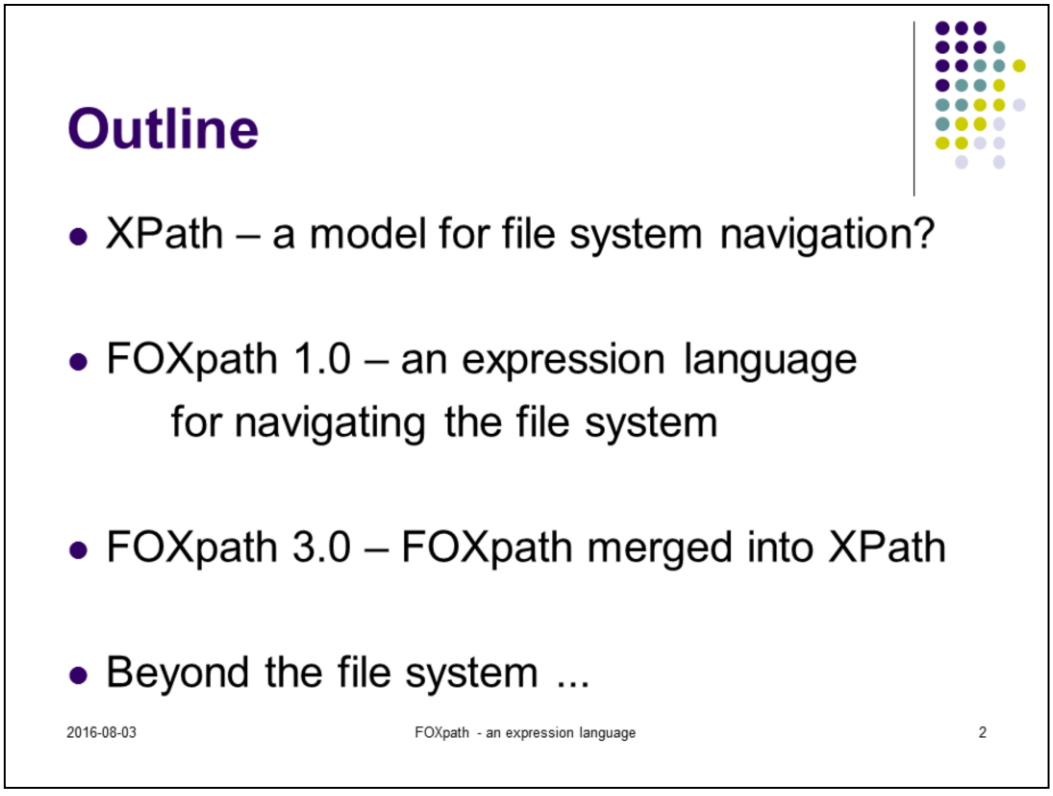
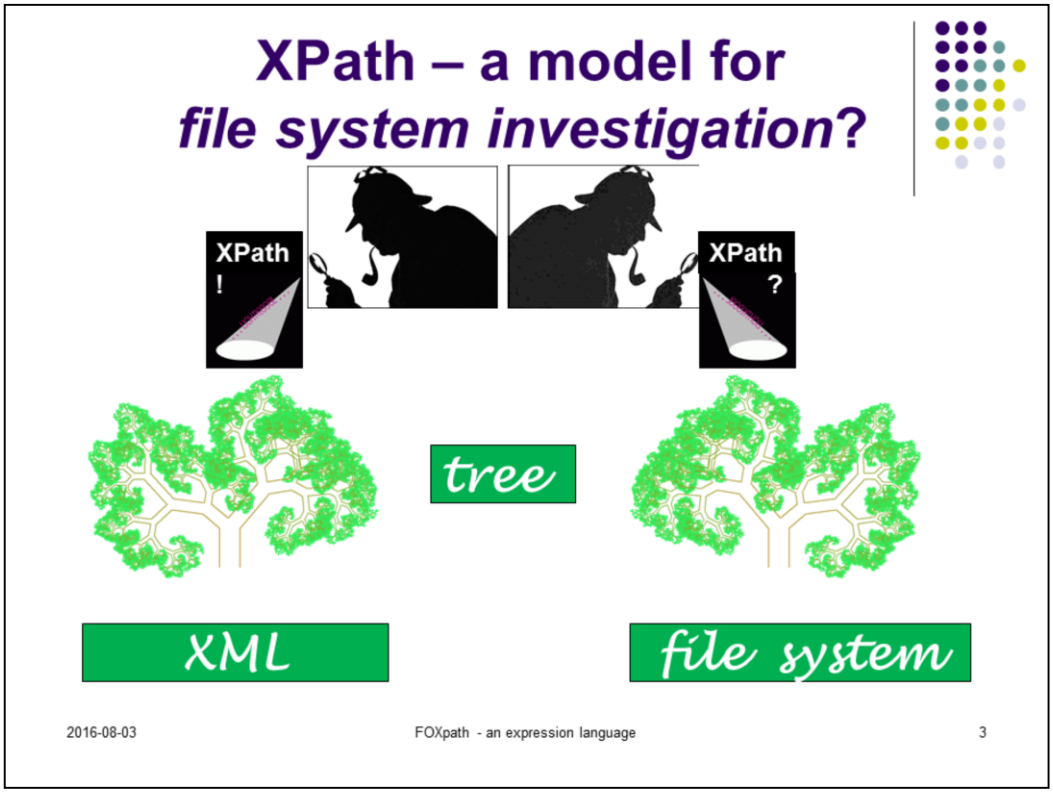
XPath is an expression language for navigating the contents of XML documents. Using XPath, we can select contents in an elegant, concise and very intuitive way. What is more: XPath is not a loose collection of expressions and rules, but a complete and fully composable expression language. Therefore there is virtually no limit to the selectiveness we may achieve, constructing XPath expressions. XML documents and a file system have much in common – they both expose a tree-structured collection of names. So – should we not have something like „XPath for files and folders“ – FOXpath for short? This presentation reports my efforts to design and implement such an expression language.



First, we check the basic assumption that XPath is an adequate model for file system navigation. Then we explore FOXpath, a new expression language for navigating the file system. We shall see that this first version of FOXpath is a modified copy of XPath. But then we merge FOXpath back into XPath, obtaining an extended version of XPath supporting file system navigation. And finally we explore how FOXpath is not restricted to physical file systems, but can navigate logical file systems as well, as for example used to organize the contents of a NOSQL database, version control repositories or the URI collection exposed by REST-ful web services.

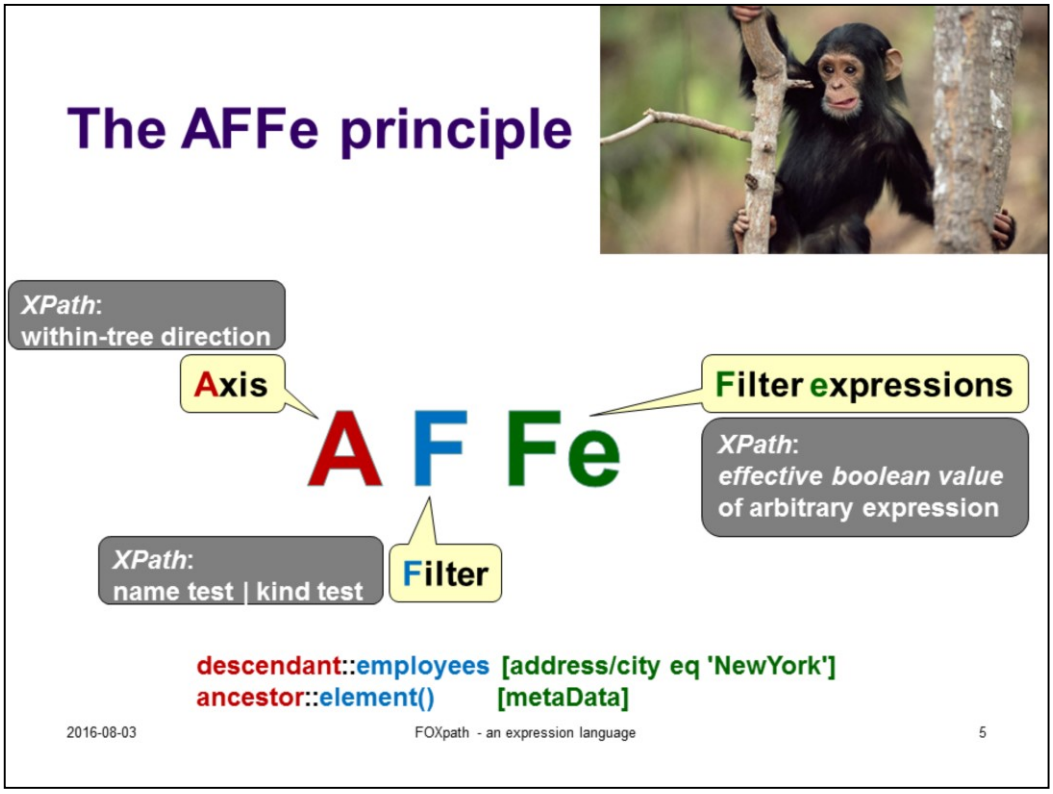


XPath îs a device for exploring and navigating XML documents. A closer look at XPath reveals core concepts not based on XML, but on the general notion of a tree-structured collection of names. XML is a particular instance of that notion, and the file system is another instance. We conclude that it should be possible to design an expression language which is similar to XPath, but deals with files and folders, rather than nodes.

