



XML (almost) all the way

Experiences with a small-scale journal publishing system

Peter Flynn

IT Services Electronic Publishing Unit
11 August 2015

**A TRADITION OF
INDEPENDENT
THINKING**



UCC

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Contents

Background: demand and supply for e-journals

Requests

Alternative scenarios

CP and friends

Serving

Lessons learned

All done

Background

E-journals: the demand



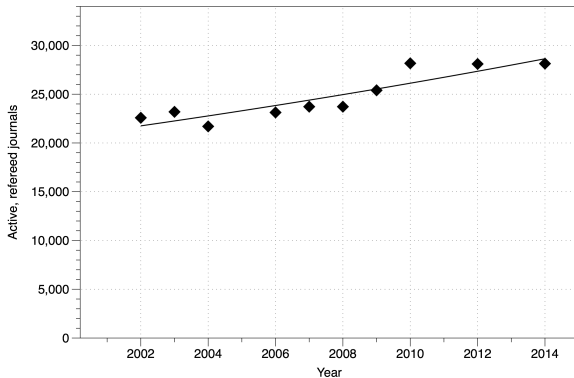
ALTHOUGH THEY KNEW THE READERSHIP WOULD BE RATHER MODEST, THE 2 STRANDED EPIDEMIOLOGISTS FELT THEY NOW HAD ENOUGH "N OF 1" TRIALS TO START A JOURNAL.

Bastian

Everyone wants an e-journal

Cartoon by Hilda Bastian,
hildabastian.net

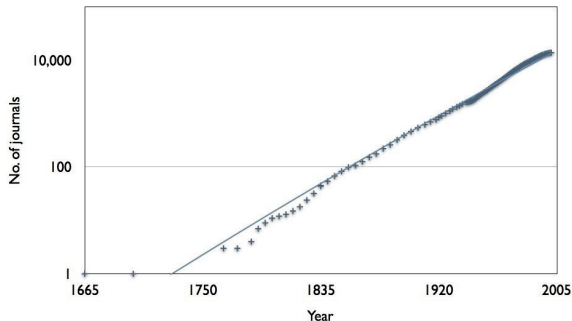
E-journals: the trend



Growth in the number of active, peer-reviewed journals recorded in Ulrich's directory, 2002-2012

STM 2015 Report, 4th ed http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

E-journals: the history



Growth in the number of active, peer reviewed learned journals since 1665 (after Mabe, 2003)

STM 2015 Report, 4th ed http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

Background

E-journals: the effort



Getting an e-journal off the ground is *hard*

Image by Isabel G. Méijome
<http://www.ejournal.fi/a-journey>




Task list

- Name
- Logo
- Tagline
- Size
- Languages
- Frequency
- Hosting
- Format
- Design
- Mug, T-shirt, etc
- Topic and constraints
- Editors
- Authors
- Articles
- Word? HTML5?
- Paper?
- Advertisements
- Announcement and PR
- Stylesheet
- Access

Original scenario

Journal for Drama and Theatre in Foreign and Second Language Education

- Name: Scenario
- Logo: 
- Tagline: as above
- Size: 10 articles/issue
- Languages: EN/DE
- Frequency: 2× a year
- Hosting: onsite
- Format: HTML and PDF
- Design: **none**
- Mug, T-shirt, etc: \Smiley
- Topic: as above
- Editors: ✓
- Authors: ✓
- Articles: underway
- Word? ✓
- Paper? ✗
- Advertisements: ✗
- Announce/PR: **none**
- Stylesheet: **what?**
- Access: Open (CC)

Original scenario

Markup

Parentheses ()

(Curly) braces { }

(Square) brackets []

Angle brackets < >

Pointy brackets < >

Ampersand &

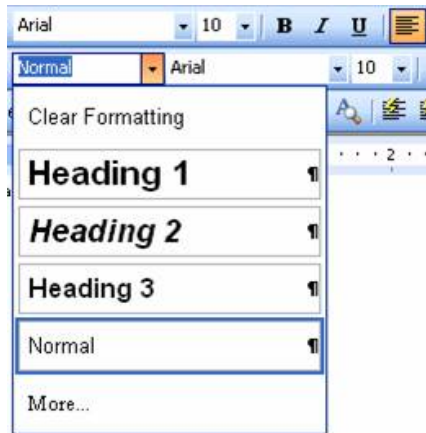
Backslash \

No-one has ever heard of these
(well, maybe parentheses)

