

BioDataScience Administration

Philippe Grosjean & Guyliann Engels

2020-05-29

Contents

Préambule	5
1 Introduction	7
2 Organisation générale	9
2.1 RGD, droit d’auteur & plagiat	12
2.2 Infos utiles concernant le DSI UMONS	13
3 Git & Github	15
3.1 Installation Git et cie	15
3.2 BioDataScience-Course	15
3.3 Github Classroom	16
3.4 Astuces Git et Github	16
4 Wordpress	19
4.1 Authentification des étudiants	20
4.2 Installation Wordpress	23
4.3 Configuration serveur et git	25
4.4 Agencement du site	28
4.5 Variables contextuelles	30
5 H5P	33
5.1 Installation de H5P	33
5.2 Enregistrement H5P xAPI	36
5.3 Gestion des contenus H5P	41
5.4 Quelques contenus H5P utiles	44
6 Bookdown	47
6.1 Intégration dans Wordpress	47
6.2 Contenu H5P dans bookdown	47
7 Shiny	51
8 Moodle	53
8.1 Gestion des cours	53
8.2 Badges	53
8.3 Notes	54
9 Vidéos	55
10 Rapport de progression	57

A Installation	59
Références	61

Préambule

Notre organisation **BioDataScience-Course** sur Github devient de plus en plus complexe avec le temps car nous combinons plusieurs outils différents pour fournir à nos étudiants un environnement de travail et du matériel pédagogique riches et variés. Ce bookdown centralise l'information nécessaire pour les enseignants et les administrateurs de ces outils.



Le matériel dans cet ouvrage est distribué sous licence CC BY-NC-SA 4.0.

Détails concernant le système utilisé pour compiler ce bookdown...

```
sessioninfo::session_info()
```

0.0.0.0.1 Information système

```

# - Session info -----
# setting value
# version R version 3.6.3 (2020-02-29)
# os      macOS Catalina 10.15.4
# system  x86_64, darwin15.6.0
# ui      X11
# language (EN)
# collate en_US.UTF-8
# ctype   en_US.UTF-8
# tz      Europe/Brussels
# date    2020-05-29
#
# - Packages -----
# package * version date lib source
# assertthat 0.2.1 2019-03-21 [1] CRAN (R 3.6.0)
# bookdown 0.18 2020-03-05 [1] CRAN (R 3.6.0)
# cli 2.0.2 2020-02-28 [1] CRAN (R 3.6.0)
# crayon 1.3.4.9000 2020-04-29 [1] Github (r-lib/crayon@dcf6d44)
# digest 0.6.25 2020-02-23 [1] CRAN (R 3.6.0)
# evaluate 0.14 2019-05-28 [1] CRAN (R 3.6.0)
# fansi 0.4.1 2020-01-08 [1] CRAN (R 3.6.0)
# glue 1.4.0 2020-04-03 [1] CRAN (R 3.6.2)
# htmltools 0.4.0.9003 2020-04-29 [1] Github (rstudio/htmltools@984b39c)
# knitr 1.28 2020-02-06 [1] CRAN (R 3.6.0)
# magrittr 1.5 2014-11-22 [1] CRAN (R 3.6.0)
# rlang 0.4.5 2020-03-01 [1] CRAN (R 3.6.0)
# rmarkdown 2.1 2020-01-20 [1] CRAN (R 3.6.0)
# sessioninfo 1.1.1 2018-11-05 [1] CRAN (R 3.6.0)
# stringi 1.4.6 2020-02-17 [1] CRAN (R 3.6.0)
# stringr 1.4.0 2019-02-10 [1] CRAN (R 3.6.0)
# withr 2.2.0 2020-04-20 [1] CRAN (R 3.6.2)
# xfun 0.13 2020-04-13 [1] CRAN (R 3.6.2)
# yaml 2.2.1 2020-02-01 [1] CRAN (R 3.6.0)
#
# [1] /Library/Frameworks/R.framework/Versions/3.6/Resources/library

```

Chapter 1

Introduction

Chaque section couvre un outil en particuliers, avec la section 2 qui présente la façon dont les différents outils sont assemblés pour réaliser notre environnement de travail pour les cours de **Science des Données Biologiques** à l'UMONS.

