Markup Languages und Anwendungen

# **Ruby XML Mapping**



Marc Seeger (20488) mail@marc-seeger.de Computer Science and Media, HdM Stuttgart, WS 08/09

## Agenda

- Ruby
- XML Mapping Libraries
- YAML

## Why Ruby?

A language that doesn't affect the way you think about programming, is not worth knowing

#### History

- Origin:
  - Yukihiro "Matz" Matsumoto
  - Japan 1993



- 1st english book: 2000
- Ruby on Rails: 2004

#### Language Basics

#### Variable Declaration:

```
text = "Hallo Welt" <== String
zahl = 3.5 <== Float
bla = 3 <== Fixnum
blubb = 23252345863465364564564563 <== Bignum</pre>
```

#### Function Declaration:

```
def do_something(text, number)
  puts text * 3
  puts number * 3
end
```

```
do_something("Marc", 4)

→

MarcMarcMarc

12
```

#### Language Basics

- ClassNames
- method\_names and variable\_names
- methods\_asking\_a\_question?
- slightly\_dangerous\_methods!
- @instance\_variables
- \$global\_variables
- SOME\_CONSTANTS or OtherConstants

#### Language Basics: Hashes

## Ruby is a syntactic sugar factory



Principle of Least Surprise

#### Language Basics: Arrays

```
1. a = [3.14159, "pie", 99, "Blubb"]
2. a.type
                 »Array
3. a.length
                   »3
 4. a[1]
                   »"pie"
5. a[4]
                   >nil
6. a[-1]
                   »"Blubb"
7. a[-2]
                   ≫99
8. a[1, 3]
                   »["pie", 99, "Blubb"]
9. a[0..2]
                   »[3.14159, "pie", 99]
10.
11. b = Array.new
12. b.type
                  »Array
13. b.length
                   »0
14. b[0] = "second"
15. b[1] = "array"
                   »["second", "array"]
16. b
```

## Give me some sugar: Array

```
people = Array.new
people << "Marc" << "Christian" << "Jakob" << "Michael"
people = ["Marc", "Christian", "Jakob", "Michael"]
people.push("Marc", "Christian", "Jakob", "Michael")
people = %w("Marc", "Christian", "Jakob", "Michael")</pre>
```

#### Give me some sugar: more...

```
5.times { print "Hallo HdM!" }
```

#### **Control Structures**

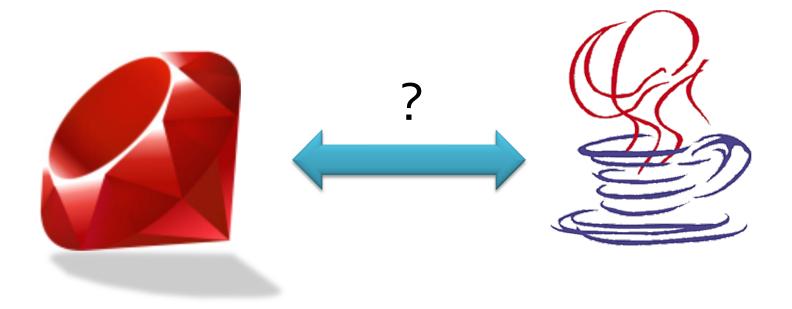
```
if expr [then]
  expr...
[elsif expr [then]
  expr...]...
[else
  expr...]
end
```

```
until expr [do]
...
end
```

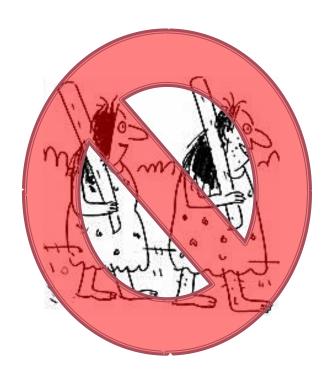
```
for i in [1, 2, 3]
  puts i*2
end
```

```
puts "Error!" if $debug
```

```
puts "Error!" unless $production_mode
```



- No Primitives, Integers and floats are objects!
  - -1.abs => 1





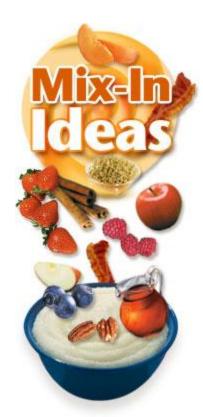
No Semi-colons

- nil, not null
- nil is an object!
  - nil.nil? => true
  - nil.class => NilClass
- nil and false are false
  - everything else, including o, is true

- Expression oriented syntax.
  - Almost everything returns a value
  - Methods automatically return their last expression.



