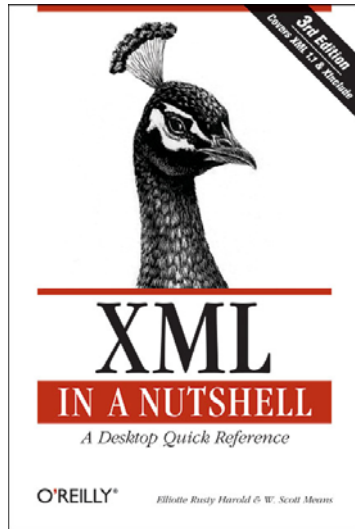
Director, Digital Content & Publishing Services

April 10, 2007 – Gilbane San Francisco

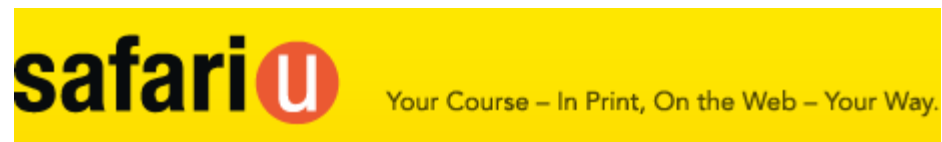


O'REILLY®

O'Reilly Media, Inc.



web 2.0
E X P O



The O'Reilly “Radar”

- **1991:** Published first book about Linux
- **1992:** Published first book about the web, when there were only 200 web sites.
- **1993:** Created GNN, the first commercial web site and first “portal” site (sold to AOL)
- **1998:** Organized the Open Source Summit, where the term "open source" was formally adopted by the leaders of most major open source projects
- **2004:** Coined the term “Web 2.0”

www.toccon.com

O'REILLY

TOC Conference
Tools of Change
for Publishing

JUNE 18-20, 2007 • SAN JOSE, CALIFORNIA



REGISTER
NOW AND
▶ SAVE!

[Home](#)

[Register Now](#)

[About TOC](#)

[Tutorials](#)

[Sessions](#)

[Speakers](#)

[Sponsors](#)

[Exhibit Hall](#)

[Media & Community
Partners](#)

[Hotel & Travel](#)

[See & Do](#)

[Wiki](#)

[News & Coverage](#)

[Newsletter](#)

[O'Reilly Conferences](#)

[O'Reilly Media](#)

Technology is transforming publishing. From the way ideas are generated to the packaging of information and delivery of products, technology and change are themes in every aspect of publishing. For publishers, these shifts are taking place so rapidly that it's challenging to keep current--let alone create new, profitable opportunities.

O'Reilly is launching its first Tools of Change for Publishing (TOC) Conference to raise the level of technology knowledge and discourse among book publishers and to spark conversation and creativity for leading change. The TOC Conference will provide insightful analysis, hands-on workshops, and engaging sessions for publishers of all kinds. Join us June 18-20, 2007 in San Jose, California for the first O'Reilly TOC Conference, and help shape the future of publishing.

Conference Speakers



Tim O'Reilly

O'Reilly Media, Inc.

Tim O'Reilly is founder and CEO of O'Reilly Media, Inc., thought by many to be the best computer book publisher in the world. In addition to publishing pioneering books like Ed Krol's *The Whole ...more*