Chapter 2

Organisation générale

Les divers outils logiciels que nous utilisons pour créer un environnement d'apprentissage riche et varié pour nos étudiants se basent pour la plupart sur des solutions Open Source et sur différents serveurs.

- R est le logiciel principal de statistique, visualisation et manipulation des données que nous enseignons à nos étudiants en même temps que les concepts de statistique et science des données eux-mêmes. C'est un logiciel puissant et gratuit que l'étudiant pourra s'appropriier plus tard pour faire ses analyses tout au long de sa carrière s'il le souhaite. R est préféré à Python ici car il est plus orienté statistiques et moins programmation, ce qui convient mieux à nos étudiants biologistes.
- RStudio en version serveur pour nos étudiants, ou en version serveur ou desktop pour les enseignants, couplé à R Markdown permet d'éditer du code R et des rapports ou autres documents au format R Markdown.
- Git est un système de gestion de version largement utilisé en programmation et de plus en plus utilisé en science des données. Combiné à R, RStudio et R Markdown, il offre la possibilité de travailler de manière *reproductible*. Nous enseignons git en même temps que les outils plus orientés science des données à proprement parler avec l'objectif d'amener nos étudiants à prendre des bonnes habitudes en matière de traçabilité et reproductibilité de leurs analyses. Nous avons choisi Github comme système d'hébergement des dépôts git sur le Net, parce qu'il est gratuit, puissant, universellement utilisé et qu'il est couplé à RStudio. L'interface git et Github de RStudio rend son utilisation simple, et Github Desktop complète l'ensemble pour effectuer des tâches non reprises dans l'interface git de RStudio. Enfin, Github Classroom est une extension permettant de gérer les nombreux dépôts créés par nos étudiants dans le cadre de nos cours.
- La SciViews Box est une machine virtuelle sous VirtualBox totalement préconfigurée. Nous créons une SciViews Box chaque année avec tous les logiciels nécessaires aux étudiants mais pas seulement : il s'agit aussi d'une machine virtuelle permettant de réaliser de nombreuses analyses de façon parfaitement reproductible. Faire installer tous les logiciels nécessaires par les étudiants n'est pas possible. Plusieurs solutions s'offrent à nous pour proposer un environnement logiciel préconfiguré pour les cours :
 - Utiliser RStudio cloud serait la solution la plus simple au départ car l'étudiant n'a qu'à se connecter et c'est parti. Même avec un Chromebook, il peut travailler sur sa session RStudio Cloud. Cependant, la session n'est disponible

que pour le cours et l'étudiant doit tout installer sur son ordinateur ensuite s'il veut continuer à utiliser R, RStudio et Git/Github (par exemple, dans le cadre de son mémoire). Or précisément, il n'a jamais appris à installer ces outils. Cela nous paraît être une barrière importante à l'adoption de ces outils logiciels au-delà du cours, même si beaucoup de (grandes) universités ont choisi cette voie.