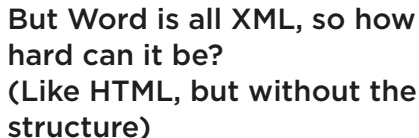
Rule 1: no author (and preferably no editor) is ever to see or
interact with raw markup (unless they ask nicely)

Original scenario

Markup: Word styles

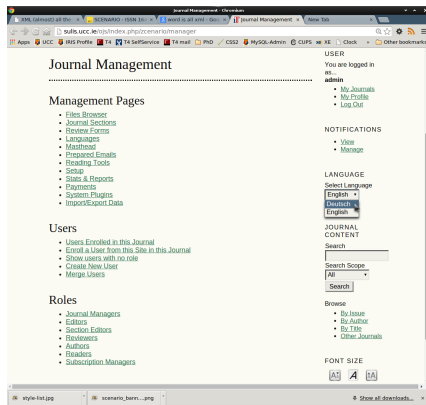


No-one has ever heard of these, either, but that never stopped anyone



Alternative scenarios

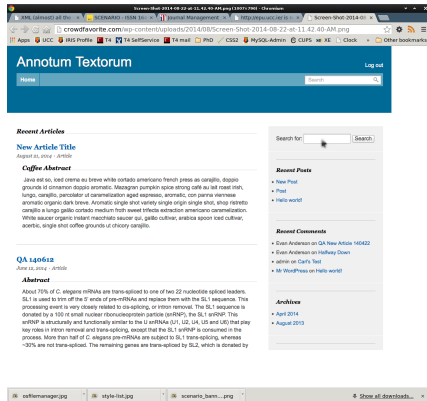
Open Journal System (OJS)



Wonderfully comprehensive but
requires extensive configuration

Alternative scenarios

Annotum in Wordpress



Almost too flexible

Alternative scenarios

Homebrew configuration

```
<?xml version="1.0"?>
<!DOCTYPE journal SYSTEM "../journal.dtd">
<journal name="footsteps" ISSN="0000-0000">
  <name>The Boolean</name>
  <uri>http://publish.ucc.ie/boolean/</uri>
  <editors type="academic">
    <editor>Alan Kelly</editor>
    <editor uri="A014/gallen">Graham
      Allen</editor>
    <editor uri="A014/omurphy">Orla
      Murphy</editor>
  </editors>
  <editors type="student">
    <editor>Aimee Brennan</editor>
    <editor>Meghan Coakley</editor>
    <editor>Anne-Marie Devlin</editor>
    <editor>Farhan Manzour</editor>
    <editor>Jane Power</editor>
  </editors>
```

```
<address uri="/graduatestudies/">Graduate
  Studies Office</address>
<email>graduatestudies@ucc.ie</email>
<description>Snapshots of Doctoral Research
  at University College Cork</description>
<icon uri="/images/boolean.png"/>
<structure volume="vol" issue="iss"
  volno="year" start="2010" maxiss="1"/>
<langs>
  <lang code="en" default="yes" name="English">
    <trans id="intro" term="introduction"/>
    <trans id="foreword" term="foreword"/>
    <trans id="vol" term="Volume"/>
    <trans id="iss" term="Issue"/>
    <trans id="abs" term="Abstract"/>
    <trans id="arts" term="Articles"/>
    <trans id="printbuttonlabel"
      term="Printer-friendly"/>
  </lang>
```

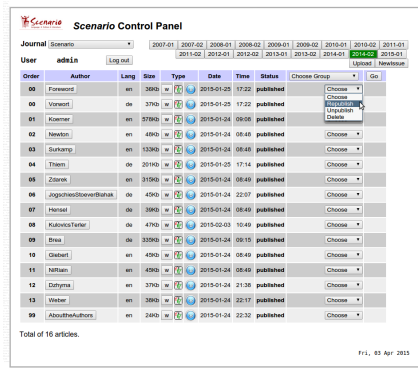

Uploader



osFileManager for uploading articles

CP and friends

Control panel



Scenario Control Panel

Journal: Scenario | 2007-01 | 2007-02 | 2008-01 | 2008-02 | 2009-01 | 2009-02 | 2010-01 | 2010-02 | 2011-01 | 2011-02 | 2012-01 | 2012-02 | 2013-01 | 2013-02 | 2014-01 | **2014-02** | 2015-01