Diamond
Sponsors

INGRAM



Gold Sponsor

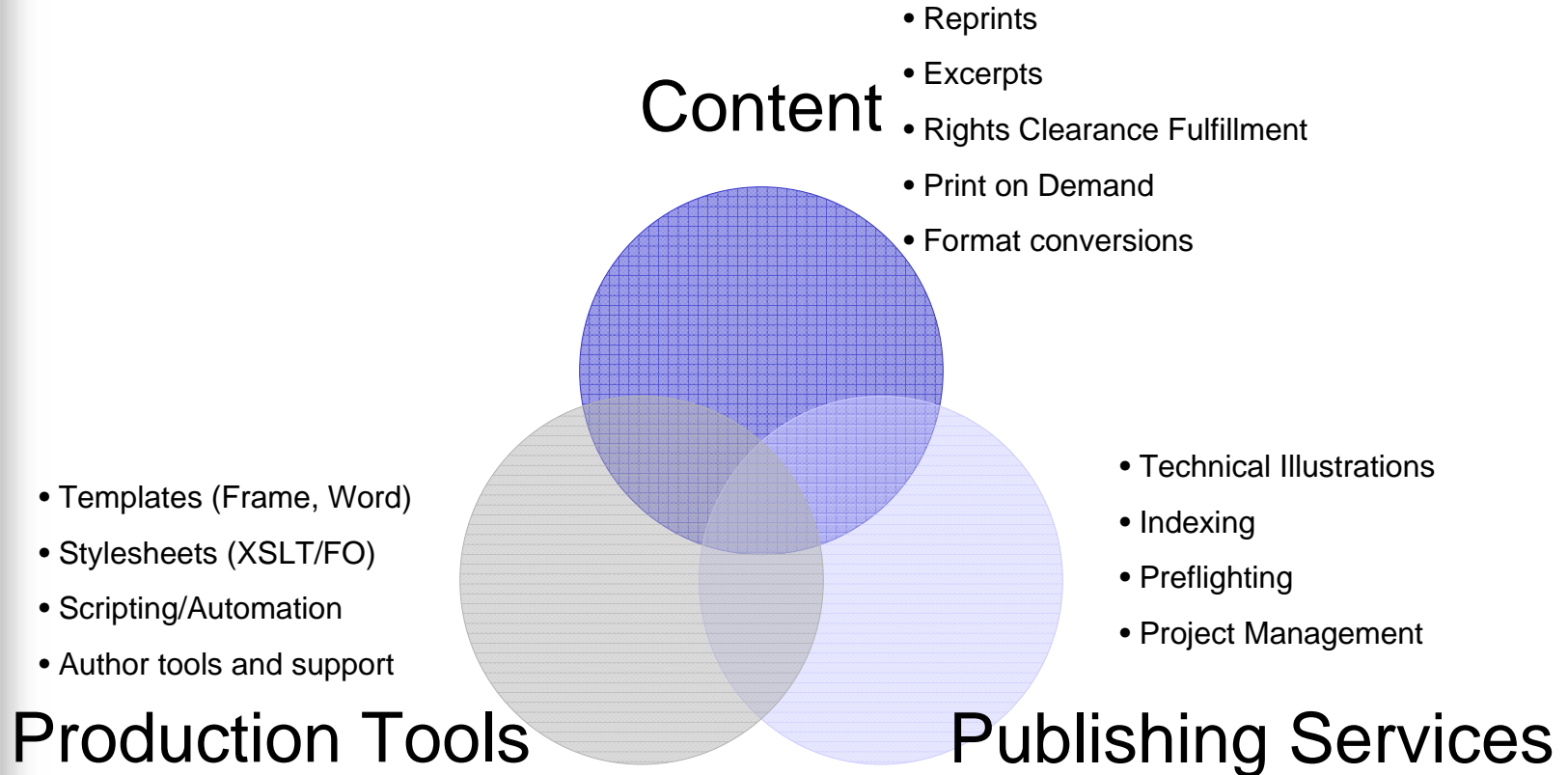


Silver Sponsors



O'REILLY®

Digital Content & Publishing Services



Automated Publishing at O'Reilly

- Safari Books Online
- SafariU
- Print Production
- “Publishing 2.0” Projects
- But, many books still produced with FrameMaker and InDesign

Building Blocks

DocBook XML

XSLT

XSL-FO

XHTML

PDF

XQuery

Safari Books Online (*safari.oreilly.com*)

The screenshot displays the Safari Books Online interface. At the top, the O'Reilly logo and 'Safari Books Online' are visible. A navigation bar includes links for Home, Network, Store, Safari Books Online, Conferences, Courses, Academic Solutions, and About. Below this, a secondary bar offers links for Safari Home, My Safari, My Favorites, My Account, Tools, Safari Labs, and Sign Out.

The main content area features a search box on the left with options for 'Entire Site', 'All Content', 'My Favorites', and 'Current Book Only'. A 'GO' button is present. To the right of the search box, there are links for 'E-Mail This Page', 'Add Note', and 'Add Bookmark'.

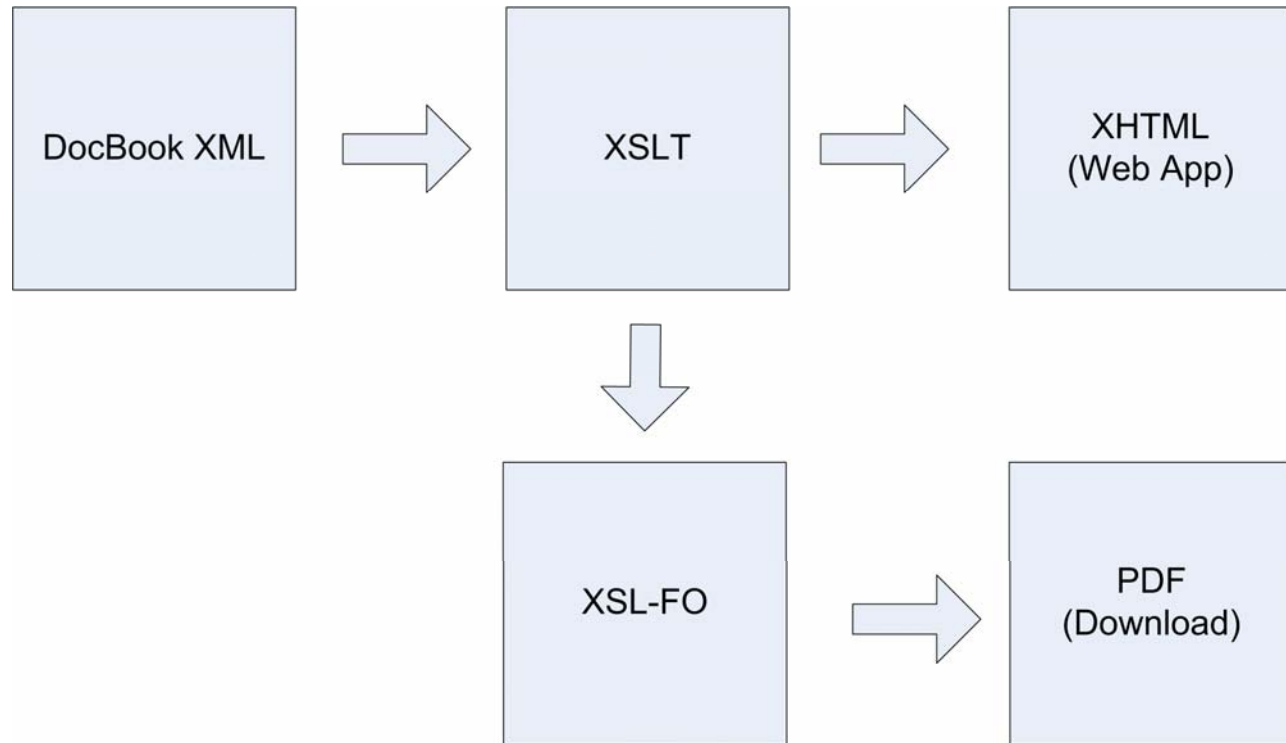
The central focus is the book 'QuickBase: The Missing Manual' by Nancy Conner. The book cover is shown on the left. To the right, the following details are provided:
Publisher: O'Reilly
Pub Date: February 01, 2007
Print ISBN-10: 0-596-52960-0
Print ISBN-13: 978-0-59-652960-4
Pages: 400

On the right side of the book details, there are three promotional buttons:
1. 'START READING ONLINE' with a right-pointing arrow.
2. 'ADD TO MY FAVORITES'.
3. 'Buy Print Version' with a '35% OFF' badge and the text 'for Safari Subscribers'. Below this is a note: 'All orders over \$29.95 qualify for FREE SHIPPING within the US*'.
Below these buttons is a 'Table of Contents' and 'Index' section.

