

Conférence Microclub

Outils collaboratifs pour la rédaction et l'édition

Nicolas Jeanmonod

11 mars 2016, Microclub, EPFL

Fil rouge Présentation des points abordés



Chaîne éditoriale

- Édition des textes en Markdown pour les cours
- Édition des textes en Markdown + HTML pour les dias
- Génération des PDF pour les cours et les dias
- Mise en ligne et collaboration avec Git et GitHub
 - Corrections
 - Pull request
 - Intégration des corrections

Fil rouge Présentation des points abordés



Contexte

- MOOC LED de Pierre-Yves Rochat (http://pyrochat.github.io/mooc-led/)
- Besoin d'outils permettant la collaboration
 - L'intégrateur doit rester maitre du projet ⇒ Forking Workflow
 - L'historique doit être disponible
 - Les révisions doivent pouvoir être comparées
- Langages de balises
 - Permettent de contrôler directement le format
 - En format texte ⇒ peuvent être comparés facilement
 - HTML
 - Markdown

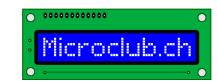
Fil rouge Pourquoi je vous raconte tout ça ?

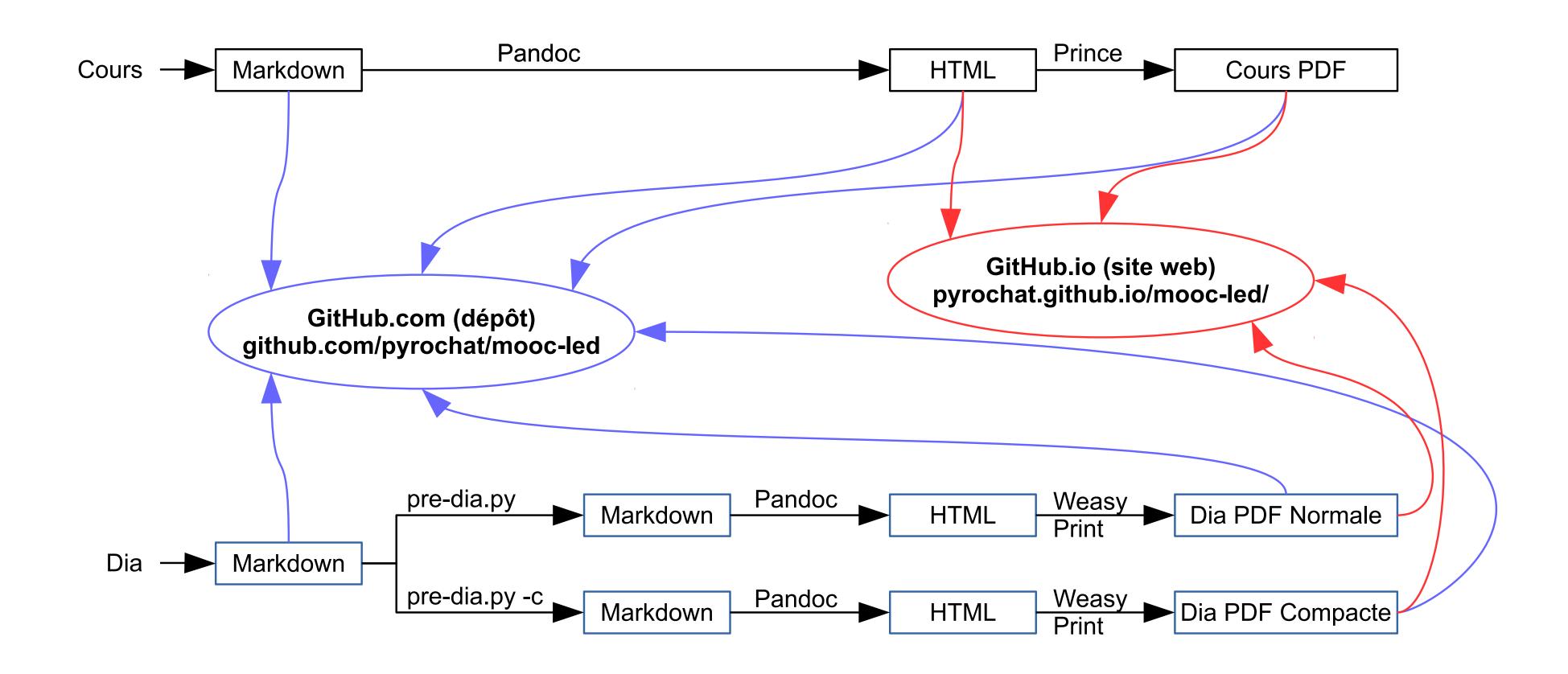


Parce que les techniques montrées dans cette conférence peuvent être utilisées lorsqu'on a besoin :

- D'une grande cohérence dans la mise en forme des fichiers finaux
- De plusieurs formats de sortie (HTML, PDF, Kindle...) avec un seul format d'entrée (Markdown ou HTML)
- De pouvoir gérer les contributions des différents collaborateurs
- D'automatiser certaines tâches
- De gérer des documents volumineux