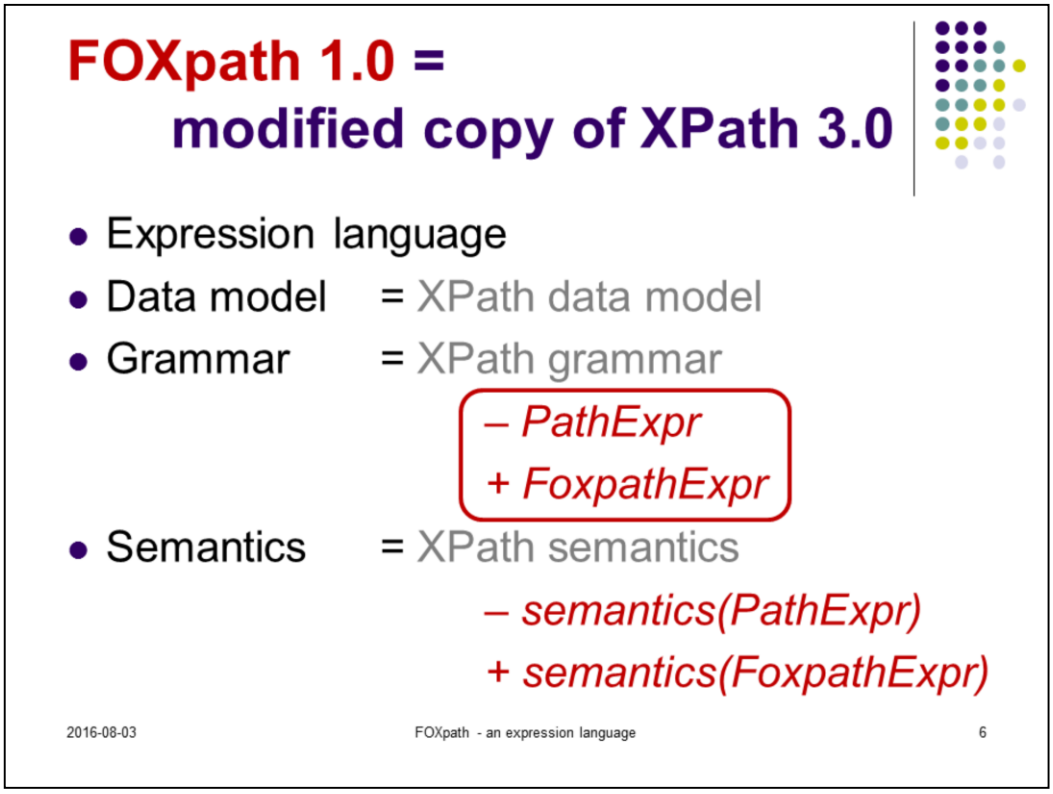
Goal: examples

```
/xsdbase/niem-3.0/descendant~::*iso*.xsd/parent~::*

/xsdbase/niem-3.0//(*.xlsx, *.txt, *.xml)

/xsdbase/niem-3.0//*[is-dir()][not(*)]

sort(distinct-values(
  /xsdbase/niem-3.0//*.xml/xroot()), lower-case#1))

/xsdbase/niem-3.0//*.xml[has-xroot('catalog')]

/xsdbase/niem-3.0//*.xsd/file-name() except
  /xsdbase/niem-2.1//*.xsd/file-name()
```

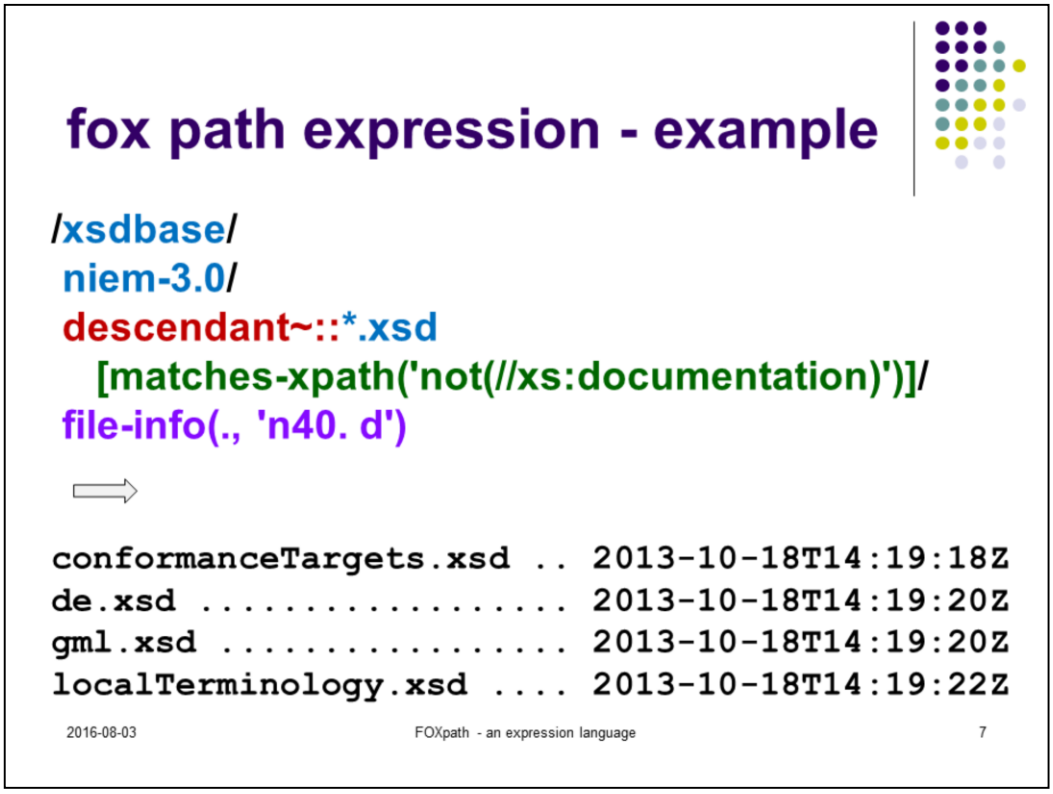
Let us start with a few examples illustrating the desired expression language. They look like XPath, and they are as powerful as XPath expressions.