- Single Inheritance
  - But mixins are available (= Interface with implemented methods)



#### Ruby for Java Coders: Mixin Example

```
module BarModule
  def hello world
    puts "Hello World"
  end
end
class BaseClass
  def class method
    puts "In class method"
  end
end
class Foo < BaseClass</pre>
  include BarModule
end
f = Foo.new
f.class method
f.hello world
```

← This module implements the mixin

←A class that doesn't do that much

←inheriting ←and mixing!

←We inherited that one ←And mixed in that one

Classes are always open (even built in classes)

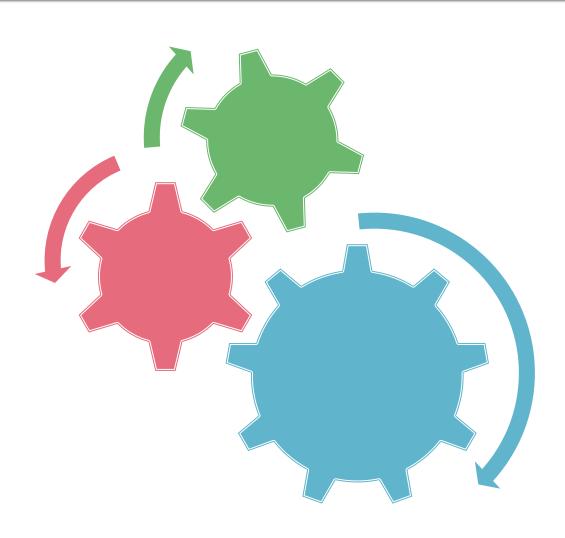
```
class String
  def foo
  "foo"
  end
end
end
puts "hdm test".foo ==> "foo"
```



Another Example from Rails:

1.hour.from\_now

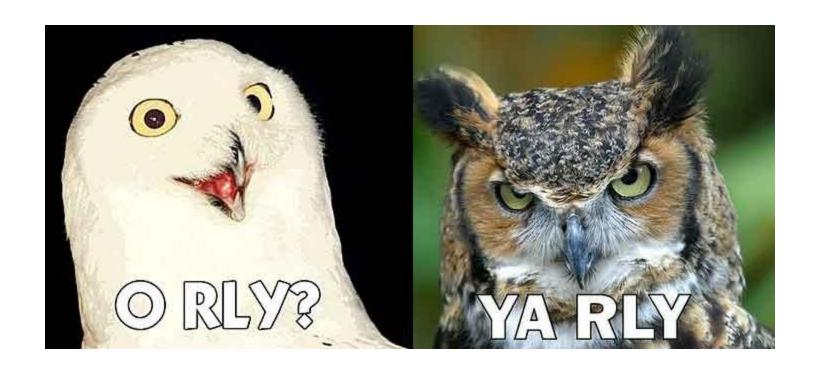
### **Features**



## **Paradigms**

- Procedural
- Object Oriented
- Functional

# **Everything is an object**



## Everything is an object

#### irb(main):001:0> 42.methods

```
=> ["%", "odd?", "inspect", "prec_i", "<<", "tap", "div", "&", "clone", ">>", "p
ublic_methods", "__send__", "instance_variable_defined?", "equal?", "freeze",
"to_sym", "*", "ord", "+", "extend", "next", "send", "round", "methods",
"prec_f,, "-", "even?", "singleton_method_added", "divmod", "hash", "/",
"integer?", "downto", "dup", "instance_variables", "|", "eql?", "size",
"object_id", "instance_eval", "truncate", "~", "id", "to_i", "singleton_methods",
[...]
```

```
irb(main):oo5:o> nil.class
=> NilClass
```

```
irb(main):002:0> nil.hallo_hdm
```

NoMethodError: undefined method `hallo\_hdm' for nil:NilClass from (irb):2

#### Typing: strong / weak

- Strong typing
  - "5"/2 → "NoMethodError"
- Weak typing
  - "16 " / 2 → 8 (e.g. in Perl)

Ruby is strongly typed! (Java too)