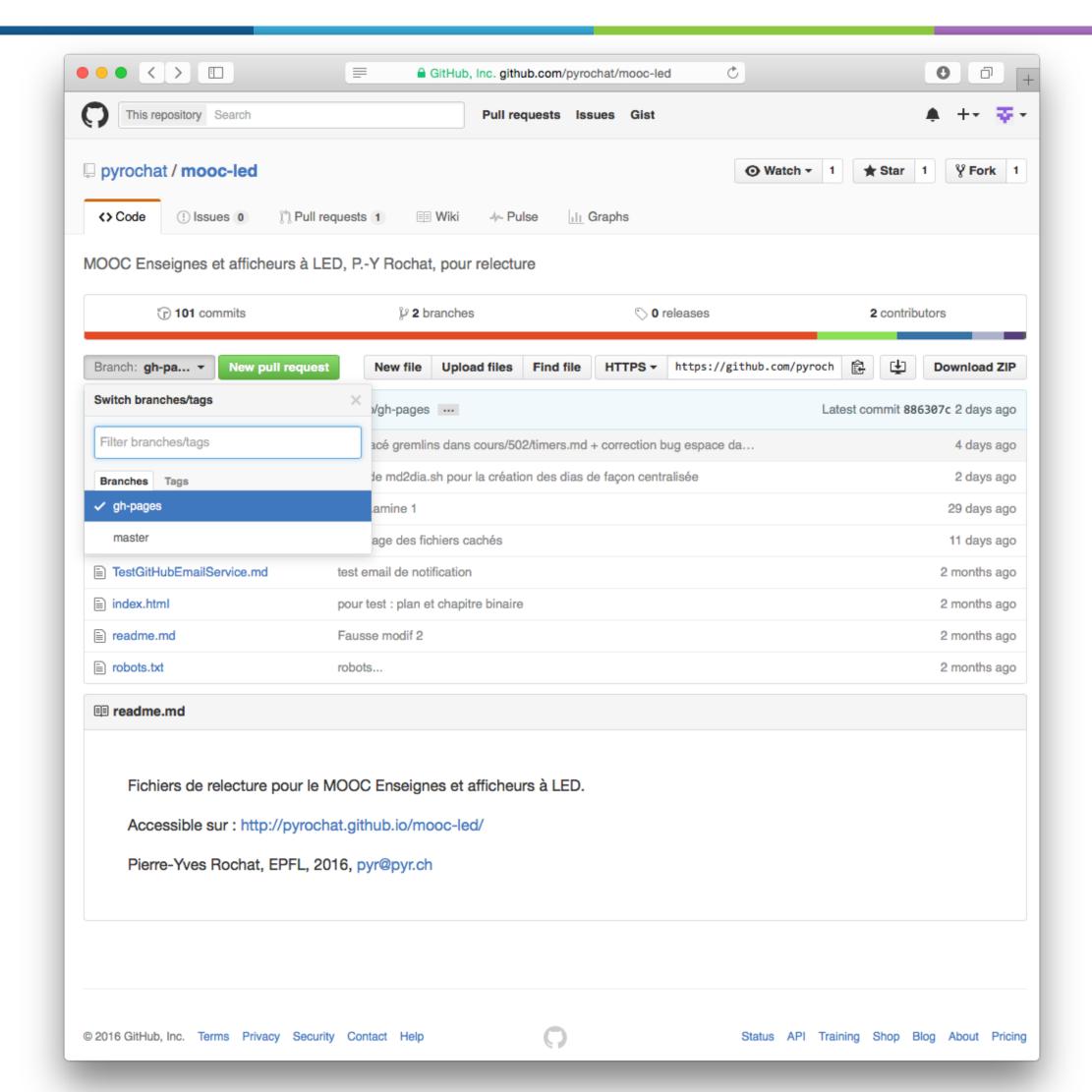
Flux de production

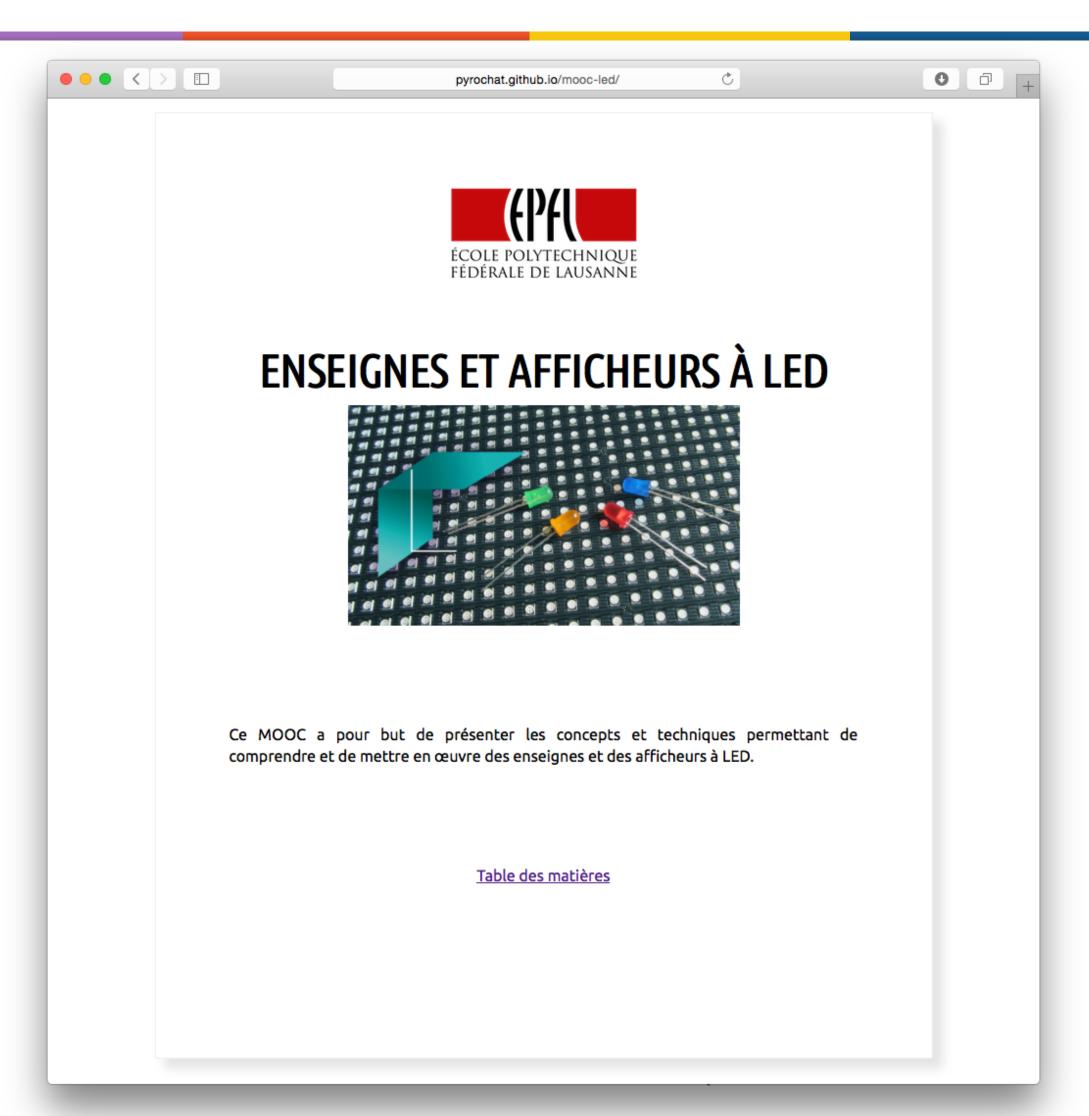




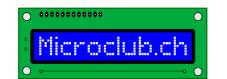
GitHub.com et GitHub.io







Langages de balisage



Un langage de balisage enrichit l'information textuelle brute d'une information sémantique.

Comparaison processus WYSIWYM vs WYSIWYG

Markdown

Un titreÉcrire un paragraphe- Faire une liste

HTML

Traitement de texte

```
Un•titre¶
Écrire•un•paragraphe¶
• → Faire•une•liste¶
```

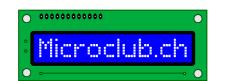
Avantages des langages de balisage



- Séparation sémantique/visuel
- Basés sur des fichiers textes
 - éditables avec des éditeurs simples
 - faciles à comparer
 - faciles à gérer avec des logiciels SCM (Git,...)
 - résilients à l'obsolescence informatique

- Facile d'écrire des scripts pour automatiser les tâches récurrentes
- Nombreuses solutions logicielles
- Nombreux formats de sortie à partir d'une source unique
- Normes libres (non-propriétaires)

Langages de balisage



- SGML (ancêtre du HTML)
- HTML
- XML (⇒ nombreuses grammaires, ex: XHTML, DocBook, SVG, …)
- TeX/LaTeX

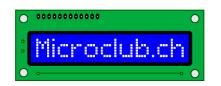
Langages de balisage légers



- Markdown
- Pandoc Flavored Markdown (PFM)
 - = Pandoc's extended Markdown
- GitHub Flavored Markdown (GFM)
- MultiMarkdown
- Markdown Extra

- Textile
- Asciidoc
- reStructuredText
- kramdown
- Fountain

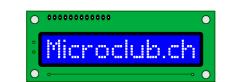
Désavantages des langages de balisage



- Difficiles d'accès aux non-techniciens
- Trop complexes pour être édités directement (HTML)
- Beaucoup de choix tue le choix
- Trop nombreuses versions d'un même vocabulaire
 - Markdown ⇒ PFM, GFM,...
 - HTML ⇒ DOCTYPE : HTML 5, HTML 4.01
 Strict, HTML 4.01 Transitional...
- Incompatibilité des agents utilisateurs (IE6...)