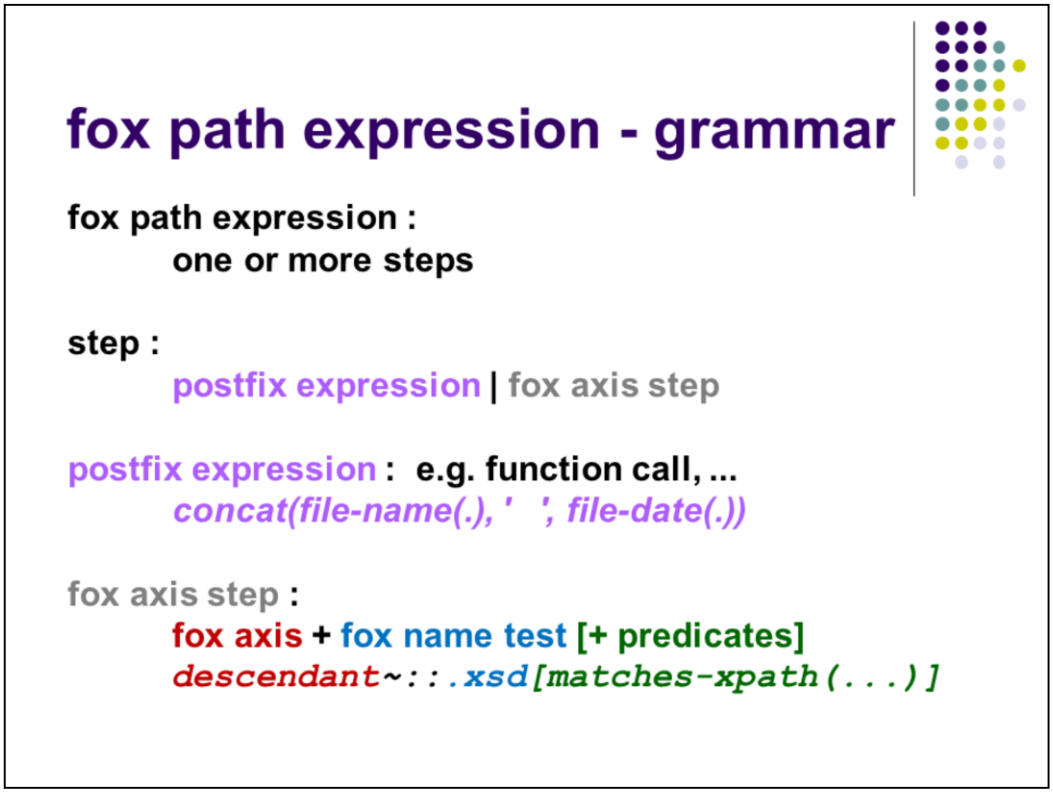
If we want to use XPath beyond XML, we must understand the core principles of XPath in an abstract way, which is independent of XML. „Affe“ is the German word for monkey, and most monkeys love to traverse trees. It‘s natural to summarize the core principles of XPath under the term AFFe. It‘s an acronym referring to the building blocks of selection: axis + filter + filter expression. The axis collects everything visited when traversing the tree in a particular direction (children, descendants, parent, ancestors, and so forth). Collected items are submitted to a simple filter, which in the case of XPath is either a name test or a kind test. If this filtering is still too crude, further filtering of unlimited distinctiveness can be achieved by adding filter expressions.



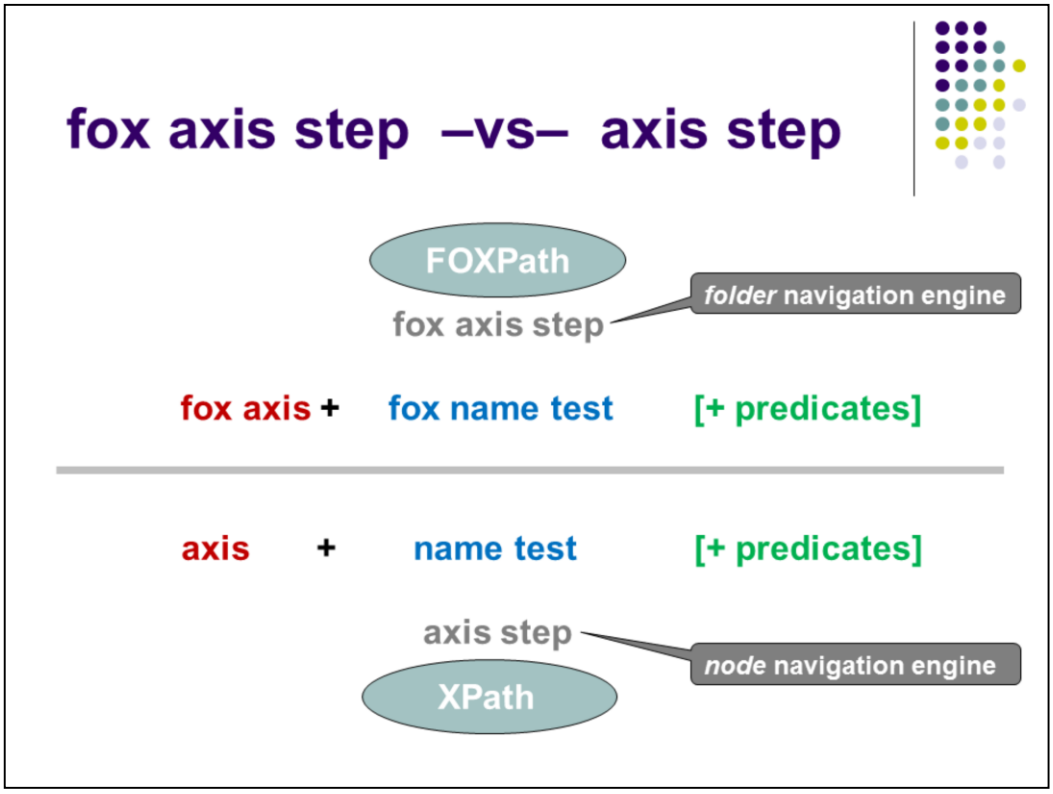
XPath is a masterly elaboration of the AFFe principle. I never cease to marvel at the expressiveness achieved by uncomprising rigour and consistency. Aspiring to a new „XPath for files and folders“ language, what would be more promising than to stick as closely to XPath as possible?



So the fox path expression is the only difference between XPath and FOXpath! Before becoming formal, let us look at an illustrative example. The expression reports all Niem XSDs without documentation. Ignoring the tilde and the use of wildcards within names, this could be an XPath expression.

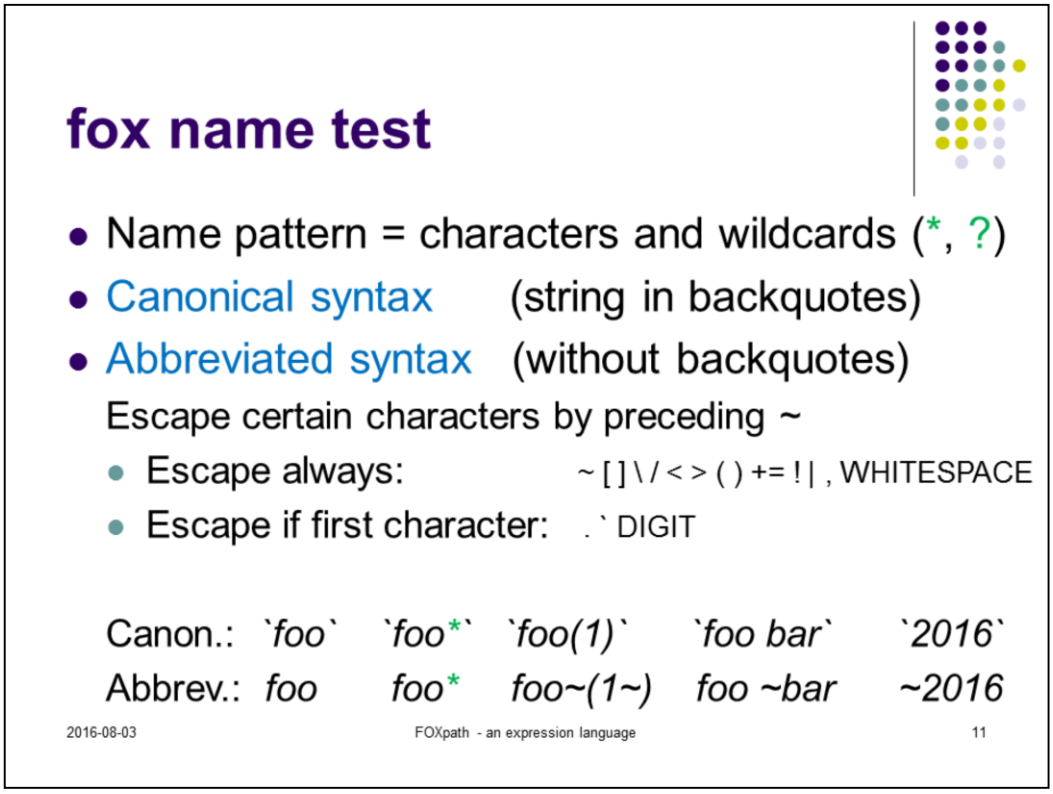


That‘s because the grammar of the fox path expression is just a modified copy of the grammar of the path expression.