#### Typing: explicit/implicit

- Explicit: int a = 5
- Implicit: a = 5

Ruby is implicitly typed! (Java explicitly)

#### Typing: static / dynamic

- Static typing
  - The compiler checks types during compilation
- Dynamic typing
  - The compiler doesn't check types during compilation

Ruby uses dynamic typing (Java uses static typing)

## Blocks



"Blocks are unnamed functions"

#### **Blocks**

```
def foo
              def foo &proc
                                             yield 2
                proc.call 2
                                             yield 4
                proc.call 4
Define:
                proc.call 6
                                             yield 6
                                           end
              end
              foo{|some number|
Call:
                puts some number * 3
              6
Result:
              12
              18
```

#### Blocks Example: Iterators

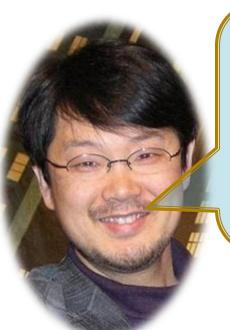
```
The current piece of the collection we are working with

some_collection.each { | item| puts item }

some_collection.select { | item| item =~ /[xz]/ }

some_collection.reject { | item| item =~ /[xz]/ }
```

#### Closures



block.call

A closure object has:

- •code to run (the executable)
- •state around the code (the scope)

So you capture the environment, namely the local variables, in the closure. As a result, you can refer to the local variables inside a closure

```
def method_that_returns_a_block( x )
  some_value = x * 12

return Proc.new { puts "The value of X *was* #{x}, causing some_value to be #{some_value}"}
end

block = method that returns a block(5)
```

#### Closures: Examples

Idea: Function returning a function

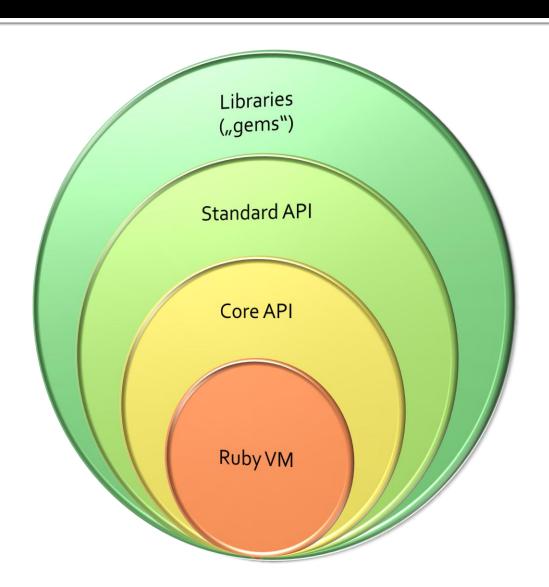
```
# Builds a function that returns true
# when 'f' returns false, and vice versa.
def complement f
  lambda {|*args| not f.call(*args) }
end
```

#### In action:

```
is_even = lambda {|n| n % 2 == 0 }
is_odd = complement(is_even)

is_odd.call(1) # true
is_odd.call(2) # false
```

# Parts of Ruby



#### **Execution Environments**

- Ruby VM (Ruby 1.8)
- YARV (aka Ruby 1.9)
- Rubinius
- MacRuby
- Jruby
- IronRuby
- Hotruby

## Package Management

#### Gem:



Search:

```
C:\Users\Marc>gem search -r mapping
*** REMOTE GEMS ***
dm-mapping (0.7.0)
xml-mapping (0.8.1)
```

Installing:

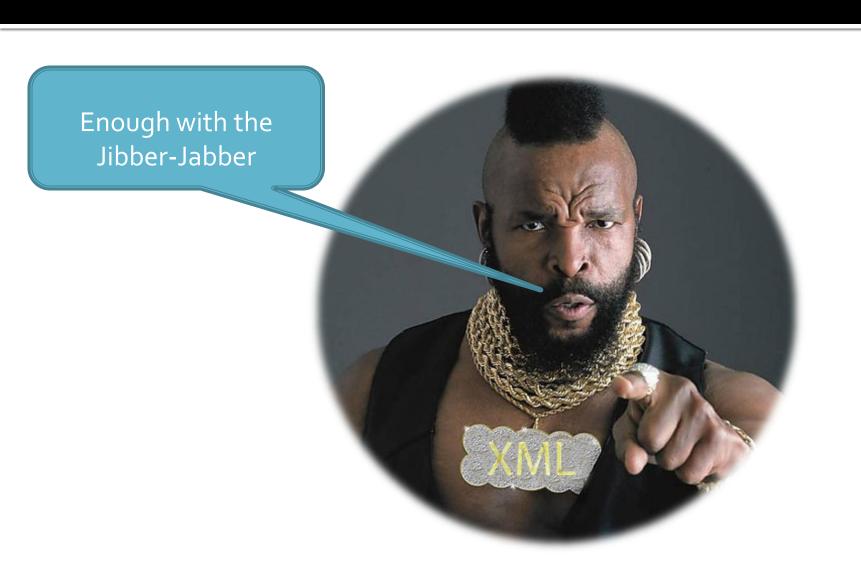
```
C:\Users\Marc>gem install xml-mapping
Successfully installed xml-mapping-0.8.1
1 gem installed
Installing ri documentation for xml-mapping-0.8.1...
Installing RDoc documentation for xml-mapping-0.8.1...
```

#### Package Management

Gem:

Usage:

```
irb(main):001:0> require "rubygems"
=> false
irb(main):002:0> require "hpricot"
=> true
irb(main):003:0> Hpricot.methods
=> ["private_class_method", "inspect", "name", "uxs", "tap", "]
ct", "clone", "public_methods", "__send__", "method_defined?",
e_variable_defined?", "equal?", "freeze", "extend", "autoload",
efined?", "methods", "ancestors", "module_eval", "hash", "dup",
s", "build_node", "public_method_defined?", "instance_variables
```



# **Mapping Libraries**

- XML → Object
  - XML-Object
  - XmlSimple
- XML  $\leftarrow \rightarrow$  Object
  - ROXML
  - XML::MAPPING
  - HappyMapper

# XML-Object

http://xml-object.rubyforge.org/

"Tools like JSON or YAML are a much better fit for this kind of job, but one doesn't always have that luxury."

### XML-Object

... attempts to make the accessing of small, well-formed XML structures convenient, by providing a syntax that fits well in most Ruby programs.

# XML-Object: Usage

13

recipe.instructions.steps.size

```
Krecipe name="bread" prep time="5 mins" cook time="3 hours">
       <title>Basic bread</title>
       <ingredient amount="8" unit="dL">Flour</ingredient>
       <ingredient amount="10" unit="grams">Yeast</ingredient>
       <ingredient amount="4" unit="dL" state="warm">Water</ingredient>
       <ingredient amount="1" unit="teaspoon">Salt</ingredient>
       <instructions easy="yes" hard="false">
8
         <step>Mix all ingredients together.
9
         <step>Knead thoroughly.</step>
         <step>Cover with a cloth, and leave for one hour in warm room.</step>
10
11
         <step>Knead again.</step>
12
         <step>Place in a bread baking tin.</step>
13
         <step>Cover with a cloth, and leave for one hour in warm room.</step>
14
         <step>Bake in the oven at 180(degrees)C for 30 minutes.
15
       </instructions>
     </recipe>
  require 'xml-object'
    recipe = XMLObject.new(File.open('recipe.xml'))
 2
 3
 4
    recipe.name
                                   => "bread"
 5
    recipe.title
                                   => "Basic bread"
 ó
 7
    recipe.ingredients.is a?(Array)
                                   => true
 8
    recipe.ingredients.first.amount => "8" # Not a Fixnum. Too hard. :(
 9
10
    recipe.instructions.easy?
                                   => true
11
12
    recipe.instructions.first.upcase => "MIX ALL INGREDIENTS TOGETHER."
```





# Ò\_ó Ambiguities?

>ruby test.rb
At first, Elements are checked: Bread Recepie
You can get the Attributes though: an awesome recepie for bread

### Features: Adapter

```
1 require 'xml-object' # REXML
2 require 'xml-object/adapters/hpricot' #hpricot
3 require 'xml-object/adapters/libxml' #libxml
```

### Features: Question notation

1 <admin>true</admin>



1 XMLFile.admin?

=> true



# Features: Collection auto folding



```
1  student = XMLObject.new(xml_file)
2
3  student.course.is_a?(Array)  => true
4  student.course.first == 'Math' => true
5  student.course.last == 'Biology' => true
```