- Nombreux jeu de caractères codés (utf-8, Windows-1256,...). Moins problématique aujourd'hui puisque l'UTF-8 s'est imposé (86% des sites internet l'utilise)
- Nécessitent souvent l'apprentissage de plusieurs vocabulaires (Markdown + HTML + CSS)
- Seul LaTeX est excellent pour le rendu des équations mathématiques, bien qu'il existe des solutions (MathJax par ex.) pour les autres langages.

Normes Markdown



- Norme de base by John Gruber
 - http://daringfireball.net/projects/markdown/
- GitHub Flavored Markdown (GFM)
 - https://help.github.com/categories/writing-on-github/
- Pandoc (PFM)
 - http://pandoc.org/README.html#pandocs-markdown
 - EPFL-ENAC-IT Jean-Daniel Bonjour ⇒ Excellente ressource ! http://enacit1.epfl.ch/markdown-pandoc/
- CommonMark ⇒ projet de standardisation
 - http://commonmark.org/



Le Markdown est conçu pour être :

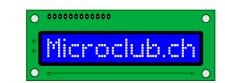
- facile à lire et à écrire
- être transformé en HTML

```
En Markdown traditionnel un paragraphe se termine avec deux retours à la ligne. J

Donc les 3 lignes ci-dessus ne forment qu'un seul paragraphe.
```



```
Autre syntaxe pour niveau 1
## Titre de niveau 2
Syntaxe alternative pour niveau 2
### Titre de niveau 3 incluant [un lien] (http://example.com/)
#### Titre de niveau 4
```



```
Paragraphe normal
> Début du bloc de citation
> * élément de liste
> * second élément
>
>> Décalage de 2ème niveau, usage de styles
Markdown, <u>balises</u> HTML
>
> Bloc de code dans la citation
Retour à l'alignement normal
```



Faire un retour à la ligne
>

- En Markdown classique, ce n'est pas possible (ce qui est quand même incroyable!)
- En MultiMarkdown et en GFM, il faut ajouter deux espaces à la fin de la ligne précédant le saut.
- Avec Pandoc, il faut ajouter une espace et un caractère \ à la fin de la ligne précédant le saut.