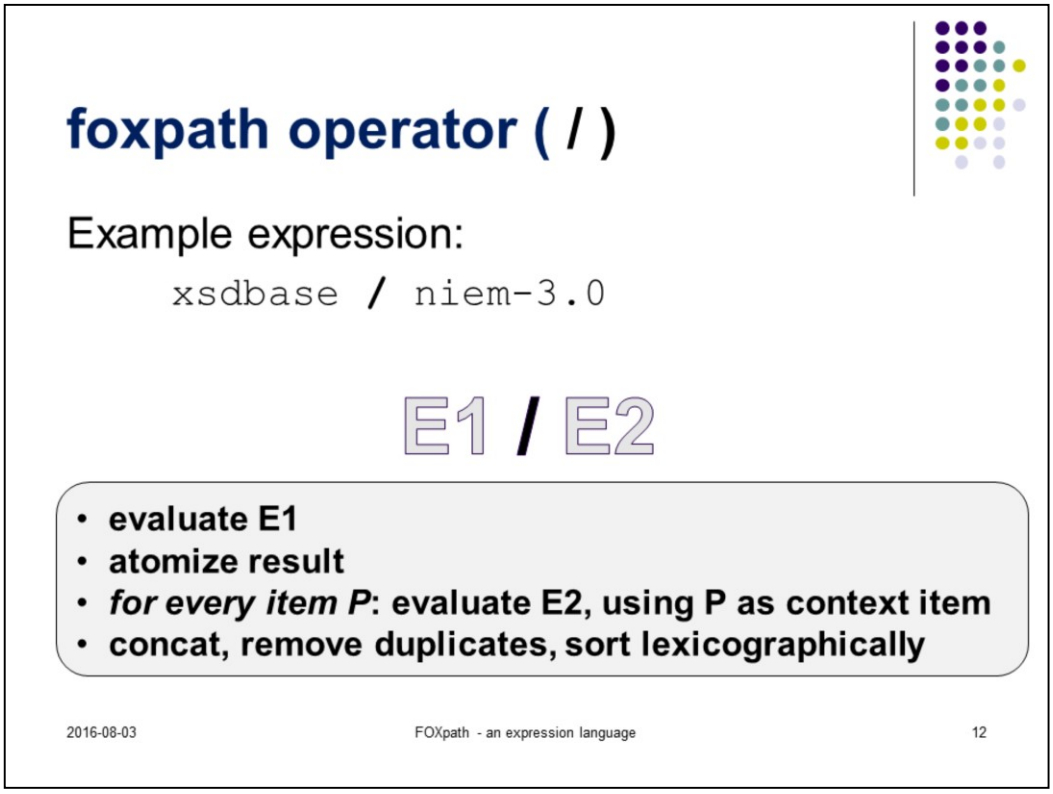


So the core functionality of folder nagivation, which is the fox axis step, is a modified copy of the axis step of XPath, which is the core functionality of node navigation..

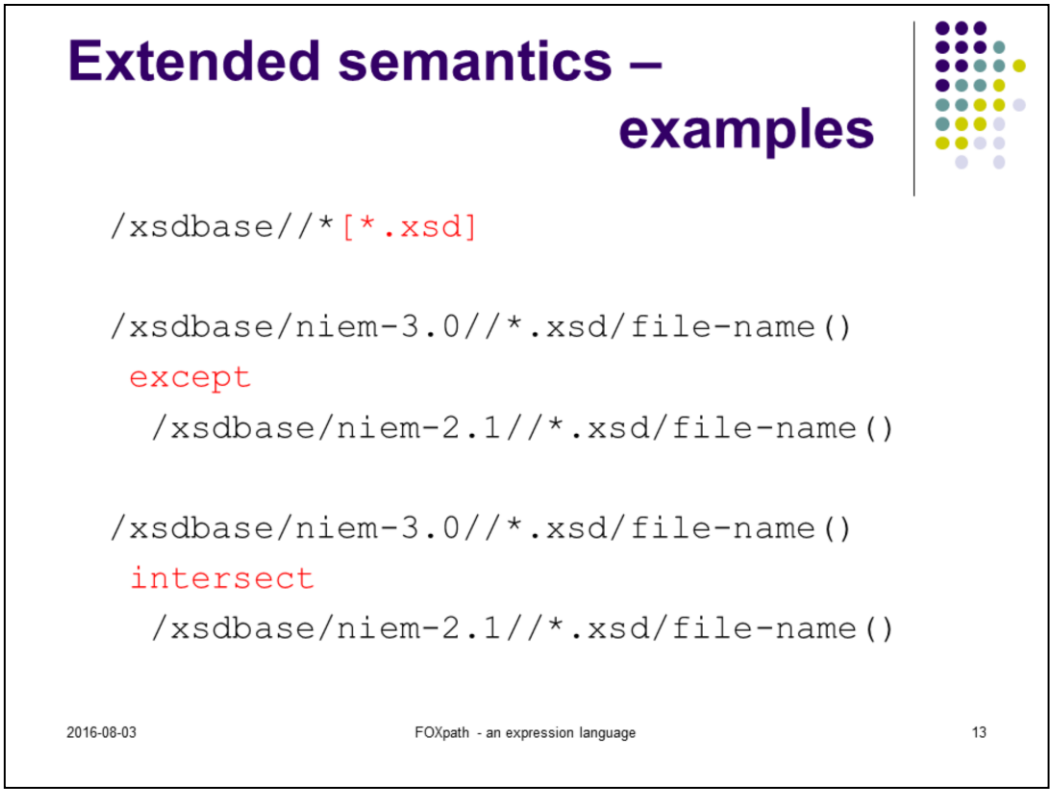
FOXpath supports most of the axes supported by XPath, excepting of course the attribute axis, as well as the preceding and following axis, which are not believed to be useful when navigating a file system. As in XPath, the child axis is the default axis.



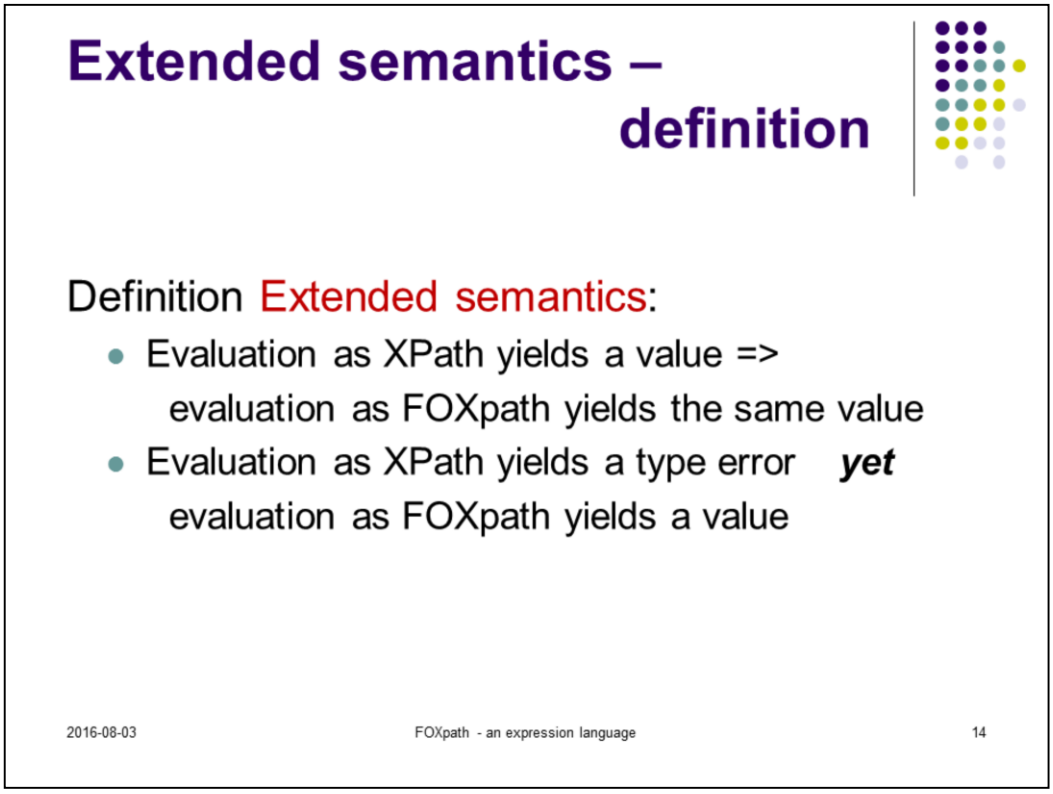
The fox name test filters by file or folder name. In contrast to the node name test of XPath, name patterns with inserted wildcards are supported. As file names can be almost arbitrary strings – not only NCNames – the syntax of a name test must make sure to avoid parsing ambiguity.



In XPath, the steps of a path expression are combined by the path operator, represented by a slash. Similarly, in FOXpath, the steps are connected by the foxpath operator, also represented by a slash. The operator semantics are an adapted copy of the semantics of the path operator. Adaptation is required, as the lefthand operand supplies URIs, not nodes.



As we want FOXpath expressions to be just as elegant and expressive as XPath expressions, we are motivated to introduce a slight extension to the expression semantics of XPath.



