

Import WordPress into TYPO3

Migrating away from WordPress after 9 years



About Reelworx

- Founded in 2014
- Based in Linz, Austria
- Main focus on TYPO3 based web applicances
- Extensions from us you should know:
 - rx_shariff
 - rx_unrollsavebuttons
 - rx_scheduled_social
 - rx_fscmigrate
 - openid
- <https://reelworx.at>
- <https://blog.reelworx.at>



Johannes Kasberger,
Co-Founder

Slack: @jkasberger

Mail: johannes.kasberger@reelworx.at

The Site: Austrian Wings

- <https://www.austrianwings.info>
- News Portal with focus on aviation, non-commercial project
- Different work tasks (photographer, writer, ...)
- Started as WordPress site with
 - > 25.000 articles
 - > 27.000 pictures
 - ~ 1 million page loads / ~ 180.000 Unique Visitors per month
- Major Problem: Quality of stored data not good enough (inline image tags), old template tightly coupled with content
- Goals: Better content structure, responsive design, keep feature set

Why TYPO3?

- Data model is much better (FAL, image references, ...)
- Translations possible (future feature requirement)
- Personal interest: How complex is switching from WordPress to TYPO3?

WordPress „Data Model“

- Users are stored in `wp_users` table
- Tags und Categories in `wp_term_taxonomy` (name in `wp_terms`)
- Posts in `wp_posts`
`posts_type='POST'` and `post_status='zpublish'`
- Picture Metadata in `wp_posts`
`post_status='inherit'` and `post_type='attachment'`
- Preview image in `wp_postmeta`
(`meta_key='_thumbnail_id'`)
- Redirects in `wp_postmeta`
`meta_key='_wp_old_slug'`

WordPress „Data Model“

- Many relations stored in HTML
- Pictures are tags
- Picture caption copied from meta data entry, but no real relation exists
- Internal links are <a> tags
- Related posts also just links (old WordPress theme)
- Updated at date also stored as text

WordPress „Data Model“

has to do with



Import process

- Picture import to FAL was easy (copy files, run Update Storage Index Task)
- General idea: Insert data in new table, store mapping old id => new id
- User import: Insert into **be_user**
 - Problem: some default values were missing (TYPO3 Bug: Core TCA != SQL defaults)
 - Hashed passwords continued to work
- Tags & Categories
 - WordPress categories imported to **sys_category**
 - WordPress tags imported to **tx_news** tag
 - Invalid references to deleted articles present

Posts

- Posts import
- New column for WordPress ID and WordPress permalink
- Create category & tag mapping (30,000 tags stored)
- Import picture meta data (title + description)
- Import preview image (`sys_file_reference` entries)
- Convert WordPress redirects to Apache redirects
- Initial import done
- Images + links still inline html!

Content import

Multiple iterations and variations of the importer necessary.

1. try: regular expressions to extract HTML content

2. try: DOMDocument to parse HTML

Time intensive task, many authors, many versions resulted in very different content

Pictures converted to **sys_file_reference**

ViewHelper replaces [picture] tags with actual pictures

Videos stored in FAL / **sys_file_reference** created

ViewHelper replaces [video] tags with video

„Other News“ imported as **tx_news_domain_model_news_related_mm**

Runtime of import: 45 minutes

Additional features

- Automated Social Media publishing with Scheduler Task
- Fetch Facebook Page Likes/Twitter Follower count (for footer, privacy)
- News / Core / RealUrl Cache Management (join our follow-up session)
- Some hooks
 - Let RealUrl generate URLs via CLI
 - Core hooks to handle old WordPress URLs
 - DataHandler: automatically fill fields to achieve similar behavior like WordPress
- Optimized Backend for our editors (Live Demo)

TYPO3 in production

- Import successful
- Old features preserved (Social Publishing, Sitemap, RSS Feed)
- Austrian Wings workflow still possible / improved

- How do the editors like TYPO3 compared to WordPress?
- What's missing in the everyday usage?
- Other problems?

Lessons learned

- New RealUrl feature necessary to generate 404 errors for incorrect requests (already released)
- File search in BE offered only very limited options
 - Multi-Keyword Search: implemented by us, already merged in CMS 8
 - Search in meta data not possible with filelist module
 - Ext:media offers meta data search but no multi-keyword search
- Options for Suggest Wizard were not sufficient (merged for TYPO3 8)
- Missing feature to preview [picture] tags embedded in news text within RTE (plugin for CKeditor in CMS 8?!)

Lessons learned (2)

- Page title on paginated pages is always the same (SEO)
- Core suggest wizard exceptions within flexforms, e.g. news tags selection (reason still unclear)
- Response time for Backend and Frontend sufficiently fast
- Editors like TYPO3 (except for the missing features ;-)
- Website user feedback also very positive

„Es wirkt einfach viel wertiger und allein der Editor für die Texte ist ein Traum.“

“... bin ich echt gut im Workflow und finde TYPO3 einfach nur saugeil. Ich denke die Umstellung war eine absolut richtige Entscheidung.“

Questions?