- Utiliser Docker et Rocker qui est une solution de containerisation pour R et RStudio. Si ces outils fonctionnent bien en production, nous les avons trouvés moins pratiques que la machine virtuelle pour nos étudiants, surtout dans un contexte où la quasi-totalité des ordinateurs utilisés sont sous Windows ou MacOS (Docker fonctionne particulièrement bien dans un environnement Linux au départ). L'installation et la gestion de containers Dockers est simple pour un utilisateur un peu technique et averti, mais nous, nous démarrons avec certains étudiants qui ont une connaissance bien moins technique de l'informatique. La moindre action d'installation ou de configuration, aussi simple soit-elle, paraît démesurément complexe à leurs yeux. Il faut donc démarrer avec de l'**ultra-simple**, autrement dit, fonctionnel directement ou s'auto-installant, auto-configurant tout seul.
 - Déléguer le calcul sur le cloud à travers Binder. Aujourd'hui, il est très facile d'exécuter du code dans un document Jupyter ou R Markdown via ce service gratuit sur le cloud. Et c'est effectivement d'un accès ultra-simple du côté de l'utilisateur qui n'a qu'à cliquer sur un lien pour lancer l'application. Au moins un format de cours en ligne y recourt (template par Ines Montani, voir par exemple pour un cours concret en R, Supervised Machine Learning Case Studies in R). Nous l'avons testé, mais l'avons jugé à l'usage trop lent au démarrage et trop limité ensuite. De plus, comme pour RStudio Cloud, les étudiants n'ont aucun logiciel utilisable sur leur PC pour travailler ensuite. La simplicité au départ ne fait donc que repousser la difficulté ultérieurement de devoir quand même installer R, RStudio, git, etc. sur son PC.
 - Enfin, la solution de la machine virtuelle avec la SciViews Box nous permet d'offrir une série de logiciels préconfigurés en version serveur et/ou locale avec interface graphique sous Xubuntu. A l'usage, cela fonctionne bien, mais l'installation était encore un peu laborieuse lorsqu'elle devait se faire à la main dans VirtualBox (on en revient pratiquement à des difficultés similaires à Docker, par exemple). Nous avons donc conçu des logiciels annexes d'installation et de lancement de la machine virtuelle de sorte que celle-ci ne soit qu'à un clic de souris (ou presque) de nos étudiants débutants ! A l'usage, cela fonctionne très bien. Nous avons un lanceur rapide par édition de SciViews Box, et ce, à la fois sous Windows et sous MacOS. L'étudiant a un logiciel réellement installé sur son PC, qu'il peut continuer à utiliser parfaitement au-delà des cours eux-mêmes.
- Moodle est la plateforme d'apprentissage en ligne utilisée par notre université, l'UMONS. Si elle offre de nombreuses possibilités, il n'est pas possible d'y intégrer les outils d'apprentissage liés à R comme les learnrs et les ouvrages en ligne au format bookdown, voire encore des documents R Markdown. Nous essayons d'intégrer autant que possible nos outils dans cette plateforme, mais nous sommes obligés de recourir à d'autres solutions pour pouvoir couvrir le panel des fonctionnalités qui nous sont nécessaires.
 - Bookdown est un type de sortie R Markdown multi-documents qui permet d'intégrer également des calculs et graphiques de R. Nous l'utilisons comme moyen de centraliser tout notre matériel pédagogique et l'organiser en modules qui sont autant

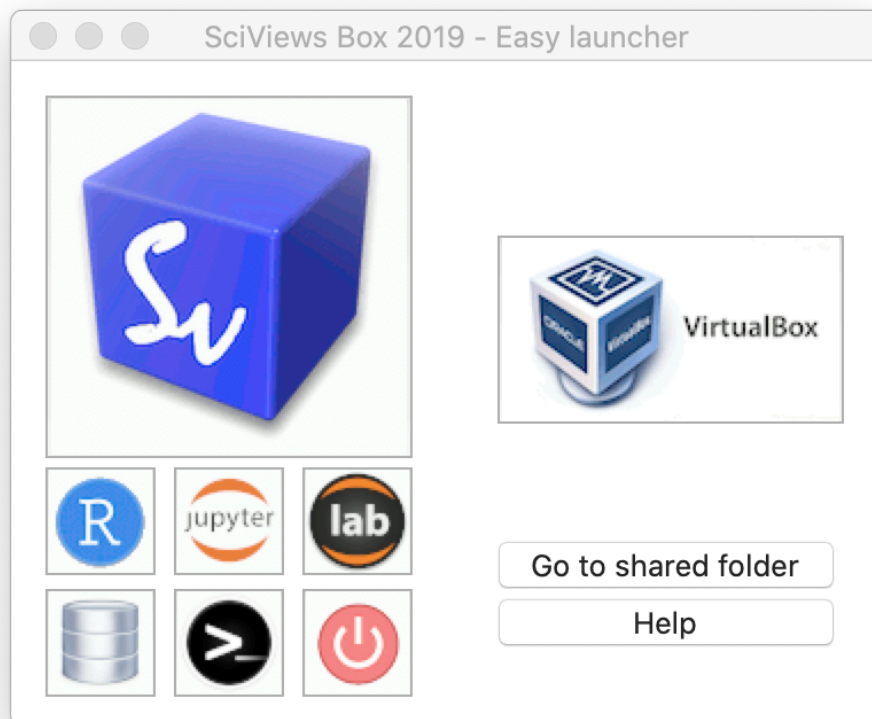


Figure 2.1: Lanceur rapide de la SciViews Box 2019

de chapitres différents du bookdown. Nous avons également développé différentes extensions pour bookdown afin d’y intégrer tout ce dont nous avons besoin (voir 6), et notamment des applications Shiny (voir 7) et des widgets h5p (voir 5). Un chapitre entier y est donc consacré.