User: **admin** | Log out | Upload | Newsletter

Order	Author	Lang	Size	Type	Date	Time	Status	Choose Group	Go
00	Foreword	en	30Kb	w	2015-01-25	17:22	published	Choose	
00	Voorwoord	de	37Kb	w	2015-01-25	17:22	published	Choose	
01	Koerner	en	578Kb	w	2015-01-24	09:06	published	Choose	
02	Newton	en	48Kb	w	2015-01-24	08:48	published	Choose	
03	Sunkamp	en	133Kb	w	2015-01-24	08:48	published	Choose	
04	Thiem	de	201Kb	w	2015-01-25	17:14	published	Choose	
05	Zisarik	en	315Kb	w	2015-01-24	08:49	published	Choose	
06	JogschiesStoeverBrahak	de	45Kb	w	2015-01-24	22:07	published	Choose	
07	Hansel	de	36Kb	w	2015-01-24	08:49	published	Choose	
08	KulovicsTanner	de	47Kb	w	2015-02-03	10:49	published	Choose	
09	Brea	de	335Kb	w	2015-01-24	09:15	published	Choose	
10	Giebert	en	45Kb	w	2015-01-24	08:49	published	Choose	
11	Narfaen	en	45Kb	w	2015-01-24	08:49	published	Choose	
12	Dzhyra	en	37Kb	w	2015-01-24	21:36	published	Choose	
13	Weber	en	36Kb	w	2015-01-24	22:17	published	Choose	
99	AbouttheAuthors	en	24Kb	w	2015-01-24	22:32	published	Choose	

Total of 16 articles.

Fri, 03 Apr 2015

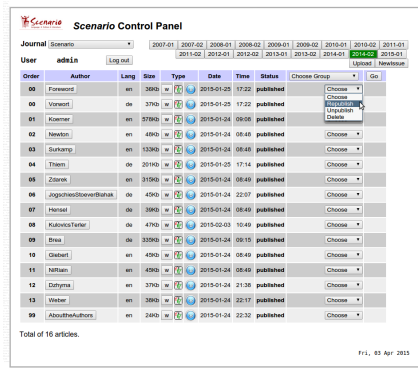
Control Panel framed in Cocoon
with background support from
shell script

CGI script

- Processes new articles :
 - Unzips document
 - Checks metadata
 - Extracts images and converts to PNG
- Sets or unsets the 'published' flag
- Stores/updates metadata in a Papers file
- Returns list of articles and metadata as XML to the CP

CP and friends

Control panel (again)



Scenario Control Panel

Journal: Scenario | 2007-01 | 2007-02 | 2008-01 | 2008-02 | 2009-01 | 2009-02 | 2010-01 | 2010-02 | 2011-01 | 2011-02 | 2012-01 | 2012-02 | 2013-01 | 2013-02 | 2014-01 | 2014-02 | 2015-01 | 2015-02 | Upload | New Issue

User: admin | Log out

Order	Author	Lang	Size	Type	Date	Time	Status	Choose Group	Go
00	Foreword	en	309Kb	W	2015-01-25	17:22	published	Choose Choose Preview Markup Delete	
00	Vorwort	de	379Kb	W	2015-01-25	17:22	published		
01	Koerner	en	578Kb	W	2015-01-24	09:06	published		
02	Newton	en	48Kb	W	2015-01-24	08:48	published	Choose	
03	Sunkamp	en	133Kb	W	2015-01-24	08:48	published	Choose	
04	Thiem	de	201Kb	W	2015-01-25	17:14	published	Choose	
05	Zsarek	en	315Kb	W	2015-01-24	08:49	published	Choose	
06	JogschiesStoeverBrahak	de	45Kb	W	2015-01-24	22:07	published	Choose	
07	Hansel	de	36Kb	W	2015-01-24	08:49	published	Choose	
08	KulovicsTarlner	de	47Kb	W	2015-02-03	10:49	published	Choose	
09	Brea	de	335Kb	W	2015-01-24	09:15	published	Choose	
10	Giebert	en	45Kb	W	2015-01-24	08:49	published	Choose	
11	Nirbaen	en	45Kb	W	2015-01-24	08:49	published	Choose	
12	Dzuma	en	37Kb	W	2015-01-24	21:36	published	Choose	
13	Weber	en	36Kb	W	2015-01-24	22:17	published	Choose	
99	AbouttheAuthors	en	24Kb	W	2015-01-24	22:32	published	Choose	

Total of 16 articles.