On the left side of the book details, there is a 'Table of Contents' section with a 'Book' icon. It lists the following chapters and sections:
Copyright
Introduction
▶ Signing Up and Taking a Quick Tour
▶ Different Ways of Viewing Your Data
▶ Working with Data and Documents
▶ Reports, Change Notifications, and Reminders
▶ Planning Your QuickBase Solution
▶ Using QuickBase's Prebuilt Applications
▶ Creating Your Own Application
▶ Managing Applications
▶ Managing Roles and Groups
▶ Creating Relationships

The 'Table of Contents' and 'Index' section on the right lists the following chapters and sections:
Copyright
Introduction
▶ Chapter 1. Signing Up and Taking a Quick Tour
 Section 1.1. Creating an Account
 Section 1.2. Using Your My QuickBase Page
 Section 1.3. Getting Started with Applications
▶ Chapter 2. Different Ways of Viewing Your Data
 Section 2.1. QuickBase's View Formats
 Section 2.2. Creating, Editing, and Printing Views
 Section 2.3. Tips for Creating Specific View Types
▶ Chapter 3. Working with Data and Documents
 Section 3.1. Getting Data and Documents into QuickBase
 Section 3.2. Finding, Changing, and Exporting Data
 Section 3.3. Editing and Collaborating on Documents
▶ Chapter 4. Reports, Change Notifications, and Reminders
 Section 4.1. Sending Reports to Yourself and Others
 Section 4.2. Modifying Report Subscriptions
 Section 4.3. Triggering Change Notification Emails
 Section 4.4. Sending Reminders
 Section 4.5. Using vCard and iCalendar with QuickBase
▶ Chapter 5. Planning Your QuickBase Solution

Safari Toolchain



Customized, on-demand downloads

Troubleshooting Moment: Help! I Didn't Get My Verification Email

As soon as you click Continue in step 5 of the registration process, QuickBase's verification email wings its way through cyberspace to your email inbox. But what if you go through the registration steps, fill in all the needed info, click Continue—and then sit back to wait for an email that never appears?

The likely culprit? Your spam filter—an overzealous junk-mail catcher that probably nabbed the QuickBase verification email before you even saw it. To find the email, try one of these strategies:

- Check your email program's Junk folder.
- Add the following address to your email program's address book:
corpsales@quickbase.com.
- Try signing in to QuickBase. If your account is unverified, QuickBase's Unverified Account page appears. This page lists the email address you used to register; double-check that address to make sure you didn't make a mistake typing it in. If your email address looks OK, click the Resend Verification Email button to try again.
- If all else fails, talk to your network administrator to make sure that your network's settings aren't filtering emails that come from QuickBase.

You can't use QuickBase until you've got a verified account, so getting that verification email is an absolute must!

Accepting an Invitation

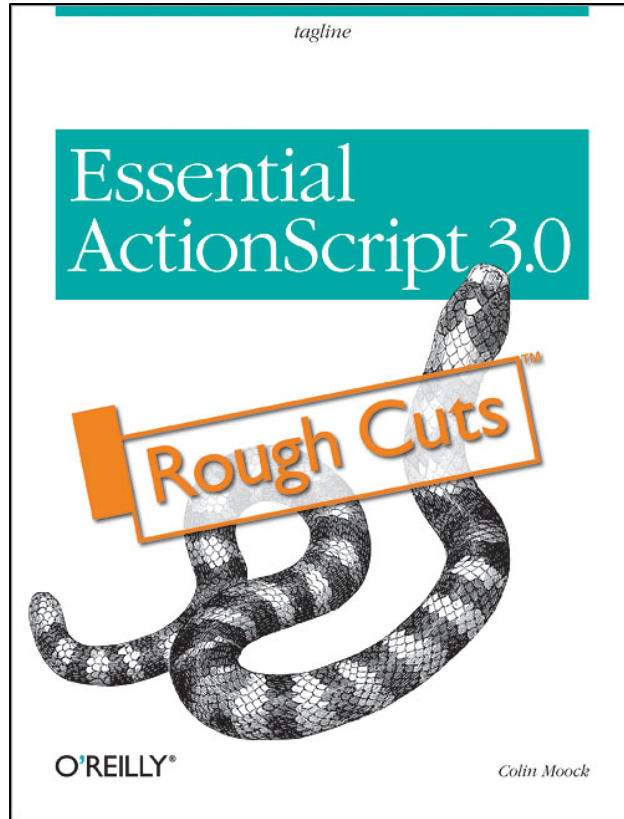
If someone with a QuickBase account wants you to start using one of their QuickBase applications, you'll get an email invitation to join QuickBase. The email has the name of the application and the words "QuickBase Invitation" in the subject line.

The body of the email contains a link to a page within QuickBase; click that link to go to a Welcome page, which tells you that QuickBase has created an account for you. Click the Continue button to register. The Registration page is identical to the one shown in [Figure 1-3](#). Enter your name and a password, and QuickBase takes you straight to the application's Dashboard page ([Creating an Application from a Template](#)). From there, you can work with the application in whatever capacity the administrator who invited you has assigned.