- Shiny permet de créer des applications sous R avec interface graphique sur un browser web. Les possibilités sont pratiquement sans limites pour créer des applications interactives et démontrer les concepts “en live” à nos étudiants. Les applications Shiny sont intégrables dans bookdown, mais elles doivent être exécutées sur un serveur dédié. Un type particulier d’application Shiny est constitué par les learnrs. Ces documents interactifs présentent des exercices dans R et des choix multiples aux étudiants. Ces learnrs forment la base des outils d’apprentissage à domicile et aussi des examens théoriques de nos cours.
- H5P propose une large variété de contenus javascript intégrables dans une page web sous forme d’iframes pour varier et rendre plus efficace l’apprentissage. Il s’agit d’un logiciel gratuit open source, mais il faut installer son propre serveur sous Moodle, Wordpress ou Drupal. Des essais ont été fait sous Moodle de l’UMONS. Cela fonctionne, mais il apparaît que les contenus H5P de Moodle ne sont *pas* intégrables dans un bookdown, en tous cas pas la version de test telle qu’elle était configurée. Nous avons donc décidé d’installer notre propre serveur Wordpress, essentiellement pour pouvoir faire fonctionner nos contenus H5P. Ce serveur est hébergé chez Gandi.net qui gère aussi le nom de domaine **sciviews.org** sous lequel nos créations sont servies sur le net, ainsi que les boîtes mails correspondantes, comme **sdd@sciviews.org**.
- Les vidéos relatives au cours sont hébergées sur Youtube dans un canal dédié **BioDataScience-Course**. La section relative aux vidéos (voir 9) explique aussi en détails comment nous créons ces vidéos et avec quels outils.
- Enfin, les données relatives à la progression des étudiants dans les learnrs sont collectées dans une base de données MongoDB hébergée sur le cloud MongoDB Atlas. La progression peut être suivie grâce aux rapports de progression (10).



Ces nombreux outils et services devant collaborer entre eux pour fournir un environnement d’apprentissage complet et cohérent, il est important de les configurer et de les inter-connecter correctement. C’est le but principal du présent document que de documenter ces opérations d’administration de l’écosystème logiciel utilisé.

Les sections suivantes détaillent l’installation et la configuration de différents outils logiciels tandis que les paragraphes suivants dans la présente section reprennent quelques informations générales utiles.

2.1 RGPD, droit d’auteur & plagiat

Les questions relatives aux données d’un point de vue plus légal sont traitées ici.

2.1.1 RGPD

Le **Règlement Général sur la Protection des Données** ou RGPD (n° 2016/679) adopté par l’Union Européenne en 2016 régit la façon dont les données à caractère personnel peuvent être collectées et utilisées. Ce règlement est en faveur de l’utilisateur (ici

l'étudiant). Il faut notamment son accord pour collecter et utiliser ses données personnelles. Or de telles données sont indispensables pour suivre la progression des étudiants, pouvoir les noter, etc.

A l'UMONS, l'étudiant signe le document adéquat lors de son inscription. Ce document mets **Moodle** en conformité par rapport au RGPD, mais pas les outils externes. Nous devons donc préciser exactement ce que nous devons faire pour être en conformité. C'est un travail que nous devons encore réaliser au stade actuel avec la juriste de l'UMONS une fois tous nos outils mis en place.