Pandoc Flavored Markdown (PFM)



Le PFM est aussi appelé "Pandoc's extended Markdown"

- Supporte les retours à ligne

 br/> en ajoutant un "\" en fin de ligne
- Peut ajouter des attribut HTML: # Exemple { attribut="10cm" }

Pour les détails

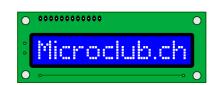
- http://pandoc.org/README.html#pandocs-markdown
- EPFL-ENAC-IT Jean-Daniel Bonjour http://enacit1.epfl.ch/markdown-pandoc/

Exemple de conversion avec Pandoc



```
pandoc
                                        # doc HTML entier
    --standalone
    --from markdown
                                        # format source
                                        # format cible
    --to html
                                       # rendu de code
    --highlight-style pygments
    --email-obfuscation references
                                        # anti-spam
    --normalize
                                        # enlève doubles espaces
    --css=$CSS FILE
                                        # fichier CSS
                                        # fichier de sortie
    -- output output.html
                                        # fichier en entrée
    input.md
```

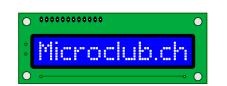
Exemple de conversion avec Pandoc



```
    microcontroleur.html

 1 % Introduction au microcontrôleur
                                                                                1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w</pre>
 2 % [Pierre-Yves Rochat](mailto:pyr@pyr.ch), EPFL
                                                                                    <html xmlns="http://www.w3.org/1999/xhtml">
 3 % rév 2016/01/30
                                                                                    <head>
                                                                                      <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
                                                                                      <meta http-equiv="Content-Style-Type" content="text/css" />
                                                                                      <meta name="generator" content="pandoc" />
       Système informatique dans un circuit intégré ##
                                                                                      <meta name="author" content="Pierre-Yves Rochat, EPFL" />
 7
    Pour commander des enseignes ou des afficheurs à LED, c'est très souvent un
                                                                                      <title>Introduction au microcontrôleur</title>
                                                                                      <style type="text/css">code{white-space: pre;}</style>
 9
    Nous sommes tous familiers avec les systèmes informatiques, à commencer par
                                                                                      <link rel="stylesheet" href="../../statiques/style.css" type="text/css" />
                                                                                10
11
                                                                                11
                                                                                    </head>
      Un **processeur**, qui exécute les instructions. Il est cadencé par une ho
                                                                                    <body>
      Une **mémoire morte** (ROM : *Read Only Memory*) qui contient les instruct 13
                                                                                    <div id="header">
      Une **mémoire vivre** (RAM : *Random Access Memory*, qui permet de stocker
                                                                                    <h1 class="title">Introduction au microcontrôleur</h1>
      Des circuits d'**Entrée-Sortie** (I/O : *Input/Output*). Ce sont les circu
                                                                                    <h2 class="author"><a href="&#x6d;&#x61;&#x69;&#108;&#116;&#x6f;&#58;&#112;&
                                                                                    <h3 class="date">rév 2016/01/30</h3>
16
    Des lignes électriques, appelées *bus* relient ces éléments entre eux.
                                                                                    </div>
                                                                                17
                                                                                    <h2 id="système-informatique-dans-un-circuit-intégré">Système informatique
18
    ![Architecture d'un système informatique](images/architecture-sys-info-200dp
                                                                                    Pour commander des enseignes ou des afficheurs à LED, c'est très souvent
                                                                                    Nous somme tous familier avec les systèmes informatique, à commencer par
20
    ## Microcontrôleur ##
                                                                                    <l
                                                                                21
21
22
                                                                                    un <strong>processeur</strong>, qui exécute les instructions. Il est cad
       **microcontrôleur** est aussi un système informatique. Sa particularité e
                                                                                    24 24 24 cli>une <strong>mémoire vivre</strong> (RAM : <em>Random Access Memory</em>
24
      La mémoire morte contient généralement de 1 kB à quelques centaines de kil 25
                                                                                   des circuits d'<strong>Entrée-Sortie</strong> (I/O : <em>Input/Output
26 * Le processeur est cadencé à des fréquences de quelques mégahertz ou dizain 26 
☐ Line 67, Column 31
                                                                                                                                          Tab Size: 4 HTML Liquid
```