### Features: Collection pluralization

```
1 \student>
2 \ \name \Bob \langle \name \rangle
3 \ \course \Math \langle \course \rangle
4 \ \course \German \langle \course \rangle
5 \ \course \Biology \langle \course \rangle
6 \langle \student \rangle
```

```
XML
```

```
student = XMLObject.new(xml_file)

student.courses.first == student.course.first => true
```



### Features: Collection proxy



```
author.publications == author.publications.books => true author.publications.map { |b| b.downcase } => ['math 101', 'biology 101']
```

### ò\_Ó

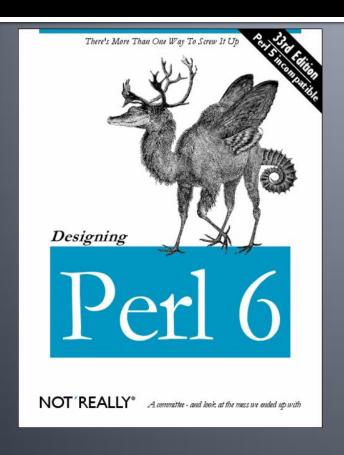
```
1 irb(main):002:0> test = XMLObject.new("2008-12-07_17-36-57.gpx")
2 [FATAL] failed to allocate memory
```

#### Recursive

The design of the adapters assumes parsing of the objects recursively. Deep files are bound to throw SystemStackError, but for the kinds of files I need to read, things are working fine so far. In any case, stream parsing is on the TODO list.

# **XmlSimple**

http://xml-simple.rubyforge.org/



a Ruby translation of Grant McLean's Perl module XML::Simple

# XmlSimple = Rexml + ...

- xml\_in()
- xml\_out()

### Our source

```
<config logdir="/var/log/foo/" debugfile="/tmp/foo.debug">
    <server name="sahara" osname="solaris" osversion="2.6">
      <address>10.0.0.101</address>
      <address>10.0.1.101</address>
    </server>
   <server name="gobi" osname="irix" osversion="6.5">
      <address>10.0.0.102</address>
   </server>
   <server name="kalahari" osname="linux" osversion="2.0.34">
      <address>10.0.0.103</address>
      <address>10.0.1.103</address>
   </server>
  </config>
```

### Our code

```
The input file

A hash containing options

require 'xmlsimple'

config = XmlSimple.xml in('foo.xml', { 'KeyAttr' => 'name' })
```

### Our result: Hash.new

```
'logdir'
            => '/var/log/foo/',
             => '/tmp/foo.debug',
'debugfile'
'server'
 'sahara'
               => {
 'osversion' => '2.6',
   'osname' => 'solaris',
 'address'
                => [ '10.0.0.101', '10.0.1.101' ]
 'gobi'
  'osversion' => '6.5',
  'osname'
             => 'irix',
   'address'
                 => [ '10.0.0.102' ]
 'kalahari'
   'osversion' => '2.0.34',
   'osname'
           => 'linux',
                 => [ '10.0.0.103', '10.0.1.103' ]
  'address'
```

# my personal opinion

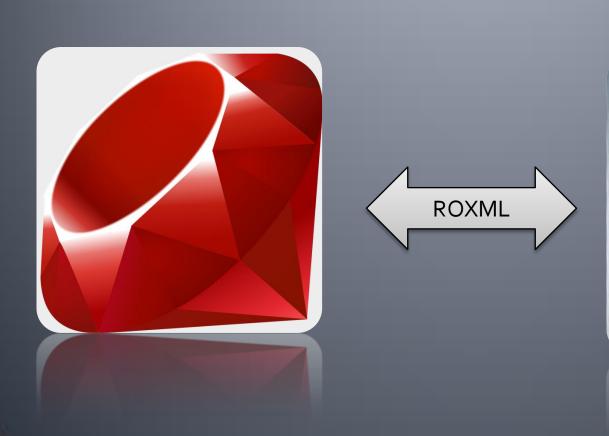
ONE DAY IN THE LIFE OF A PERL PROGRAMER



READING THE CODE FROM THE PREVIOUS DAY

### **ROXML**

http://roxml.rubyforge.org/



```
211 Z Label- Geselist.
<level2_l label="Einkor
    <level3_1 label="P
    <level3 2 label="V
    <level3_3 label="T
    <level3_4 label="
</le>e12 1>
<level2_2 label="Gesu
     <level3_1 label='
     <level3_2 label=
         --13 3 label=
```