L'utilisateur doit avoir la possibilité d'effacer intégralement ses données personnelles s'il le souhaite lorsqu'il efface son compte d'un site. C'est indiqué explicitement dans le RGPD. Cependant, des restrictions à ceci sont indispensables pour la bonne gestion du suivi des étudiants à l'Université. C'est pour cela que l'étudiant a du signer un accord lors de son inscription. Afin d'être en conformité avec cette directive, nous avons rajouté un bouton **Effacer mes données personnelles** dans la première page des cours bookdown. S'il clique sur ce bouton, l'utilisateur peut ensuite lire le contenu de manière anonyme et les activités learnr/Shiny ne sont **pas** enregistrées. Par contre, H5P est toujours enregistré et les données historiques antérieures sont toujours dans notre base de données.

TODO: il faudra peut-être mettre en place un bouton qui permette de demander l'effacement de *toutes* les données, et qui pourrait être utilisé pour les visiteurs non-UMONS pour être totalement en conformité RGPD de ce point de vue.

L'anonymisation des données est de mise lorsque ces données servent à une étude générale (par exemple, évolution des performances des cohortes d'étudiants avec les outils progressivement mis en place, études scientifiques, ...). Cela passe par l'effacement des données de toute information à caractère personnel. Le nom, numéro de matricule ou adresse email de l'étudiant sont remplacés par un identifiant générique, par exemple, "étudiant #1", "étudiant #2", etc. Nous mettons cette pratique en œuvre pour toutes les études rétrospectives générales visant à estimer l'impact pédagogique des outils mis en place.

2.1.2 Droit d'auteur

...

2.1.3 Plagiat

L'UMONS s'est doté d'un outil de détection du plagiat intégré à Moodle. Il reconnaît les fichiers Word, PowerPoint et PDF. Malheureusement, il ne reconnaît pas les PDF générés via LaTeX. Donc, il faut sortir soit en Word, soit faire un PDF depuis Word à partir de documents R Markdown. Il faut intégrer les documents sous forme de devoirs dans Moodle pour pouvoir bénéficier de cette fonction, ce qui est très contraignants pour nous.

TODO: réfléchir à des façons plus pratiques de mettre en œuvre l'anti-plagiat dans nos cours de Science des données.

2.2 Infos utiles concernant le DSI UMONS

- Les documents **Comment ?** expliquent comment réaliser différentes tâches liées à l'utilisation des outils informatiques à l'UMONS (connexion, gestion des mails,

partage de fichiers, etc.). Ces documents sont disponibles depuis l’Intranet UMONS uniquement.

Chapter 3

Git & Github

Nous utilisons abondamment et enseignons l’usage de git et Github dans le cadre de nos cours de **Science des Données Biologiques** à l’UMONS. De plus, nous utilisons Github Classroom pour gérer la distribution des travaux aux étudiants sous forme de dépôts Github. Cette section couvre ces logiciels.

3.1 Installation Git et cie

Dans la SciViews Box, git, Github Desktop et RStudio server (incluant une interface vers git) sont préinstallés. Hors machine virtuelle, il faut installer git, RStudio et Github Desktop. Le reste se fait en interagissant avec Github via le navigateur web.

3.2 BioDataScience-Course

L’organisation Github BioDataScience-Course regroupe tous les dépôts relatifs à nos cours, y compris les travaux des étudiants.

3.2.1 Convention de noms

Comme l’organisation contient plusieurs centaines de dépôts, il est utile de les nommer correctement pour s’y retrouver. Les dépôts créés via Github Classroom seront traités plus loin. Voici quelques règles :

- Les dépôts important (matériel pédagogique, gestion des étudiants et de leur progression, ...) commencent tous par **sdd**. Pour les dépôts moins importants ou de test, utiliser un autre préfixe.
- Lorsqu’un dépôt est relatif à une cohorte d’étudiants en particulier, nous l’indiquons par un suffixe reprenant les deux derniers chiffres de l’année initiale, suivi de **m** pour les étudiants montois et **c** pour les étudiants de Charleroi. Par exemple, **sdd_students19m** concerne la cohorte d’étudiants montois inscrits en premier cours (Bab2 de biologie) en 2019-2020. De même **sdd_students20c** concerne la cohorte d’étudiants de Charleroi inscrits en Bab2 en 2020-2021.

3.3 Github Classroom

Nous utilisons Github Classroom pour générer la distribution des travaux aux étudiants sous forme de dépôts Github.

TODO: Guyliann: les éléments importants de la gestion de Github Classroom sont à présenter ici...