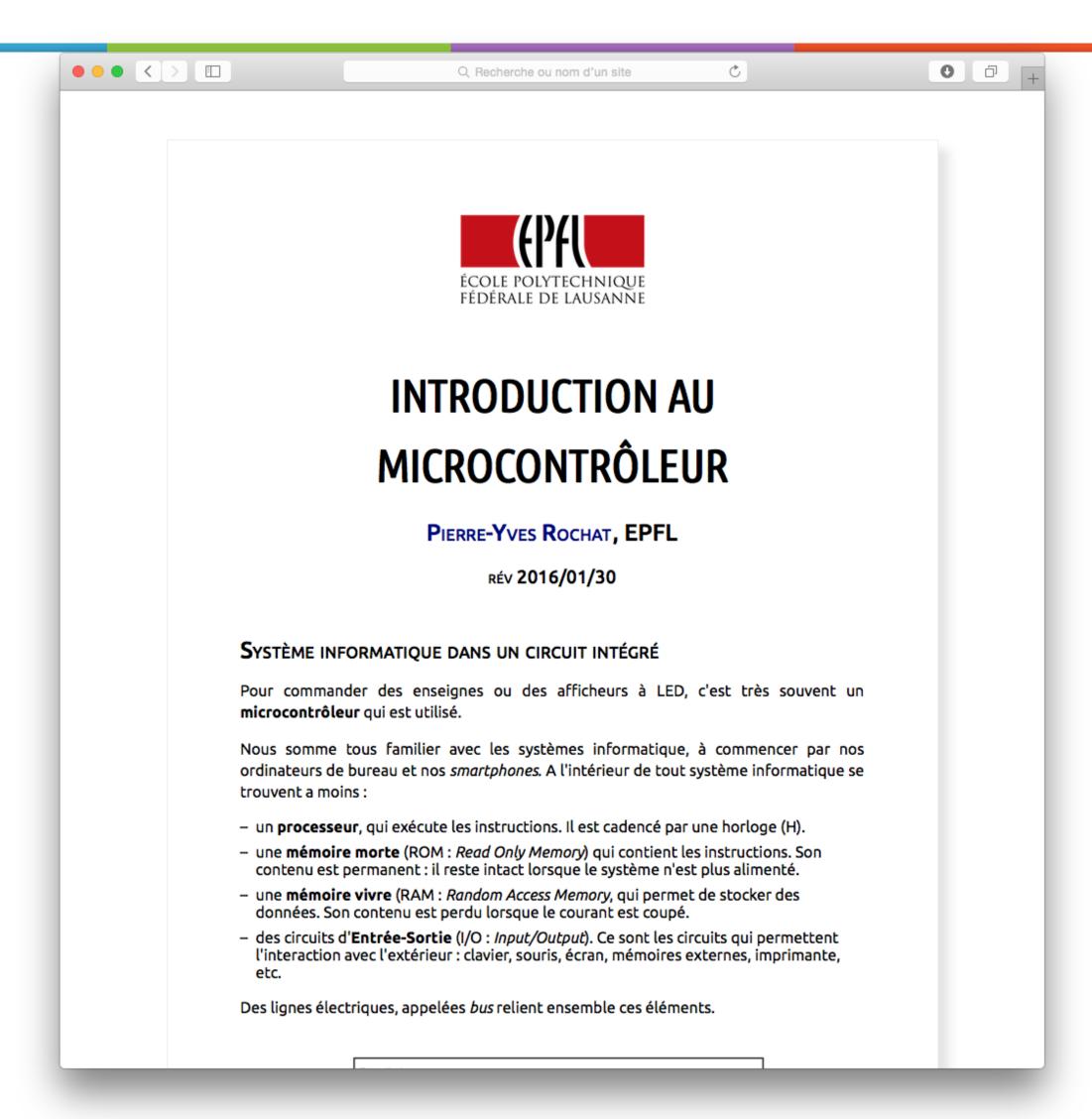
Exemple de CSS



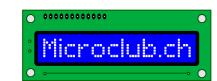
```
exemple.css
     @charset "utf-8";
                                                                            1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w</pre>
     @import url(http://fonts.googleapis.com/css?family=Ubuntu+Mono:400,700,400
                                                                            2 <html xmlns="http://www.w3.org/1999/xhtml">
  3
     body {
                                                                            3
                                                                                <head>
         font-family: Ubuntu;
                                                                                 <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
  4
                                                                                 <meta http-equiv="Content-Style-Type" content="text/css" />
         width: 17cm;
                                                                                 <meta name="generator" content="pandoc" />
         font-size: 12pt;
  6
                                                                                 <meta name="author" content="Pierre-Yves Rochat, EPFL" />
  7
         string-set: strMOOC
                              "EPFL — MOOC AFFICHAGES À LED"
                                                                            7
                    strPage_nb "- " counter( page ) " -"
  8
                                                                                 <title>Introduction au microcontrôleur</title>
                    strSEPARATOR " •";
                                                                                 <style type="text/css">code{white-space: pre;}</style>
  9
                                                                                 <link rel="stylesheet" href="../../statiques/style.css" type="text/css" />
 10
                                                                           10
                                                                                </head>
     h1, h2, h3, h4, h5, h6 {
 11
                                                                           11
         page-break-after: avoid;
                                                                                <body>
 12
                                                                            12
                                                                                <div id="header">
         font-weight: 500;
 13
                                                                                <h1 class="title">Introduction au microcontrôleur</h1>
 14
                                                                                <h2 class="author"><a href="&#x6d;&#x61;&#x69;&#108;&#116;&#x6f;&#58;&#112;&
     div#header {
 15
                                                                                <h3 class="date">rév 2016/01/30</h3>
         text-align: center;
 16
                                                                               </div>
 17
                                                                           17
                                                                                <h2 id="système-informatique-dans-un-circuit-intégré">Système informatique
     h2.author {
 18
         page-break-inside: avoid;
                                                                                Pour commander des enseignes ou des afficheurs à LED, c'est très souvent
 19
                                                                                Nous somme tous familier avec les systèmes informatique, à commencer par
  20
     ul li {
                                                                                <l
  21
                                                                            21
         line-height: 1.2em;
                                                                                un <strong>processeur</strong>, qui exécute les instructions. Il est cad
 22
                                                                                23 }
     @media screen { body { ... } }
                                                                                @media print { body { ... } }
                                                                                des circuits d'<strong>Entrée-Sortie</strong> (I/O : <em>Input/Output</e>
  25
 26
                                                                            26 
☐ Line 36, Column 17
                                                                                                                          UTF-8
                                                                                                                                   Tab Size: 4
                                                                                                                                               LESS
```