Note: If you're invited to work with someone else's QuickBase application, you may or may not have the ability to create your own applications. It depends on the permissions (usage rights) that you've been granted. If you don't see a Create a New Application button in the upper-right part of your My QuickBase

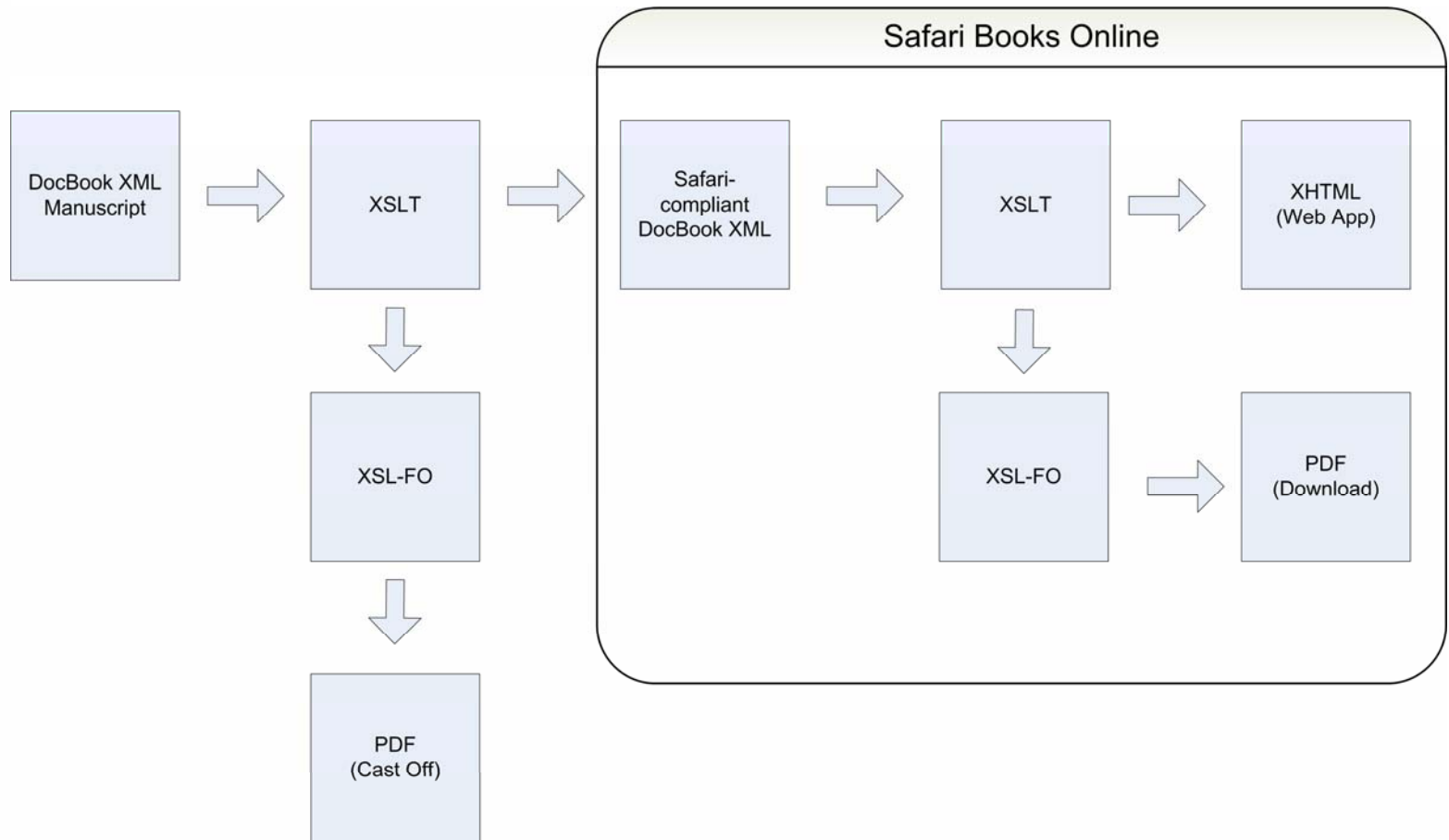
Prepared for safariproduct O'Reilly, Safari ID: safariproduct@oreilly.com
User number: 2604 Copyright 2007, Safari Books Online, LLC.
permission from the publisher is prohibited.

Rough Cuts



- Early access to hot technologies
- Customers get online and PDF access
- Analogous to regular software project “builds”
- Highly accurate “cast-off” PDFs for use by author and editor

Author-controlled updates




SafariU (safariu.com)

Current Project Title: **Untitled project**

1 Gather Content 2 Organize Content 3 Choose Options 4 Review 5 Finalize

Search Books Articles Learning Objects Public Projects for: Search [Advanced Search](#)

All Books > O'Reilly > XSLT Cookbook, 2nd Edition > XPath

 **Book: XSLT Cookbook, 2nd Edition**
Author(s): Sal Mangano
Publisher: O'Reilly
Date: 1 December 2005

[Add Entire Book to Project](#)
[Adopt This Book](#)

Table of Contents

- ▶ Copyright
- ▶ Preface
- ▶ 1. XPath
 - Introduction
 - 1.1 Effectively Using Axes
 - 1.2 Filtering Nodes
 - 1.3 Working with Sequences
 - 1.4 Shrinking Conditional Code with If Expressions
 - 1.5 Eliminating Recursion with for Expressions
 - 1.6 Taming Complex Logic Using Quantifiers
 - 1.7 Using Set Operations
 - 1.8 Using Node Comparisons
 - 1.9 Coping with XPath 2.0's Extended Type System
 - 1.10 Exploiting XPath 2.0's Extended Type System
- ▶ 2. Strings
- ▶ 3. Numbers and Math
- ▶ 4. Dates and Times
- ▶ 5. Selecting and Traversing
- ▶ 6. Exploiting XSLT 2.0
- ▶ 7. XML to Text
- ▶ 8. XML to XML
- ▶ 9. Querying XML
- ▶ 10. XML to HTML
- ▶ 11. XML to SVG
- ▶ 12. Code Generation
- ▶ 13. Vertical XSLT Application Recipes
- ▶ 14. Extending and Embedding XSLT
- ▶ 15. Testing and Debugging

Chapter 1: XPath

[Add Chapter to Project](#)

Introduction

Neo, sooner or later you're going to realize just as I did that there's a difference between knowing the path and walking the path.