Fri, 03 Apr 2015

- Upload articles
- Create new issue
- Preview article
- Article download
- Markup tree
- Styles list
- Publish/test/delete
- Adds page numbers to Papers file when published

Style management maps to HTML and \LaTeX

```
<styles id="boolean">
  <style word="Title" html="h1" latex="title"/>
  <style word="Author" html="h2" latex="author"/>
  <style word="MainText" html="p" latex="par"/>
  <style word="BodyText" html="p" latex="par"/>
  <style word="Normal" html="p" latex="par"/>
  <style word="Affiliation" html="address" latex="affil"/>
  <style word="Acknowledgement" html="address" latex="ack"/>
  <style word="Bibliography" html="li" wrapper="ol"
    prefix="References" prefix-element="lh" latex="bibitem{}"
    latex-wrapper="thebibliography" wrapper-style="{99}"/>
  <style word="ListBullet" html="li" wrapper="ul" latex="item{}"
    latex-wrapper="itemize"/>
  <style word="ListNumber" html="li" wrapper="ol" latex="item{}"
    latex-wrapper="enumerate"/>
  <style word="Footer" treatment="omit" html="" latex=""/>
  <style word="Heading1" html="h2" link="ToC" latex="section"/>
  <style word="Heading2" html="h3" latex="subsection"/>
```


HTML (and preview)

System und Bedeutung: Zum Kulturtheater in Deutschland

Stefan Hajduk

ISSN 1649-8528 - Jahrgang 2007 - Ausgabe 02

Sprache • Kultur • Literatur

Index of articles

Abstract

Das deutsche Theater zeichnet sich durch sein fischschwimmendes System kommunaler Bühnen in öffentlicher Trägerschaft aus. Die weitgehend aus Stauspielern geleistete Finanzierung wirft indes immer wieder Fragen nach der gesellschaftlichen Funktion und gesamtgesellschaftlichen Bedeutung auf. Diese lassen sich adäquat nur aus einer spezifischen Nationalgeschichte verstehen, in der ästhetische Bildungsideen politische Entwicklungsstrukturen überlagert haben. Die zumal im internationalen Vergleich hohe Qualität des deutschen Kulturtheaters einerseits, seine öffentliche Wahrnehmung andererseits weisen allerdings einen Mangel an reflexiver Funktionalität auf. Dieses Defizit an kultureller Selbstreflexion scheint selbst ein Inszenierungseffekt der Mediengesellschaft zu sein, in der Theatralität und Realität zunehmend konvergieren.

1. Barock und Gegenwart?

„All the world's a stage“, sagte Shakespeare um 1600. Haben Sie schon einmal darüber nachgedacht was er damit im elisabethanischen England meinte und was es für unsere Welt heute heißen könnte? Sah er etwa bereits am Beginn der Neuzeit eine unaufhaltsame Verflüchtigung des Realen in den Simulakren der Theaterästhetik am Werk? Darin ähnlich einer totalisierten Simulationswirklichkeit wie sie in apokalyptischen Gegenwartsdiagnosen für die poröse Welt der Postmoderne ebenso kennzeichnend wie unwiderrufbar sein soll? Oder können wir mit Shakespeare und seit dem Barockzeitalter unsere individuellen und sozialen Lebenswelten mit ihren Inszenierungsformen inmitten einer Mediengesellschaft nur besser verstehen lernen, indem wir ihre Realitätsansprüche in Analogie zur Bühnensatzbetschen Illusion dekonstruieren? Die pompöse Repräsentation des Adels einerseits, die reflexive Verunsicherung der subjektivierten Existenz im 17. Jahrhundert andererseits legen eine solche Diachronie von theater-, medien- und gesellschaftsgeschichtlichen Überlegungen nahe.

Auch aus dem Spanien des Siglo de oro verbreitete sich die Rede vom *El arca teatro del mundo* (Calderón de la Barca)

Article on preview



- Articles typeset on demand
- Cached for speed
- Catenated for issue PDF
- Page numbers extracted back to Papers file

Lessons learned

Don't do this at home, folks

- Simple is hard
- Hiding pointy brackets means error-compensation, not error-correction
- Authors believe they can control and invent the format inconsistently: editors' job to mediate
- Word has many hidden nooks and crannies; if you proceed, you will likely be eaten by a grue
- Issue management, DOI assignment, EPUBs
- Ultimately the editors' choice of systems

Thank you