# Is it alive?

roxml_1.o_beta	roxml-1.o_beta	June 28, 2006
ROXML 1.0	roxml-1.0.zip	July 1, 2006
ROXML 1.1 Beta	ROXML 1.1 Beta	September 24, 2006
ROXML 1.2	ROXML 1.2	November 10, 2007
ROXML	2.2.0	November 3, 2008



### **ROXML**

#### **ROXML Features**

- Read Ruby objects from XML
- Write Ruby objects to XML
- Annotation-style methods for XML mapping
- One-to-one (composition) Ruby to XML
- One-to-many (aggregation) Ruby with array to XML

### It's all about annotations

```
library>
 <NAME><![CDATA[Favorite Books]]></NAME>
 <books>
   \cook ISBN='0201710897'>
   <title>The PickAxe</title>
   <description><![CDATA[Best Ruby book out there!]]></description>
   <author>David Thomas, Andrew Hunt, Dave Thomas
   </book>
 </books>
</library>
book = Book.new()
 book.isbn = "0201710897"
 book.title = "The PickAxe"
 book.description = "Best Ruby book out there!"
 book.author = "David Thomas, Andrew Hunt, Dave Thomas"
 lib = Library.new()
 lib.name = "Favorite Books"
 lib.books << book
```

### I/O

```
#SAVE
File.open("library.xml", "w") do |f|
lib.to_xml.write(f, 0)
end

#LOAD
lib = Library.parse(File.read("library.xml"))
```

### **XML::MAPPING**

http://xml-mapping.rubyforge.org/



### code, code, code

XML::Mapping::ChoiceNode

```
class Client
  include XML::Mapping
  text node :name, "Name"
  object node :home address, "Address[@where='home']", :class=>Address
  object node :work address, "Address[@where='work']", :class=>Address, :default value=>nil
end
              XML::Mapping::ArrayNode
              XML::Mapping::BooleanNode
              XML::Mapping::HashNode
              XML::Mapping::NumericNode
                                                               single-attribute nodes
              XML::Mapping::ObjectNode
              XML::Mapping::SingleAttributeNode
              XML::Mapping::SubObjectBaseNode
              XML::Mapping::TextNode
```

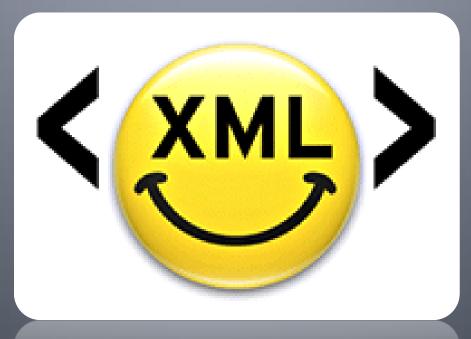
# choice node example

```
class Publication
  include XML::Mapping
  choice node :if, '@author', :then, (text node :author, '@author'),
               :elsif, 'contr', :then, (array node :contributors, 'contr', :class=>String)
 end
 ## usage
p1 = Publication.load from xml(REXML::Document.new('<publication author="Jim"/>').root)
=> #<Publication:Oxb7ad3f38 @author="Jim">
p2 = Publication.load from xml (REXML::Document.new('
<publication>
 <contr>Chris</contr>
<contr>Mel</contr>
 <contr>Toby</contr>
</publication>').root)
=> #<Publication:0xb7ac7ee0 @contributors=["Chris", "Mel", "Toby"]>
```

# HappyMapper

http://happymapper.rubyforge.org/

"Making XML fun again"



## code, code, code

has many

```
xml = << EOF
cproducts>
 oduct>
    <title> A Title</title>
  <features bullets>
     <feature>This is feature text</feature>
     <feature>This is feature text</feature>
   </features bullets>
</product>
</products>
FOF
class FeatureBullet
  include HappyMapper
  tag 'features bullets'
                                  typecasts
 element :feature, String
end
class Product
  include HappyMapper
 element :title, String
 has many :features bullets, FeatureBullet
end
Product.parse(xml).each do |product|
 puts product.title
 product.features bullets.each { |fb| puts " - #{fb.feature}" }
end
# outputs:
 A Title
    - This is feature text
```

### noteworthy:

Camel Case XML Tags to Ruby method names

```
element :total_pages, Integer, :tag => 'TotalPages'
```