Morpheus (The Matrix)

Introduction

XPath is an expression language that is fundamental to XML processing. You can no more master XSLT without mastering XPath than you can master English without learning the alphabet. Several readers of the first edition of *XSLT Cookbook* took me to task for not covering XPath. This chapter has been added partly to appease them but more so due to the greatly increased power of the latest XPath 2.0 specifications. However, many of these recipes are applicable to XPath 1.0 as well.

In XSLT 1.0, XPath plays three crucial roles. First, it is used within templates for addressing into the document to extract data as it is being transformed. Second, XPath syntax is used as a pattern language in the matching rules for templates. Third, it is used to perform simple math and string manipulations via built-in XPath operators and functions.

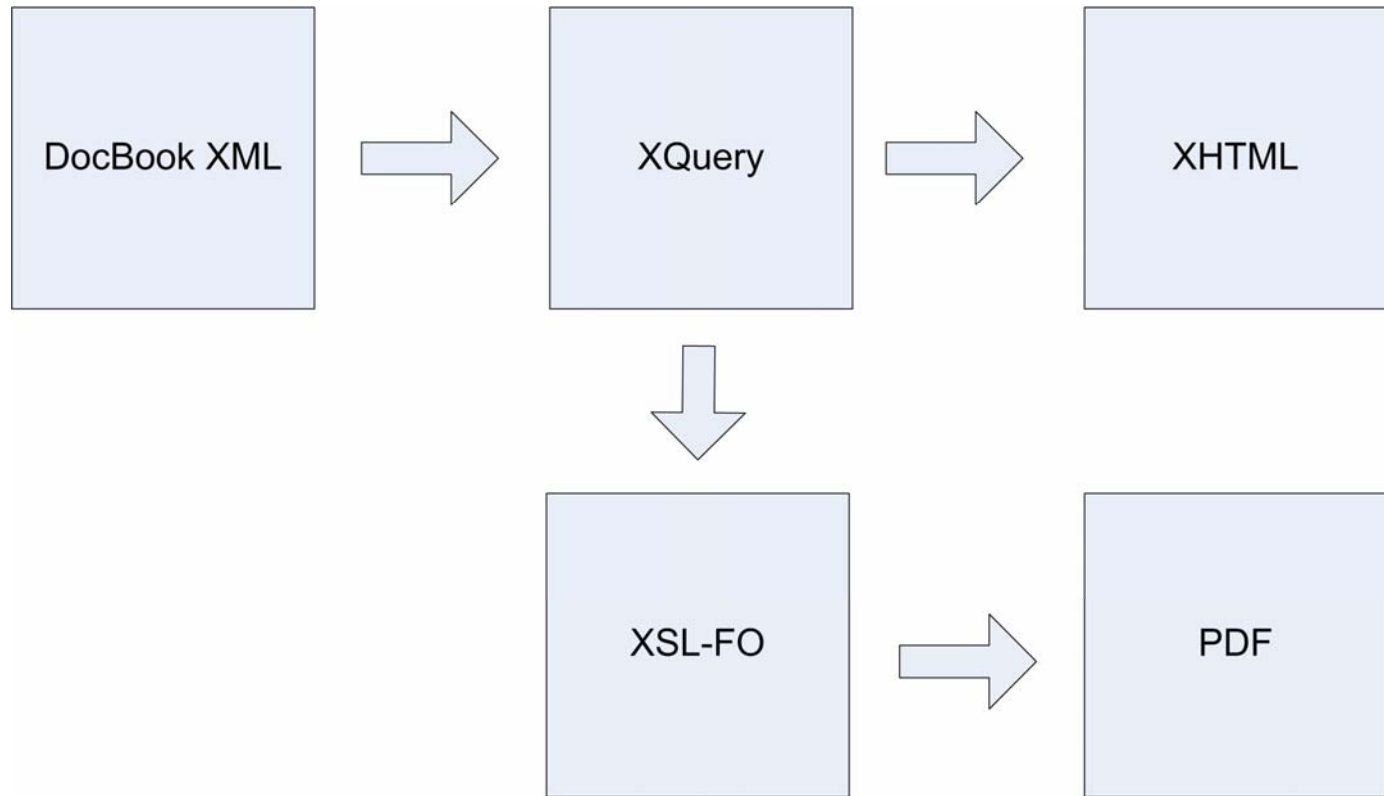
XSLT 2.0 retains and strengthens this intimate connection with XPath 2.0 by drawing heavily on the new computational abilities of XPath 2.0. In fact, one can make a reasonable argument that the enhanced capabilities of XSLT 2.0 stem largely from the advances in XPath 2.0. The new XPath 2.0 facilities include sequences, regular expressions, conditional and iterative expressions, and enhanced XML Schema compliant-type system as well as a large number of new built-in functions.

Each recipe in this chapter is a collection of mini-recipes for solving certain classes of XPath problems that often arise while using XSLT. We annotate each XPath expression with the XPath 2.0 commenting convention (: comment :) but users of XPath XSLT 1.0 should be aware that these comments are not legal syntax. When we are showing the result of an XPath evaluation that is empty, we will write (), which happens to be the way one writes a literal empty sequence in XPath 2.0.