Extended semantics are defined in terms of a comparison between the outcomes of evaluating the expression as an XPath or as an FOXpath expression.

Extended semantics – rules

- Effective boolean value

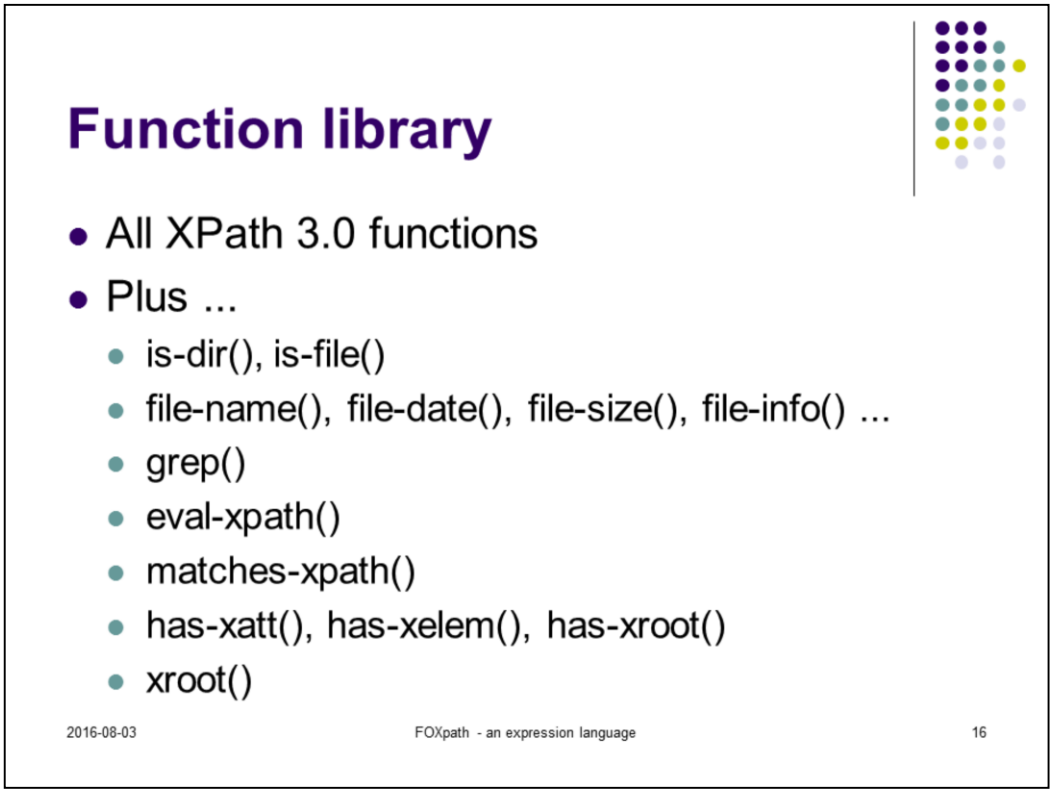
  Case: value(E) = sequence of >1 atomic items

  XPath: EBV(E) raises type error

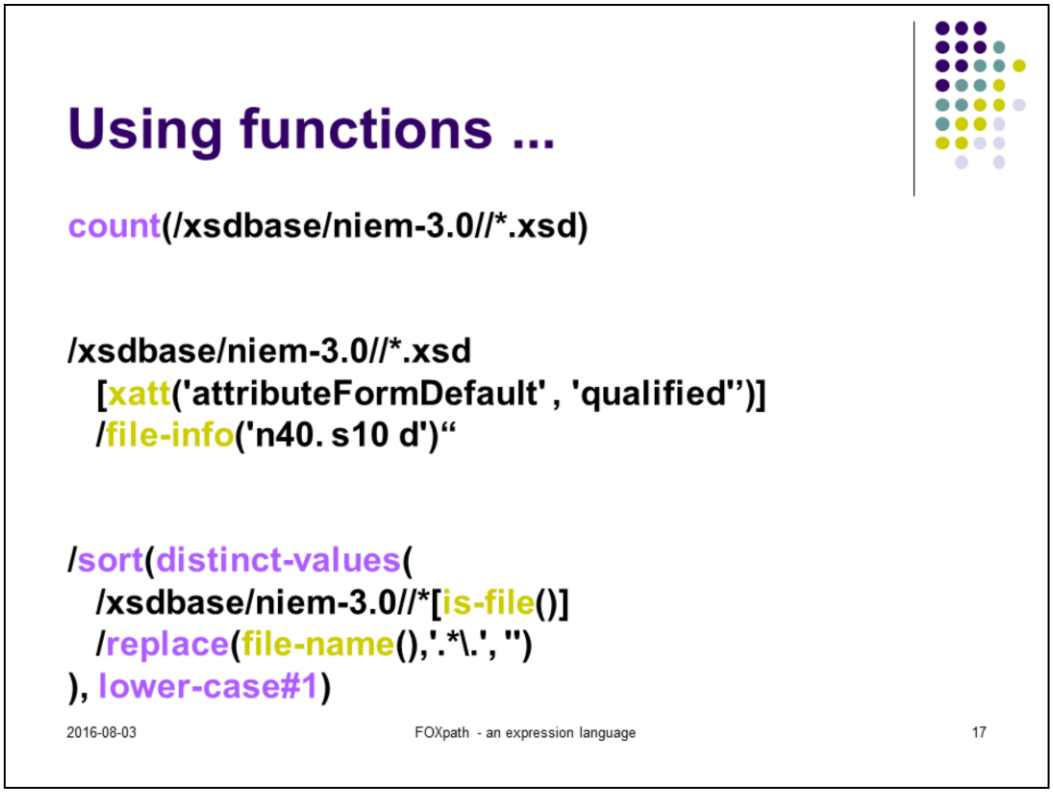
  FOXpath: EBV(E) = EBV(E[1])

- E1 or E2 contains atomic item =>
  - E1 union E2 = distinct-values ((E1, E2))
  - E1 except E2 = distinct-values (E1[not(. = E2)])
  - E1 intersect E2 = distinct-values (E1[. = E2])

FOXpath defines four semantic extensions.



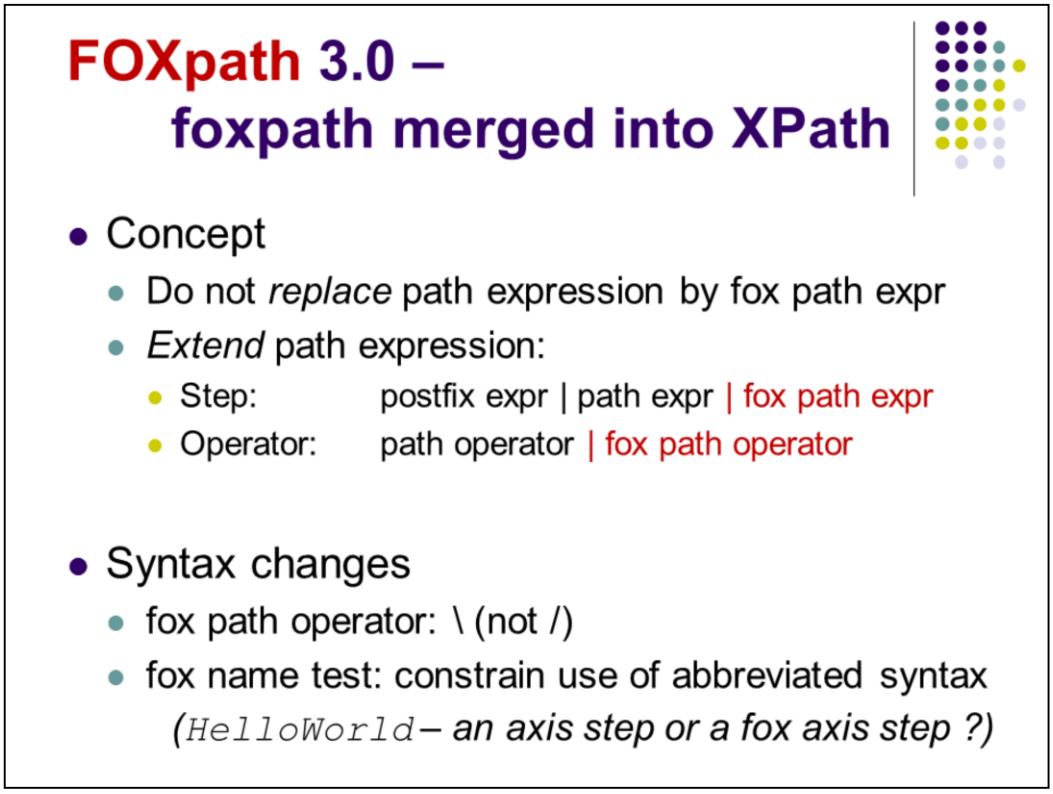
The power of XPath relies on a library of built-in functions. FOXpath supports all XPath functions, plus a few additional functions believed to be especially useful when navigating and reporting file system contents.



