

# XPath - A practical guide

Arne Blankerts <arne@thephp.cc>, Tobias Schlitt <toby@php.net>

IPC 2009

2009-11-17

1 Welcome

2 Introduction

3 In details

4 Quiz

5 The end



- Arne Blankerts <arne@thephp.cc>
- PHP since 1999 (10 years of PHP!)
- Co-Founder of thePHP.cc
- ballyhoo. werbeagentur.
- Open source addicted
  - Inventor and lead developer of fCMS site system
  - Contributor and translator for the PHP manual



- Tobias Schlitt <toby@php.net>
- PHP since 2001
- Freelancing consultant
- Qualified IT Specialist
- Studying CS at TU Dortmund  
(finishing 2010)
- OSS addicted
  - PHP
  - eZ Components
  - PHPUnit



- 1 Welcome
- 2 Introduction
  - Overview
  - Basics
- 3 In details
- 4 Quiz
- 5 The end



## XPath

Enables you to select information parts from XML documents.

- Traverse the XML tree
- Select XML nodes
- W3C recommendation
- Version 1: November 1999
- Version 2: January 2007
- Fields of application
  - XSLT (XML Stylesheet Language Transformations)
  - Fetching XML nodes within programming languages

- Every XPath expression matches a **set of nodes** (0..n)
- It encodes an “address” for the selected nodes
- Simple XPath expressions look similar to Unix file system addresses
- Absolute vs. relative addressing

Let's dig into the code





1 Welcome

2 Introduction

3 In details

- Syntax
- Specialities
- Functions
- Axis
- PHP function integration

4 Quiz

5 The end



- / Level separator
- // Multi-level separator
- \* Wildcard
- @ Attribute prefix
- @\* Attribute wildcard
- [...] Filter / condition

- Every expression step creates new context
- Next evaluated in context created by previous step

- / One level down
- // Any number of levels down
- ../ One level up

Let's dig into the code



Access nodes by position: [2]

## Start index

- Indexing generally **1** based
- Some Internet Explorer versions start with **0**

Union the node sets selected by multiple XPath expressions.

Let's dig into the code





## String

- `string-join()`
- `substring()`
- `starts-with()`
- `contains()`
- `string-length()`

## Math

- `round()`
- `floor()`

## Context

- `position()`
- `last()`

## Aggregate

- `count()`
- `min()` / `max()`

## Node

- `local-name()`

## Logic

- `not()`
- `true()`

## Function overview

An overview on all functions can be found on  
<http://www.w3.org/TR/xpath-functions/>

## 13 dimensions

Imagine an XML document to be a 13 dimensional space...

Axis	Shortcut
ancestor	
ancestor-or-self	
attribute	@
child	-
descendant	
descendant-or-self	//
following	
following-sibling	
namespace	
parent	..
preceding	
preceding-sibling	
self	.

Let's dig into the code



- Register PHP callbacks to enhance functionality
- Possible since PHP 5.3.0
- Works with functions and static methods

Let's dig into the code



- 1 Welcome
- 2 Introduction
- 3 In details
- 4 Quiz**
- 5 The end



- `//node`

- `//node`
- `/root/element`



- `//node`
- `/root/element`
- `/root/element[1]`

- `//node`
- `/root/element`
- `/root/element[1]`
- **`@id`**

- `//node`
- `/root/element`
- `/root/element[1]`
- `@id`
- `/root//node/@name`

- `//node`
- `/root/element`
- `/root/element[1]`
- `@id`
- `/root//node/@name`
- `/root/element[@name = 'php']/@*.`

- `//node`
- `/root/element`
- `/root/element[1]`
- `@id`
- `/root//node/@name`
- `/root/element[@name = 'php']/@*.`
- `./some/*[@class='tek']/@id`

- 1 Welcome
- 2 Introduction
- 3 In details
- 4 Quiz
- 5 The end**



- Are there any questions left?
- Please give us some feedback!

*Thank you!*

- Thanks for being here!
- We hope you enjoyed the session!
- Slides and material
  - Delivered by Software & Support
  - <http://schlitt.info/opensource>
  - On slideshare: <http://www.slideshare.net/tobyS>
- Contact us:
  - Arne Blankerts <[arne@thephp.cc](mailto:arne@thephp.cc)>
  - Tobias Schlitt <[toby@php.net](mailto:toby@php.net)>
- Please rate our talk at: <http://joind.in/1042>