Section 1.1: Effectively Using Axes

- Academic market
- True custom publishing
- PDF output, printed and shipped to campus bookstore
- Additional content can be included as online-only
- Students pay for books and semester access

SafariU Toolchain



Custom PDF output

```
$x << $y
```

```
(: Test if $x follows $y in document order. :)
```

```
$x >> $y
```

Discussion

The new XPath 2.0 node comparison operators are likely to be more efficient and certainly easier to understand than the XPath 1.0 counterparts. However, if you are using XSLT 2.0, you will not find too many situations where these operators are required. There are many situations where you think you need << or >> when the `xs1::for-each-group` element is preferred. See *Recipe 6.2* for examples.

Coping with XPath 2.0's Extended Type System

Problem

XPath 2.0's stricter type rules have you cursing the W3C and longing for Perl.

Solution

Most incompatibilities between XPath/XSLT 1.0 and 2.0 come from type errors. This is true regardless of whether a schema is present or not. You can eliminate many problems encountered in porting legacy XSLT 1.0 to XSLT 2.0 with respect to XPath differences by running in 1.0 compatibility mode.

```
<xsl:stylesheet version="1.0">
```

```
  <!-- ... -->
```

```
</xsl:stylesheet>
```

In my opinion, eventually you will want to stop using compatibility mode. XPath 2.0 provides several facilities for dealing with type conversions. First, you can use conversion functions explicitly.

```
(: Convert the first X child of the context to a number. :)
number(X[1]) + 17
```

```
(: Convert a number in $n to a string. :)
concat("id-", string($n))
```

XPath 2.0 also provides type constructors so you can explicitly control the interpretation of a string.

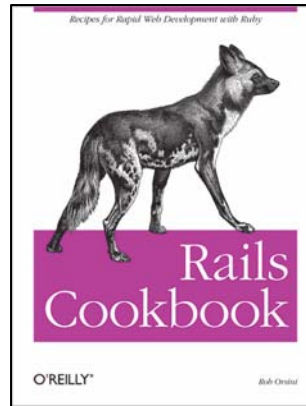
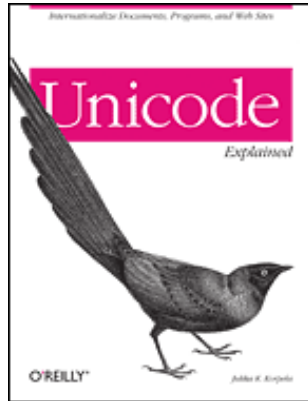
```
(: Construct a date from a string. :)
xs:date("2005-06-01")
```

```
(: Construct doubles from strings. :)
xs:double("1.1e8") + xs:double("23000")
```

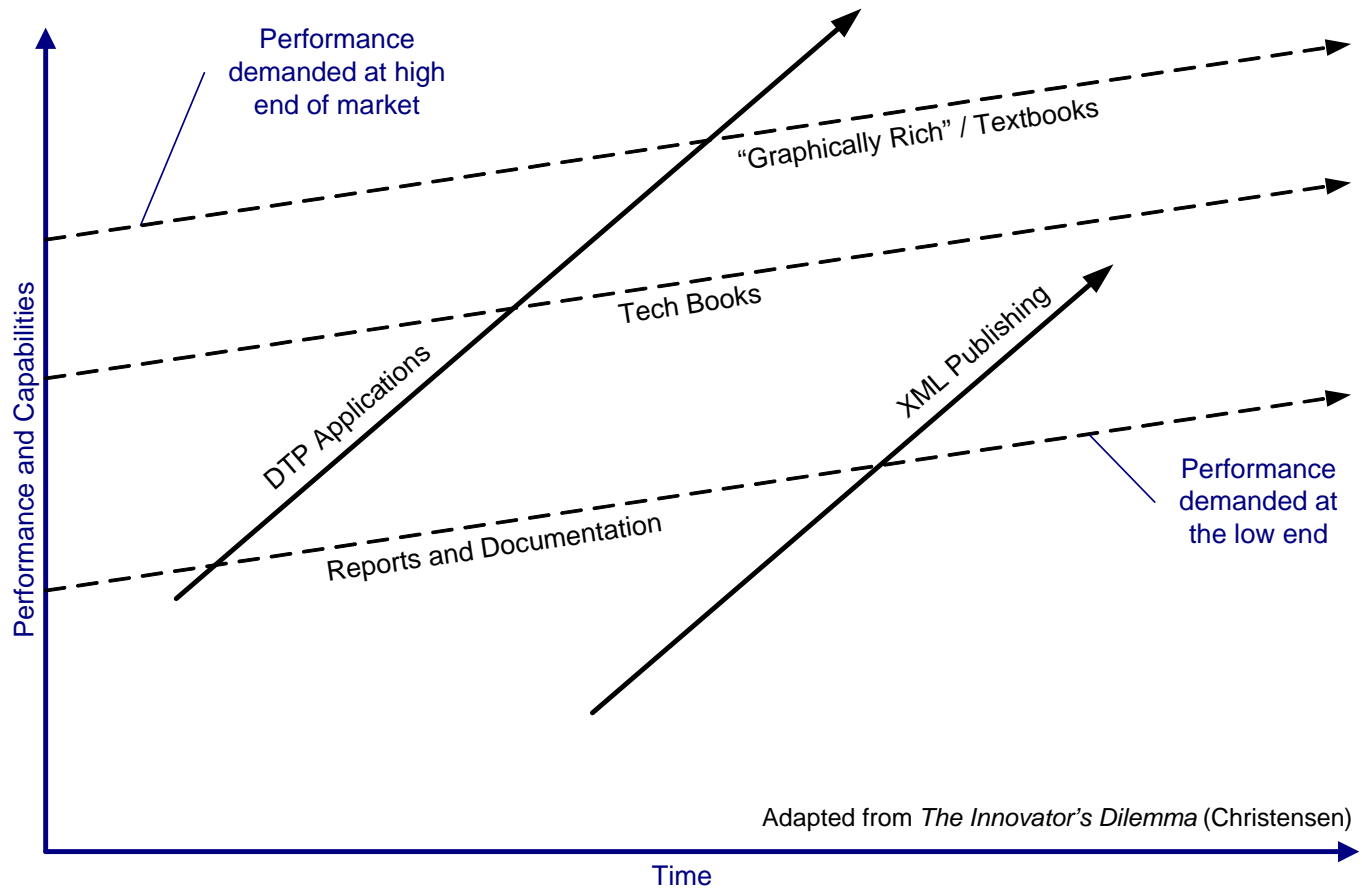
Finally, XPath has the operators `castable as`, `cast as`, and `treat as`. Most of the time, you want to use the first two.

```
if ($x castable as xs:date) then $x cast as xs:date else xs:date("1970-01-01")
```