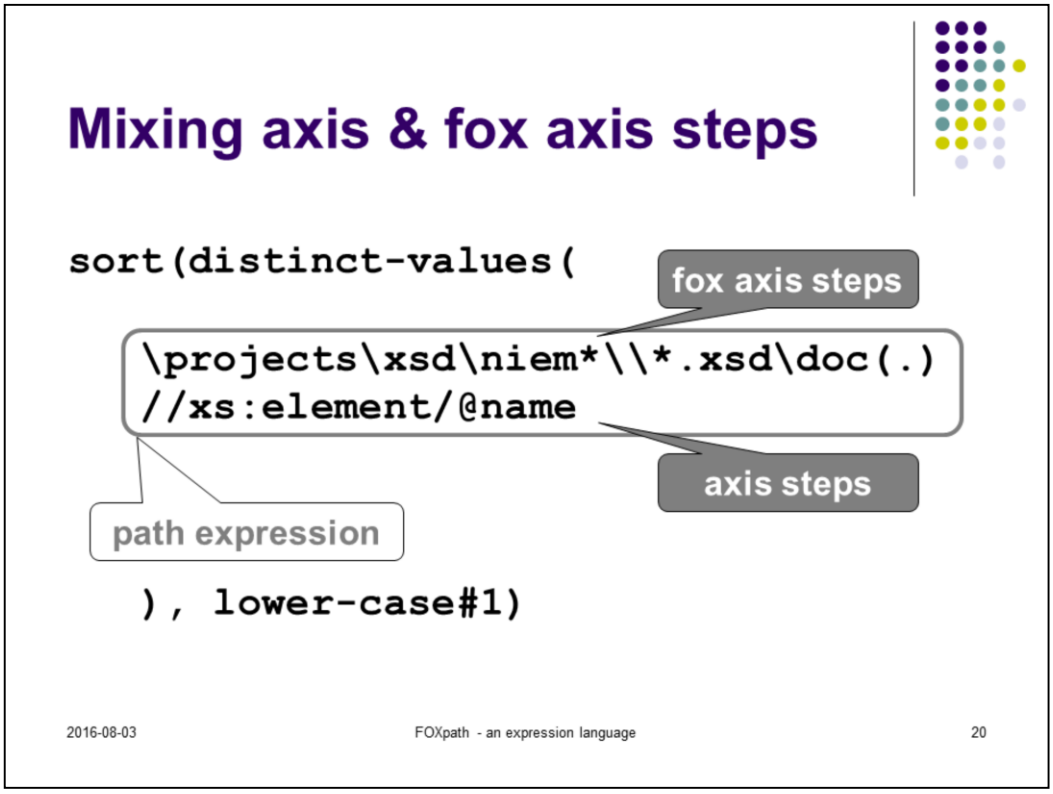
A few examples illustrate the use of XPath as well as FOXpath functions.

Bringing It All Back Home

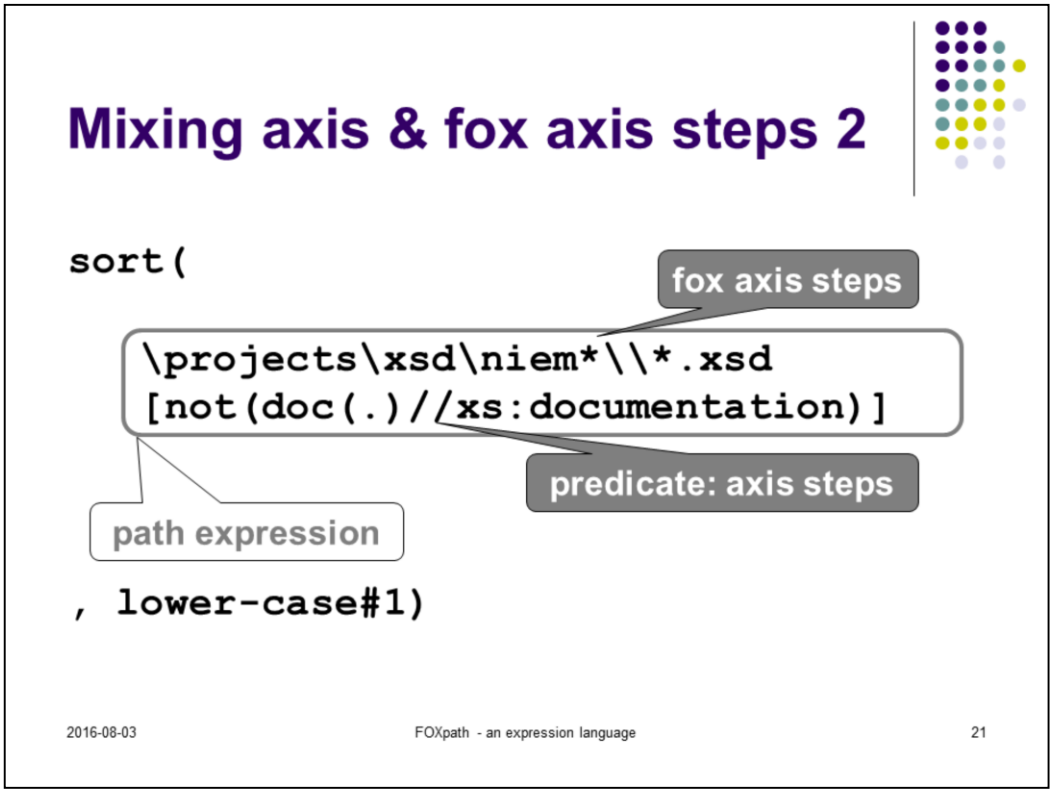
If FOXpath is almost identical to XPath, it is natural to ask if we cannot merge FOXpath back into XPath, obtaining an extended version of XPath, rather than a modified copy..



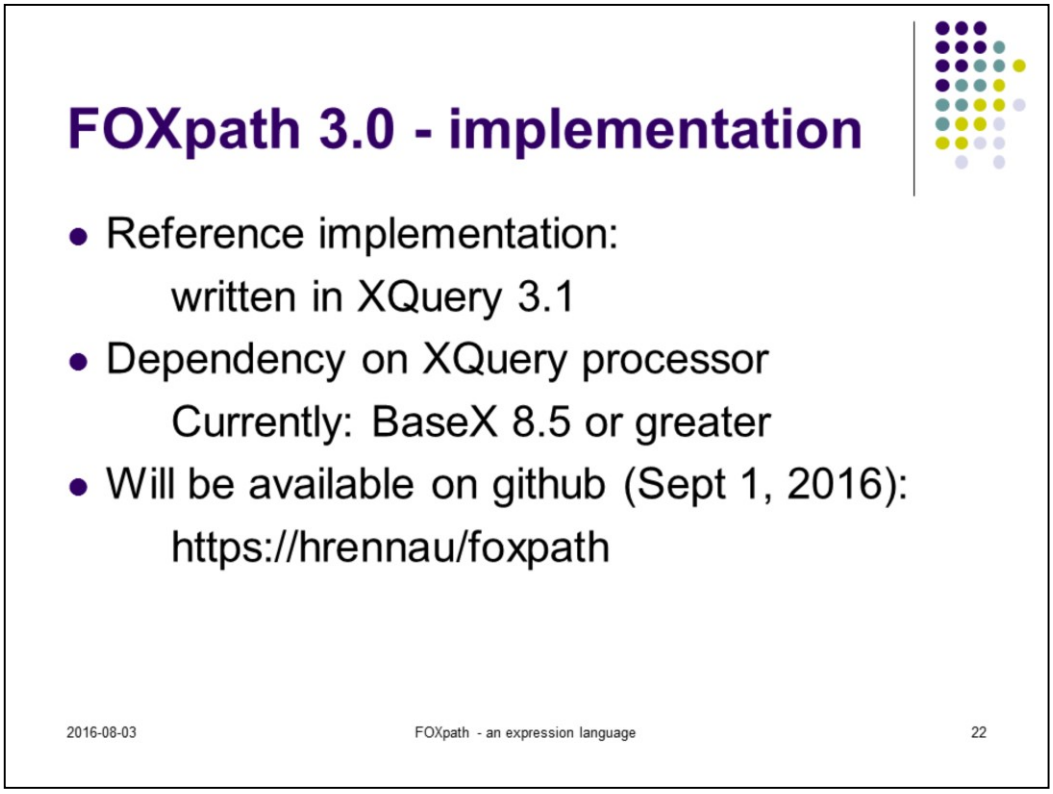
Indeed – it‘s possible! This new version of FOXpath is called FOXpath 3.0, as it is an extended version of XPath 3.0.