Exemple de conversion avec Pandoc





Conversion HTML – PDF avec Prince princexml.com



Avantages

- Résultats magnifiques
- Supporte de nombreuses possibilités des CSS
- Peut exécuter des JavaScript lors de la génération
- Très rapide

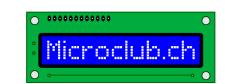
Désavantages

- Coûteux (495\$)
- Pas utilisable tel quel par les non-techniciens car uniquement en ligne de commande

Utilisation

```
prince \
  input.html \
  -o output.pdf
```

Conversion HTML – PDF avec WeasyPrint weasyprint.org



Avantages

Gratuit (licence BSD)

Désavantages

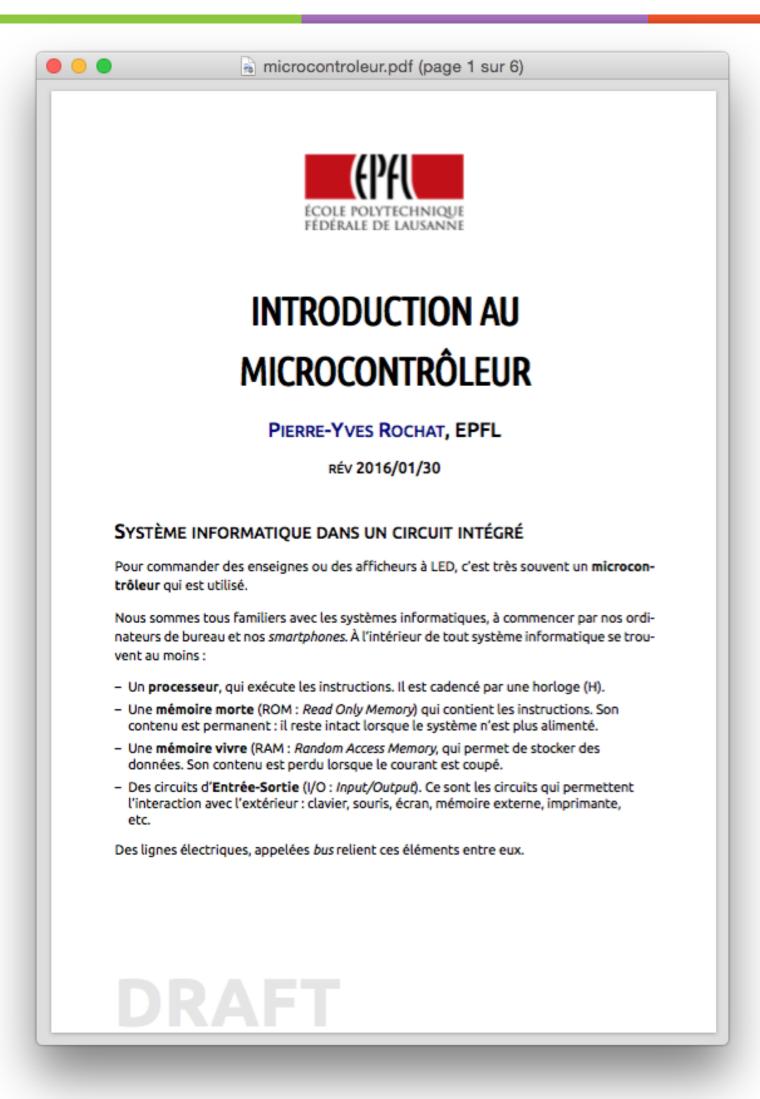
- Résultats parfois imprévisibles
- Ne supporte pas autant de possibilités CSS que Prince
- Pas de JavaScript lors de la génération
- Très lent
- Pas utilisable tel quel par les non-techniciens car uniquement en ligne de commande
- Potentiellement infernal à installer

Utilisation

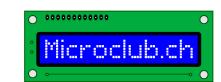
```
weasyprint \
  input.html \
  output.pdf
```

Exemple de conversion avec Prince



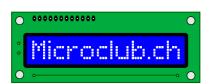


Ressources HTML



- W3School
 - http://www.w3schools.com/html/
- Mozilla
 - https://developer.mozilla.org/fr/docs/Web/HTML
- W3C
 - Doctype XHTML1 : https://www.w3.org/TR/xhtml1/
 - Doctype HTML5 : https://www.w3.org/TR/html5/
 - Validation : https://validator.w3.org/

Outil "pre-dia"

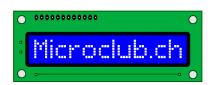


Génération de dias pour les vidéos du MOOC LED

- Mode "Normal" ⇒ avec séquences
- Mode "Compact" ⇒ sans séquences (applaties)

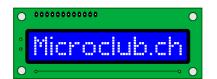
```
<section>
<!-- A -->
# Outil "pre-dia"{ class="en_tete" }
<div style="font-size:48pt; left:2cm; width:28.0cm; top:6.5cm;">
## Génération de dias pour les vidéos du MOOC LED
<!-- 23 -->- Mode "Normal" ⇒ avec séquences
<!-- 3 -->- Mode "Compact" ⇒ sans séquences (applaties)
</div>
<div style="font-size:48pt; left:31cm; width:30.0cm; top:6.5cm;">
<!-- 23 -->![](./images/exemple-pre-dia-1.png){ height=20cm }
</div>
</section>
```

Git



- SCM (Source Code Management) ⇒ Indispensable pour tout projet informatique
 - Permet de gérer l'historique
 - Permet de gérer la collaboration avec les participants
- Exemples
 - Git
 - Perforce
 - Mercurial
 - Bazaar
 - CVS
 - Subversion
 - Liste de logiciels de SCM