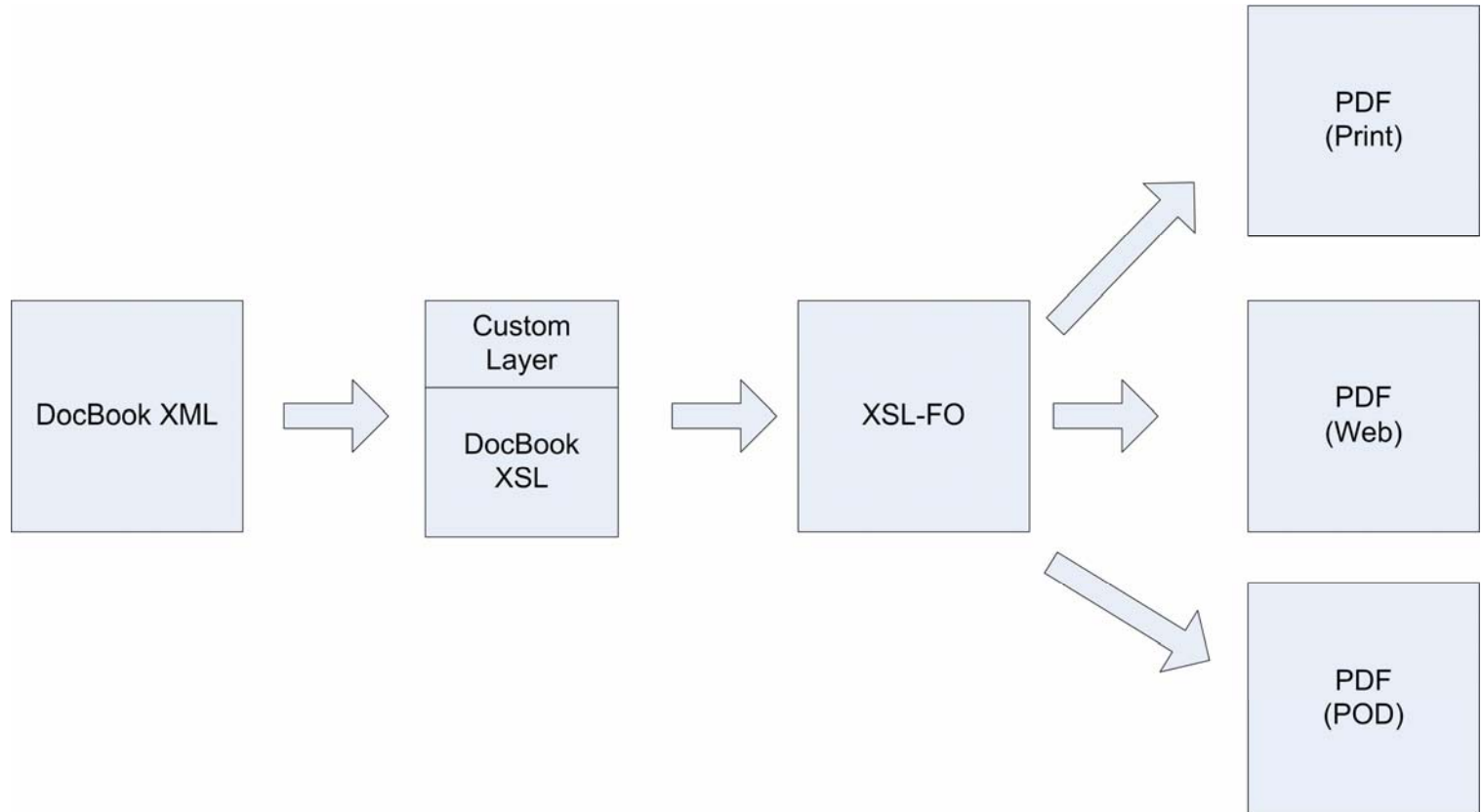
Print Production



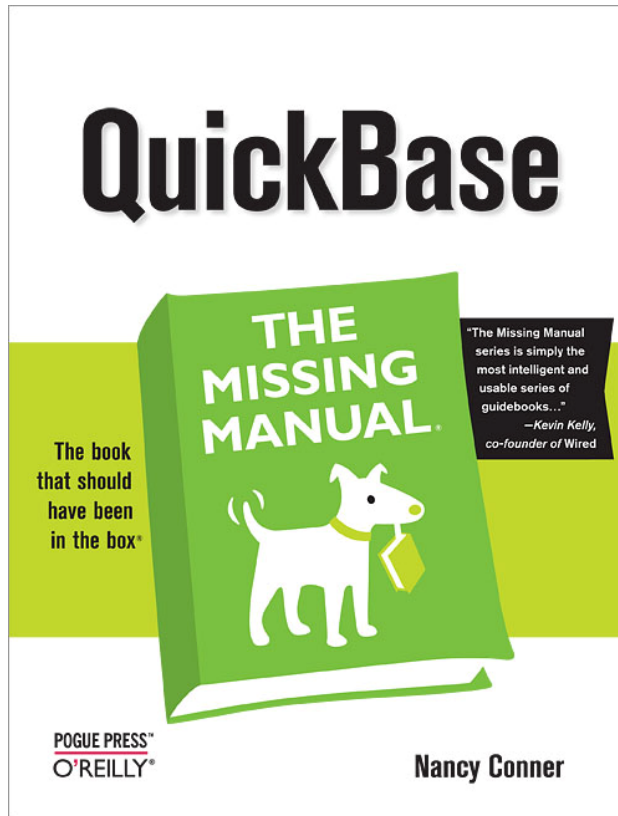
Performance Improvements



Print Production Toolchain

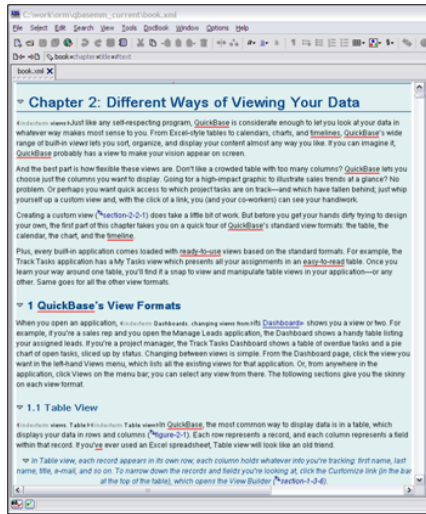


New Publishing Models



- Safari first
- PDF online
- Print-on-Demand Option
- True single sourcing

QuickBase: The Missing Manual