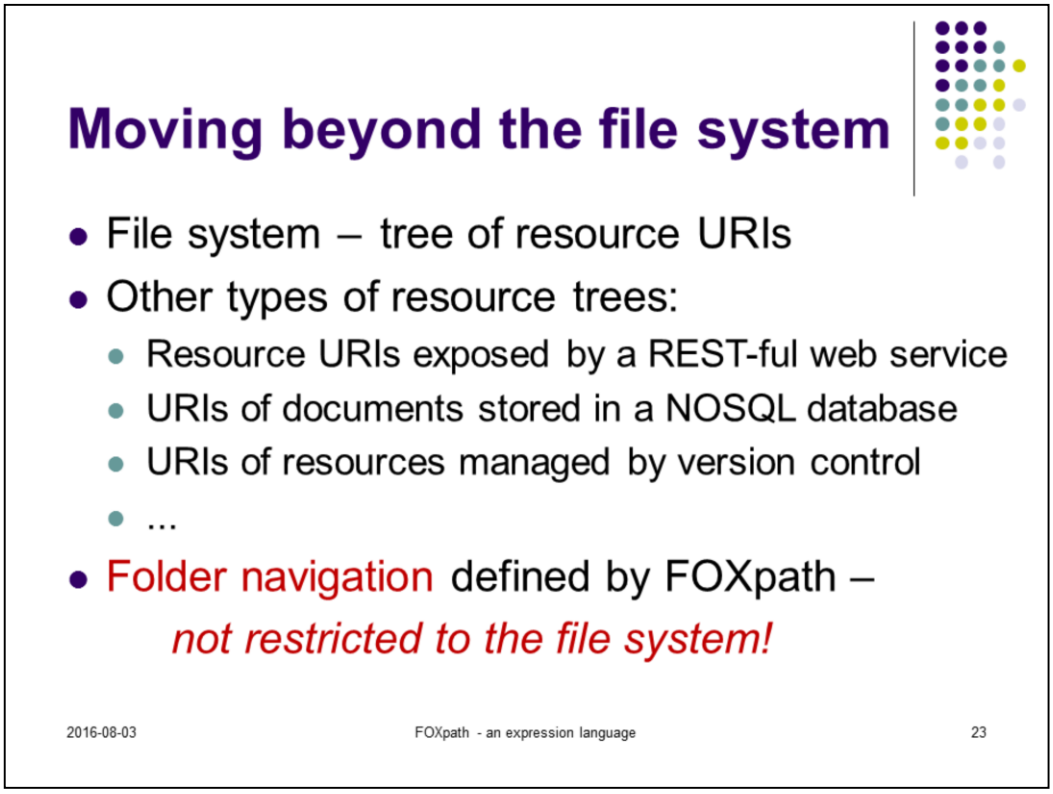
With FOXpath 3.0, you can not only use fox axis steps, you can also mix axis and fox axis steps within a single path expression. This enables expressions accomplishing a two-phase navigation: initial steps select resources, following steps navigate into their contents.



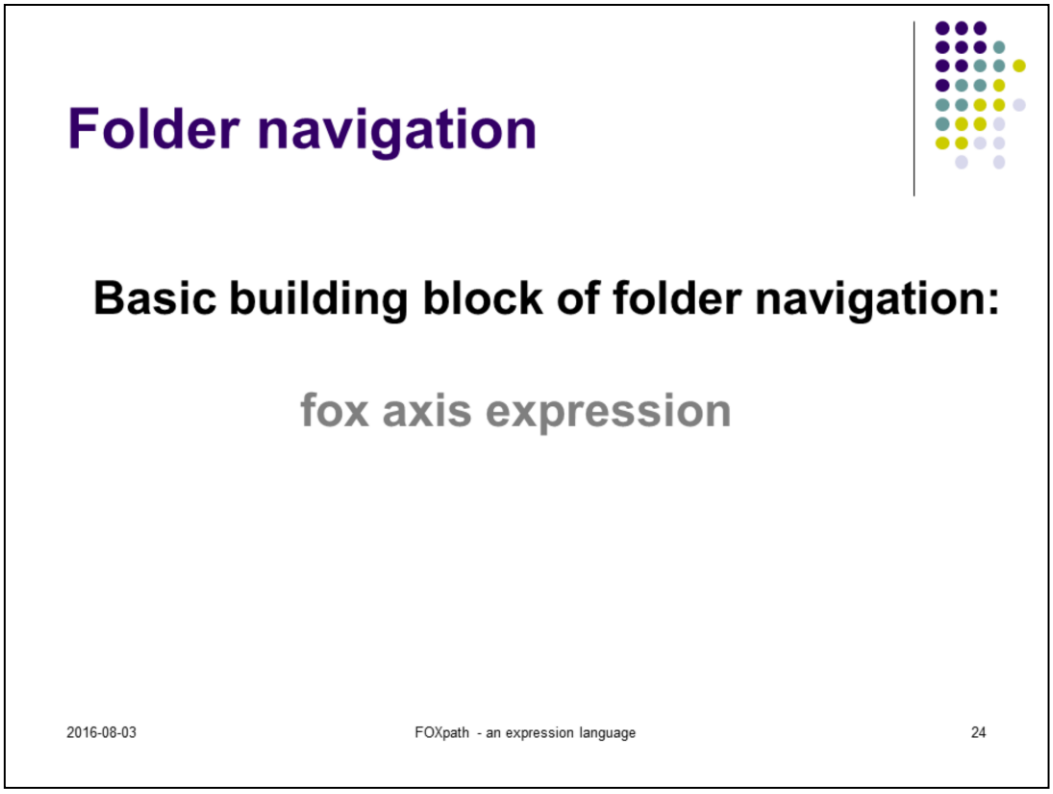
Another example of mixing is the use of fox axis steps with predicates containing axis steps, which means resource selection based on resource contents.



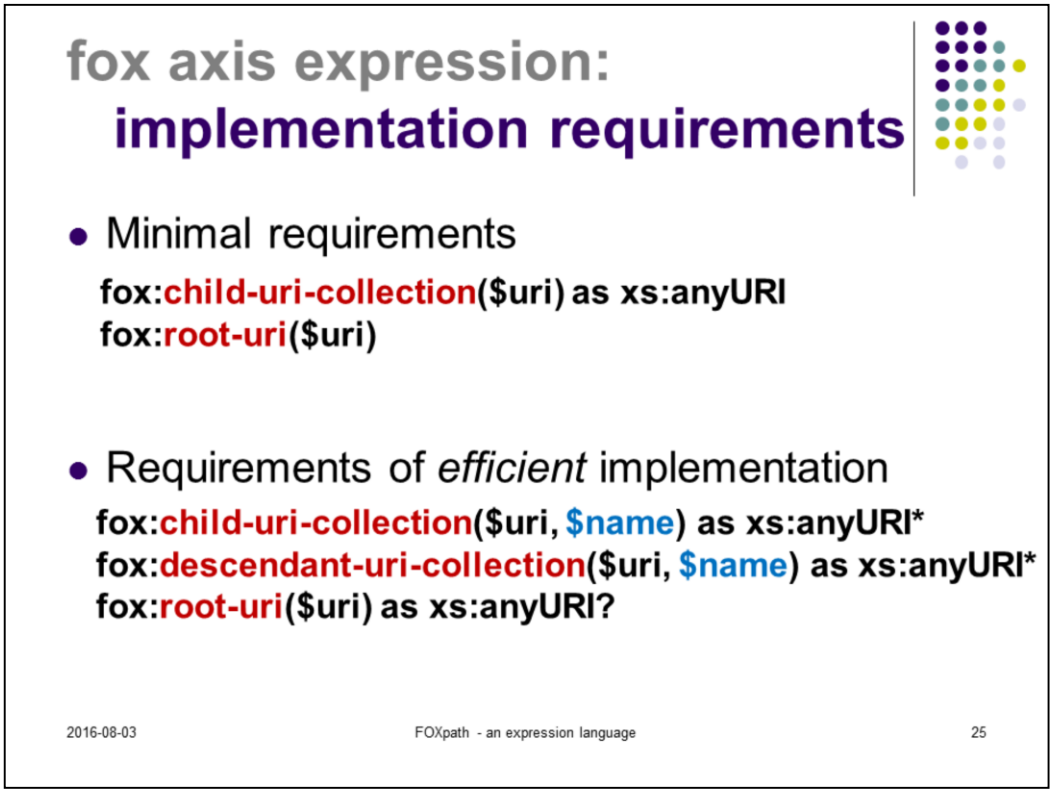
A reference implementation of FOXpath 3.0 is available.