Git – avantages et inconvénients



Avantages

- Communauté très large ⇒ donc expérience réutilisable dans d'autres contextes
 - + facilité de trouver des réponses
- Open Source
- Très rapide
- Facile à installer
- Disponible pour tous les OS courants
- Disponibilité de services d'hébergement gratuits comme GitHub, Bitbucket ou git.epfl.ch

Git – avantages et inconvénients



Inconvénients

- L'interface utilisateur est inutilement complexe. Aucune commande ne peut être déduite par la logique. Les noms des commandes sont troublants (par ex. git blame pour voir qui a fait quoi sur un fichier.)
- La gestion des fichiers volumineux est une option (GLFS Git Large File Storage)
- Pas de diff de fichiers binaires
- Il n'existe que des GUI qui n'aiderons pas les gens allergiques à la ligne de commande ⇒ difficile à utiliser par des non-techniciens

GitHub – avantages et inconvénients



Avantages des dépôts et des sites web

- Gratuit
- Pas de limite sur le nombre de dépôts
- Possibilité de site web (autant qu'on veut à priori)
- Bande passante OK
- Diff visuelles
- Projet public = publicité gratuite
- Facilité de clonage

GitHub – avantages et inconvénients



Inconvénients des dépôts

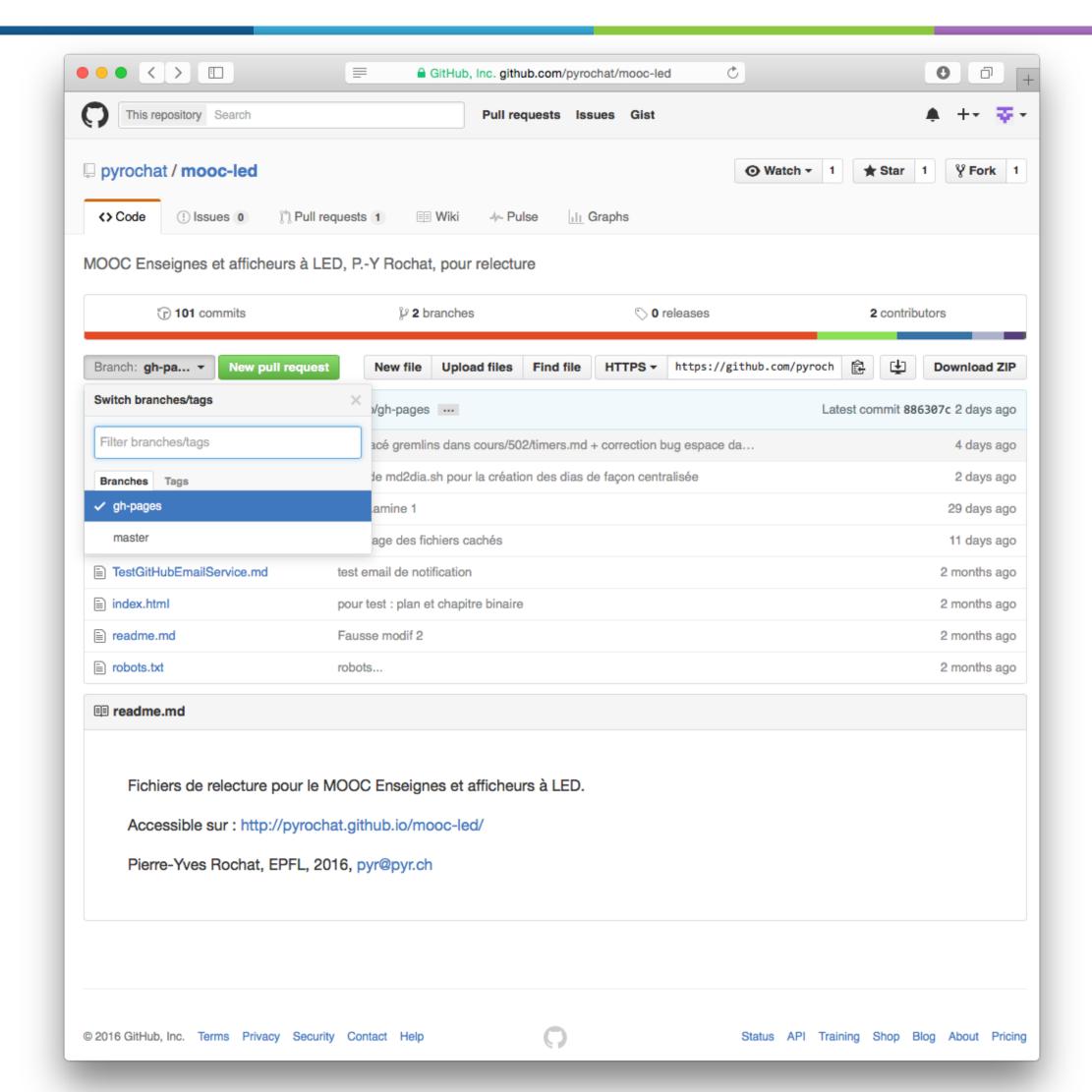
- Fichiers sensibles qui doivent rester dans l'entreprise ou sur le territoire
- Limite de 1 GB
- Sinon même limitations que Git