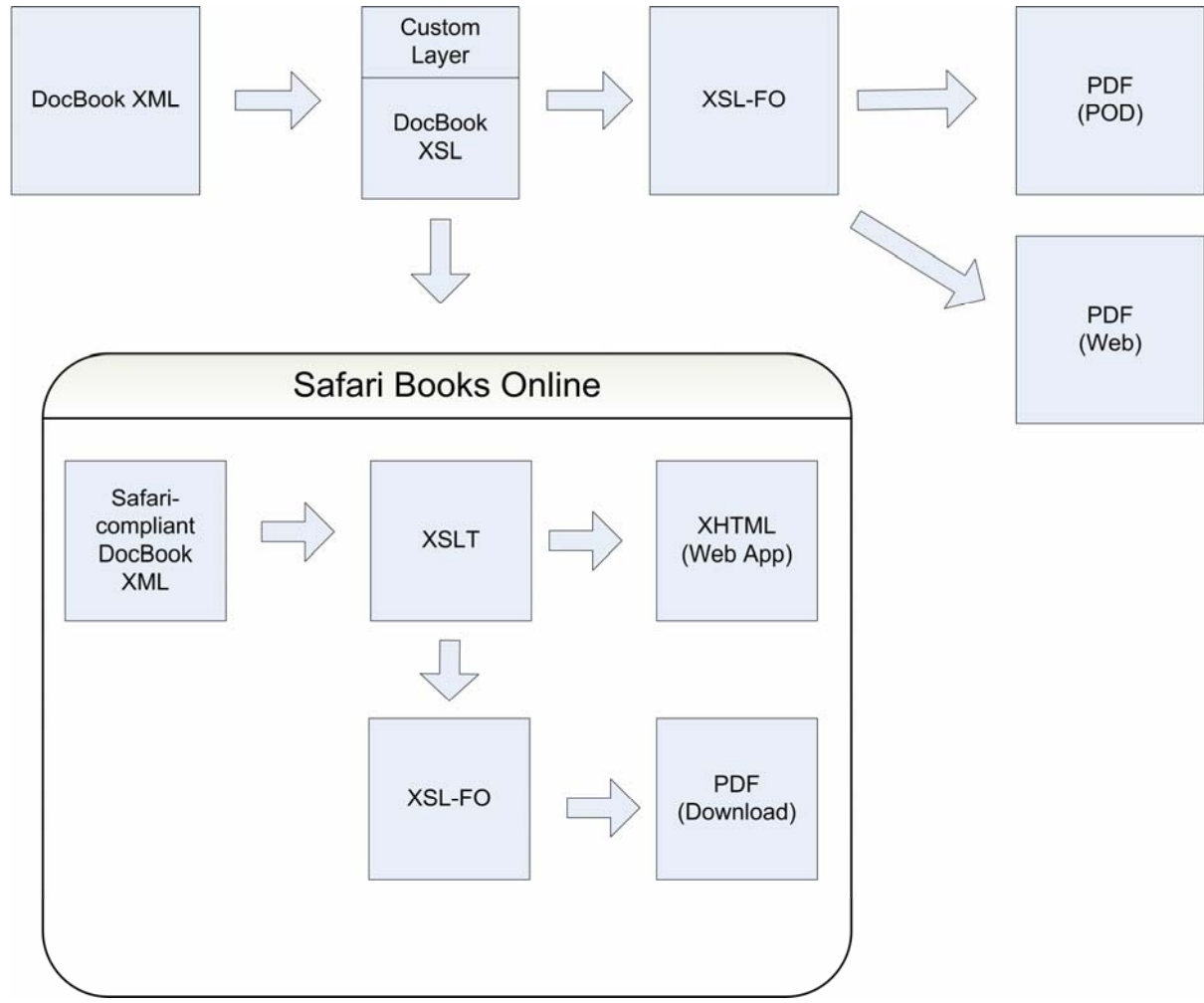


HTML Online

PDF Online



Single Source Toolchain



Lessons Learned

- Standardization around DTD/Schema
- “PDF” means a lot of different things
 - Print
 - Web-friendly
 - Amazon/Google
 - Print on Demand
- Need management buy-in for print trade-offs
- Not appropriate for all (or even most) titles