### twitter

```
<statuses type="array">
      <status>
        <created at>Sat Aug 09 05:38:12 +0000 2008</created</pre>
 3.
        <id>882281424</id>
 4.
        <text>I so just thought the guy lighting the Olympi
 5.
    the wall. Wow that would have been catastrophic.</text>
6.
        <source>web</source>
7.
        <truncated>false</truncated>
        <in reply to status id>1234</in reply to status id>
8.
        <in reply to user id>12345</in reply to user id>
9.
10.
        <favorited></favorited>
11.
        <user>
12.
          <id>4243</id>
13.
          <name>John Nunemaker
14.
          <screen name>jnunemaker</screen name>
15.
          <location>Mishawaka, IN, US</location>
16.
          <description>Loves his wife, ruby, notre dame foo
17.
          file image url>http://s3.amazonaws.com/twitte
    /Photo 75 normal.jpg</profile image url>
          <url>http://addictedtonew.com</url>
18.
19.
          cted>false
20.
          <followers count>486</followers count>
        </user>
21.
22.
      </status>
23.
    </statuses>
```

```
1. class User
      include HappyMapper
 3.
      element :id, Integer
      element :name, String
      element :screen name, String
      element :location, String
      element :description, String
      element :profile image url, String
      element :url, String
10.
      element :protected, Boolean
11.
      element :followers count, Integer
12.
13. end
14.
    class Status
      include HappyMapper
17.
18.
      element :id, Integer
      element :text, String
19.
      element :created at, Time
20.
21.
      element :source, String
22.
      element :truncated, Boolean
23.
      element :in reply to status id, Integer
      element :in reply to user id, Integer
24.
      element :favorited, Boolean
25.
      has one :user, User
26.
27. end
28.
    statuses = Status.parse(xml string)
    statuses.each do |status|
      puts status.user.name, status.user.scree
32. end
```

# YAML

(rhymes with "camel")



## **Design Goals**

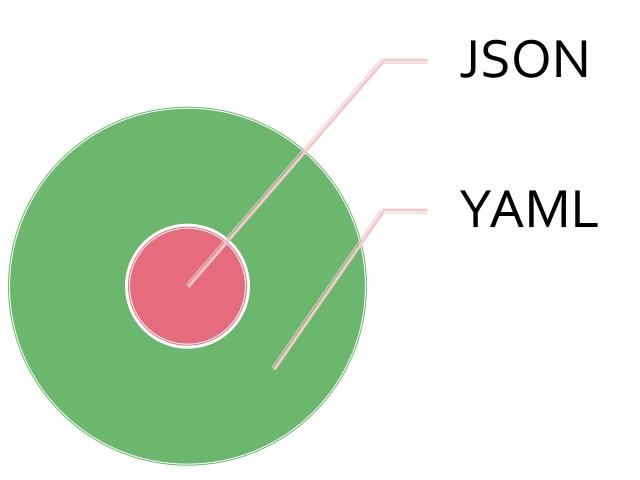
- easily readable by humans.
- matches the native data structures of agile languages.
- portable between programming languages.
- 4. consistent model to support generic tools.
- supports one-pass processing.
- expressive and extensible.
- 7. easy to implement and use.

### JSON $\leftarrow \rightarrow$ YAML

JSON/YAML = human readable data interchange format

- JSON == simplicity + universality
  - trivial to generate and parse
  - reduced human readability
- YAML == human readability + serializing native data structures
  - harder to generate and parse
  - easy to read

### JSON $\leftarrow \rightarrow$ YAML



### **YAML**

#### …is Sequences, Maps, Scalars

- Seq = Array
- Map = Hash
- Scalars = String, Integer, Float, Time, NilClass

## Sequences

```
Sequence:

- apple
- burrito
- egg salad sandwich
```

```
Array: Ruby: ['apple', 'burrito', 'egg salad sandwich']
```

### Maps

Map:

```
YAML:
event: RubyConf.new(2002)
location: Seattle, WA, U.S.A.
start: Nov. 1st, 2002
end: Nov. 3rd, 2002
```

Hash:

```
Ruby:

{
  'event' => 'RubyConf.new(2002)',
  'location' => 'Seattle, WA, U.S.A.',
  'start' => 'Nov. 1st, 2002',
  'end' => 'Nov. 3rd, 2002'
}
```