TODO: intégration de Github Classroom avec Moodle, voir <https://classroom.github.com/help/setup-moodle> & <https://classroom.github.com/help/import-roster-from-lms>.



Une autre ressource intéressante dans le contexte de l'utilisation de Github/Github Classroom pour l'enseignement est Github Education.

3.4 Astuces Git et Github

Quelques astuces qui nous sont utiles...

- Comparaison de branches sur Github : ajouter `/compare` à la fin de l'URL du dépôt et une page de comparaison de branche apparaît. Sinon, cette comparaison est aussi possible dans Github Desktop (menu **Branch** -> **Compare to Branch**).
- Lorsqu'une branche n'est plus utile, elle peut être facilement éliminée depuis Github Desktop (entrée de menu **Branch** -> **Delete...**). Cette branche doit être activée d'abord, et tous les commits doivent avoir été pushés. Il est conseillé d'éliminer la branche sur le serveur en même temps que la branche locale.

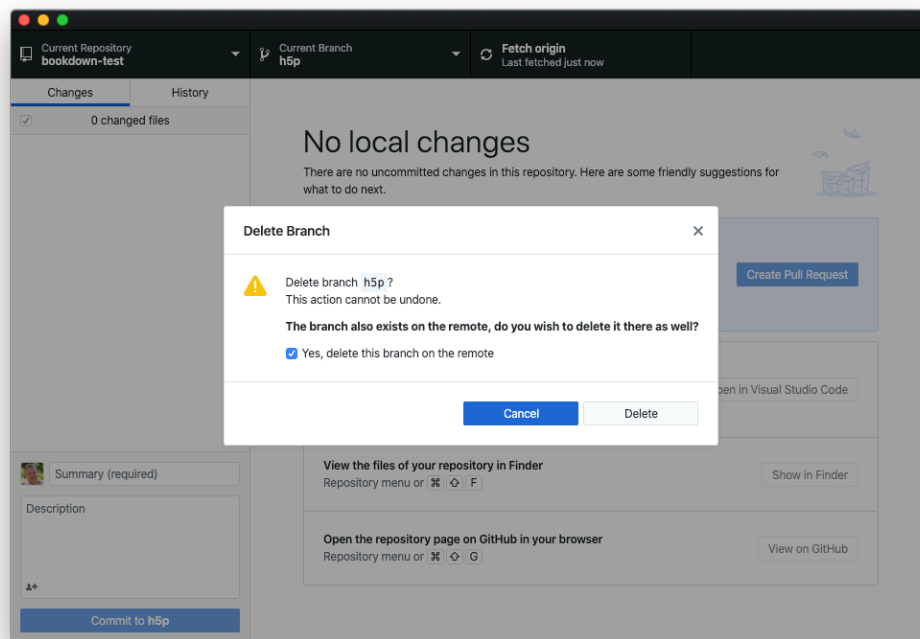


Figure 3.1: La boîte de dialogue d’élimination d’une branche dans Github Desktop.

Chapter 4

Wordpress

Nous utilisons Wordpress principalement pour gérer nos contenus **H5P**. Cependant, étant basé sur un serveur LAMP (Linux-Apache-MySQL-PHP), il permet également d'implémenter des fonctions basées sur des scripts côté serveur qui seraient difficiles, voire impossibles avec un site statique tel que l'offre bookdown, par exemple. Ainsi, nous pouvons aussi y intégrer un enregistrement de l'utilisateur via Github (protocole OAuth2). Ceci est bien pratique pour faire le lien entre le compte UMONS/Moodle et le compte Github de l'étudiant.



Nous avons choisi de privilégier l'authentification Github à celle de l'UMONS pour les raisons suivantes :

- Certains étudiants arrivent au cours sans être encore complètement inscrits, avec leur propre PC portable. Ils n'ont pas encore de login UMONS et ne pourraient rien faire si on dépendait exclusivement de ce login pour travailler. Avec Github, ils créent leur compte librement et sont immédiatement opérationnels.
- Nos outils sont prêts et utilisables en dehors du contexte de l'UMONS également (si nécessaire, par exemple, pour une formation dans le cadre d'un workshop international).
- Comme Github et Github Classroom occupent une place importante dans les outils d'apprentissage de nos cours, les étudiants doivent de toute façon créer et utiliser un compte Github. L'authentification via Github permet d'être certain que ce compte est créé et fonctionnel et que l'étudiant fournit effectivement son propre login.