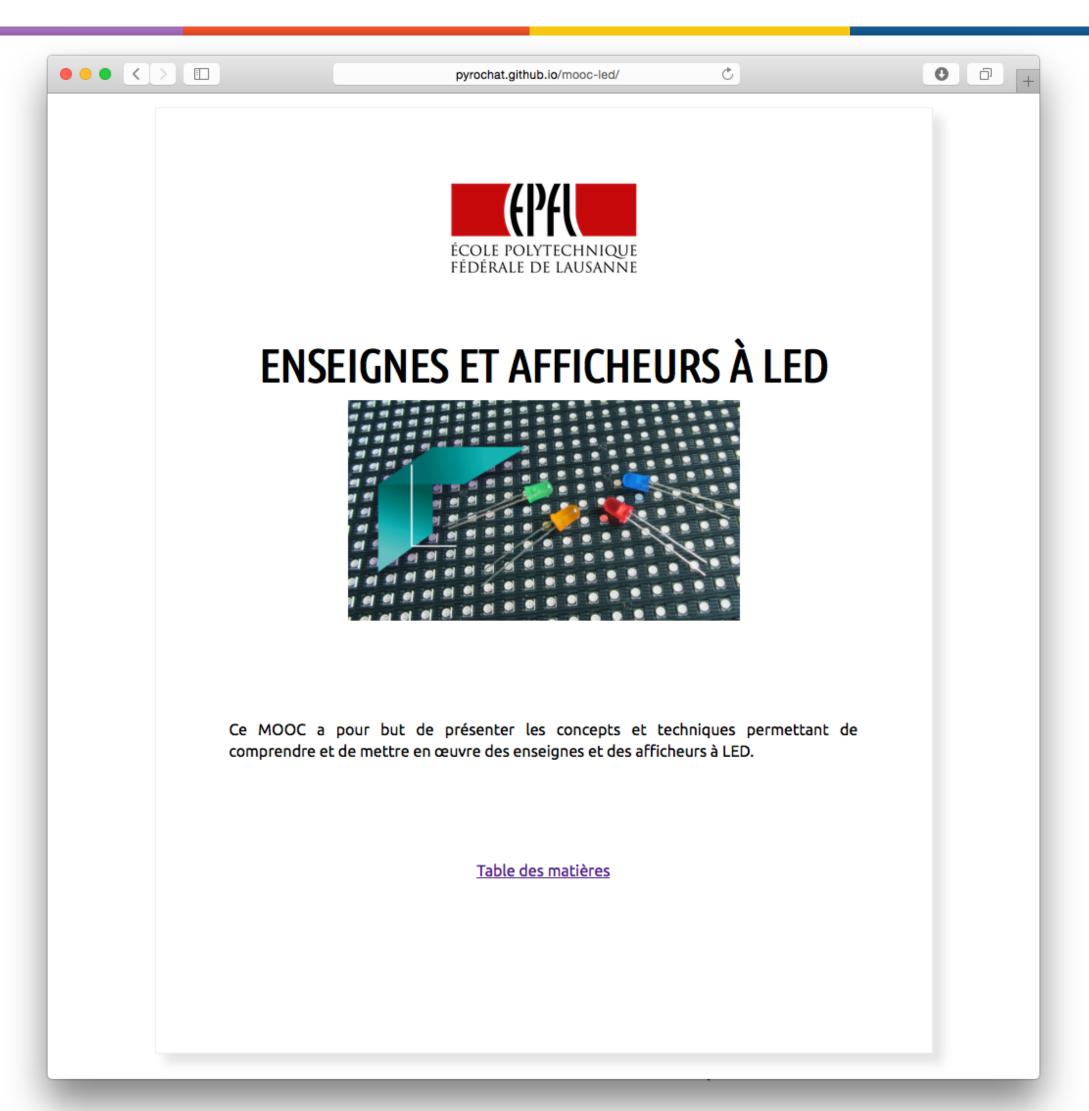
Inconvénients des sites web

- Uniquement statiques (pas de scripts côté serveur)
- Pas de possibilité d'utiliser des branches différentes (gh-pages ou master uniquement)
- Gestion du <u>CNAME</u> difficile quand on "forke"

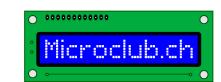
GitHub.com et GitHub.io







GitHub – flux de travail (avec site web sur GitHub)



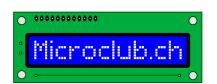
Forking Workflow

- Le responsable de projet
 - crée un dépot sur GitHub
 - le clone sur un ordi
 - crée la branche "gh-pages"
 - défini cette branche comme défaut
 - "add+commit+push" les fichiers
 - accepte ou refuse les "pull request"

Les collaborateurs

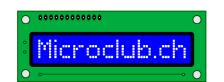
- clonent le dépôt du responsable (optionnel)
- "fork" le dépôt et clonent le "fork" sur un ordi
- modifient ou ajoutent des fichiers
- "add+commit+push" les fichiers vers leur "fork"
- envoient des "pull request" au reponsable pour qu'il intègre les modifications
- synchronisent leurs "fork"

Exemples de sites web sur GitHub



- http://pyrochat.github.io/mooc-led/
 - https://github.com/pyrochat/mooc-led
- http://ouilogique.com/
 - https://github.com/NicHub/ouilogique.com
- http://cnl-naturopathie.ch/
 - https://github.com/cnl-naturopathie/site-web-v4
- http://getbootstrap.com/
 - https://github.com/twbs/bootstrap ufgsgfui

Quelques outils pratiques



- Sublime Text <u>sublimetext.com</u>
- Atom <u>atom.io</u>
- Dterm <u>decimus.net/DTerm</u>
- Antidote <u>druide.com</u>
- Le guide du typographe romand typoguide.ch
- Validation HTML <u>validator.w3.org</u>
- Image Optim imageoptim.com
- Marked 2 <u>marked2app.com</u>
- Ukelele <u>scripts.sil.org</u>