### **Scalars**

Map of Scalars:

```
integer: 12
float: 766.05
date: 2002-11-01
time: 2002-11-01T15:30:00.00Z
string: Begins with an alphabetic or numeric character.
single-quoted: '12'
double-quoted: "12"
```

Hash of Objects:

```
Ruby:

{ 'integer' => 12,
    'float' => 766.05,
    'date' => Date.new( 2002, 11, 01 ),
    'time' => Time.utc( 2002, 11, 01, 15, 30, 00, 00 ),
    'string' => 'Begins with an alphabetic or numeric character.',
    'single-quoted' => '12',
    'double-quoted' => '12'
}
```

Native typing is implicity determined in plain scalars.

# Code Sample?



## Input/Output

#### YAML output:

```
irb(main):003:0> require "yaml"
=> true
irb(main):004:0> ["Goik", "Kriha", "Schmitz", "Maucher"].to_yaml
=> "--- \n- Goik\n- Kriha\n- Schmitz\n- Maucher\n"
irb(main):005:0> puts ["Goik", "Kriha", "Schmitz", "Maucher"].to_yaml
---
- Goik
- Kriha
- Schmitz
- Maucher
=> nil
```

#### YAML input:

```
irb(main):001:0> require "yaml"
=> true
irb(main):002:0> profs = YAML::load( File.open( 'profs.yml' ) )
=> ["Kriha", "Goik", "Maucher", "Schmitz"]
irb(main):003:0> profs.class
=> Array
```

### >1?



# Ruby and YAML

More than 1 document

```
from: Marc
to: Audience
message: >
Hallo, ich hoffe ihr seid noch wach :) ?

from: Audience
to: Marc
message: >
Klar, bei dem super Vortrag!
```

#### Ruby code

```
YAML::load_documents( File.open( 'message.yml' ) ) { |msg|
  puts "A message from #{msg['from']} to #{msg['to']}:"
  puts msg['message']
}
```

#### Output

```
A message from Marc to Audience:
Hallo, ich hoffe ihr seid noch wach :) ?
A message from Audience to Marc:
Klar, bei dem super Vortrag!
```

## But what about objects?

Won't somebody please think of the children objects!



## Too much text!

Live Demo :D



### Namespace

Problem: The !ruby/object type is only understood by YAML.rb.

```
Solution:
```

```
require 'yaml'
require 'bigdecimal'
#Marshal
class BigDecimal
  def to yaml (opts={})
    YAML::quick emit(object id, opts) do |out|
      out.scalar ("tag:induktiv.at, 2007:BigDecimal", self.to s)
    end
  end
end
#Unmarshal
YAML.add domain type ("induktiv.at, 2007", "BigDecimal") { | type, val |
  BigDecimal.new(val)
```

## Are they allowed to do that?

http://www.kuwata-lab.com/kwalify/

 YAML and JSON are simple and nice format for structured data and easier for human to read and write than XML. But there have been no schema for YAML such as RelaxNG or DTD. Kwalify gets over this situation. Fragen?

# **KTHXBYE**

#### Sources

- Each project's website
- Some useful closures in Ruby <a href="http://www.randomhacks.net/articles/2007/02/01/some-useful-closures-in-ruby">http://www.randomhacks.net/articles/2007/02/01/some-useful-closures-in-ruby</a>
- Kai Jäger: Ajax in der Praxis Grundlagen, Konzepte, Lösungen ISBN-10: 3-540-69333-5
- Using Ruby An Introduction to Ruby for Java Programmers <u>http://onestepback.org/articles/usingruby/index.html</u>
- Ruby for Java Programmers
   http://www.softwaresummit.com/2006/speakers/BowlerRubyForJavaProgrammers.pdf
- Happy Mapper: Making XML fun again:
   <a href="http://railstips.org/2008/11/17/happymapper-making-xml-fun-again">http://railstips.org/2008/11/17/happymapper-making-xml-fun-again</a>
- YAML Working draft 1.2 <a href="http://yaml.org/spec/1.2/">http://yaml.org/spec/1.2/</a>
- YAML Cookbook:
  <a href="http://www.nt.ntnu.no/users/haugwarb/Programming/YAML/YAML\_for\_ruby.html">http://www.nt.ntnu.no/users/haugwarb/Programming/YAML/YAML\_for\_ruby.html</a>