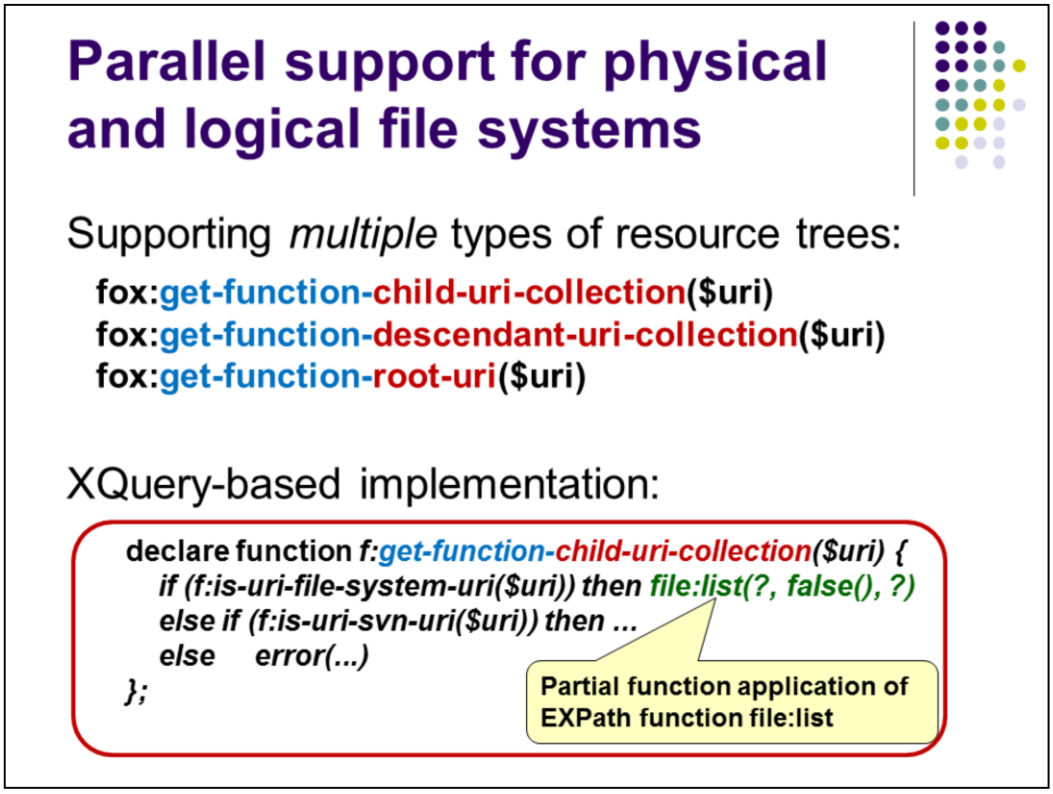
The folder navigation supported by FOXpath is not restricted to physical file systems – it can also deal with other kinds of resource trees, which may be regarded as logical file systems.



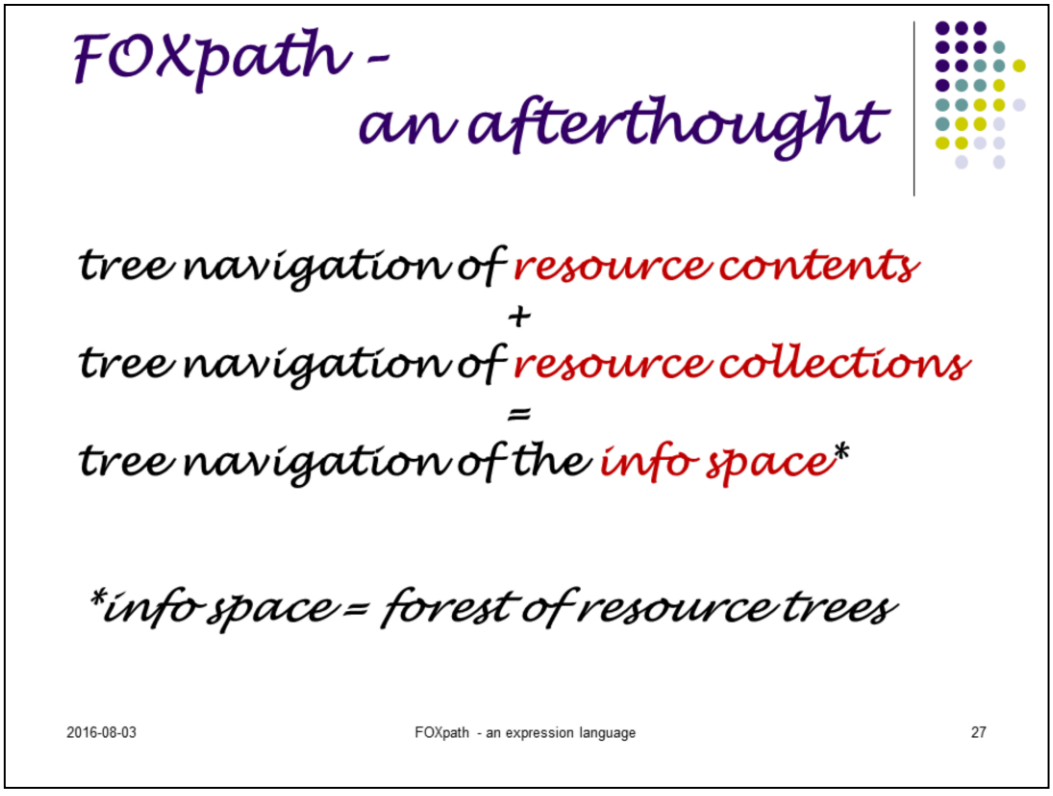
It is important to recognize the basic building block of folder navigation – which is the fox axis expression. The question arises whether the fox axis expression can deal with logical file systems as well.



As you remember, the fox axis expression combines a fox axis with a fox name test. The implementation relies on navigation primitives, two in principle, and three when we want to ensure efficiency. These primitives are functions mapping an input URI to output URIs – child URIs, descendant URIs and the root URI.

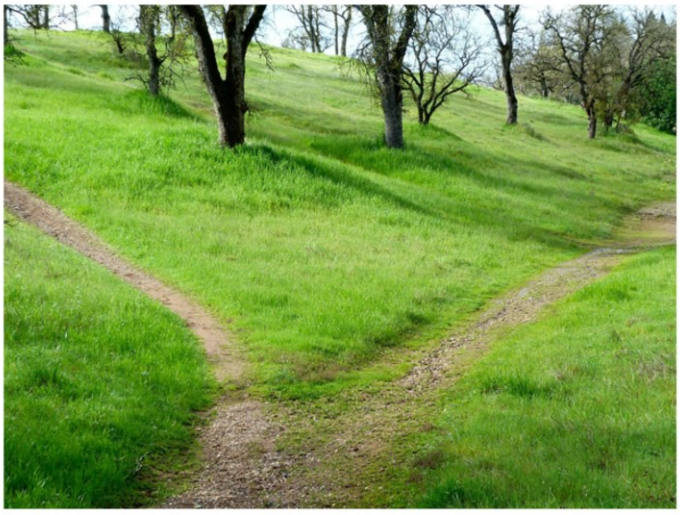


Implementations of those primitives will typically be specific for a particular type of resource tree – for example a physical file system, or an SVN repository. An implementation of the FOXpath language can, however, support in parallel any number of resource tree types. Any resource tree type can be included which meets the following conditions: (a) there is an implementation of the three navigation primitives available, (b) any URI belonging to an instance of the resource tree type is recognized to belong to that resource tree type. Supporting multiple resource tree types can be elegantly achieved using higher order functions delivering for a given URI the appropriate function implementations.



An afterthought: FOXpath complements tree navigation of resource contents with tree navigation of resource collections. FOXpath encourages us to perceive a forest of information, consisting of multiple resource trees, whose leaves are resources exposing node trees. As the complete structure can be navigated in a seamless way, it may be experienced as a single, coherent space – the info space.

Bye, bye!



There is a choice to be made: either XPath remains an ingeneous tool for navigating XML documents, or it will be extended to become the engine of the info space.