Toutes ces raisons sont suffisantes pour que l'authentification Github soit considérées comme l'élément clé dans le contexte de nos cours de Science des Données.

En 2018-2019 et 2019-2020, nous avons travaillé uniquement avec un serveur statique bookdown hébergé sur Github. Si le résultat est correct, nous avons noté plusieurs problèmes :

- Pas d'enregistrement des étudiants (suivi de progression impossible),
- Pas de H5P. Nous pourrions utiliser <https://h5p.org>, mais cette plateforme n'est prévue que pour des tests, et ne permet pas de récupérer les activités des étudiants,
- Pas de service d'applis Shiny et de learnrs. Nous exécutons ces applis en local dans la SciViews Box via les packages R BioDataScience et BioDataScience2. Cela permet

d'utiliser les learnrs, mais sans être vraiment très pratique pour les étudiants. Les learnrs qui y sont intégrés se connectent à la base de données MongoDB via une petite appli Shiny intégrée qui demande de rentrer manuellement dans des zones de texte son login Github et son email pour s'identifier. Toutefois, comme aucun contrôle n'est fait sur ce que les étudiants entrent dans ces champs, nous avons eu beaucoup d'enregistrements fantaisistes qu'il a fallu régler grâce à une table de corrections appliquée aux données de la base.

Au final, même si l'approche bookdown statique sur Github + learnr/Shiny en local dans la SciViews Box permet d'obtenir le minimum de fonctions requises pour le suivi de la progression d'apprentissage des étudiants, ce n'est pas pratique à l'usage et nous avons dû dépanner souvent des étudiants... sans compter leur stress de savoir si leur activité est bien enregistrée dans la base de données ou non (dans nos cours à l'UMONS, deux points sur vingt dans la note totale sont réservés à l'exécution de ces learnrs, d'où l'importance de bien enregistrer qui fait quoi).

4.1 Authentification des étudiants

En 2020-2021, nous avons décidé de mettre en œuvre un serveur Wordpress pour fournir le contenu H5P qui nous manque et dont l'utilisation est enregistrée pour chaque étudiant afin de pouvoir suivre sa progression (protocole xAPI). Ce serveur Wordpress nous donne aussi l'occasion de mettre en œuvre une authentification des étudiants plus fiable. Notre objectif est double au niveau de l'authentification de nos étudiants. D'une part, nous voulons obtenir de manière fiable à la fois leurs données UMONS (via Moodle) et leurs données Github, ce qui implique que ces données doivent être soit fournies automatiquement, soit vérifiées durant l'authentification. D'autre part, nous voulons pouvoir exécuter des applications Shiny en mode sécurisé pour, par exemple, leur offrir un rapport de progression calculé en direct. La nouvelle authentification se fait comme suit :

1. Nous intégrons nos outils dans Moodle via une URL qui transmet des informations relatives à l'utilisateur authentifié UMONS ("query string", c'est-à-dire la partie "?..." de l'URL). On obtient de cette façon, entre autres, son numéro de matricule, son adresse email UMONS, son nom et son prénom, le cours qu'il suit, ... Les détails sont dans la section correspondante (voir 8) et ne seront donc pas discutés ici.
2. Par ailleurs, nous utilisons abondamment Github pour les exercices des étudiants. Leur login y est différent, et il est aussi possible qu'ils y aient indiqué une adresse email différente (même si nous leurs demandons de mettre la même, ce qui est d'ailleurs indispensable pour l'identification correcte de l'auteur des commits). Dans Wordpress, nous pouvons intégrer une authentification sur base de leur compte Github via le protocole OAuth. C'est doublement intéressant car nous obtenons ainsi leur login Github à tous les coups, et aussi, nous sommes certains que leur compte Github est créé et actif dès le moment où ils sont enregistrés dans Wordpress. Que des avantages, donc. De plus, dans l'implémentation choisie, la connexion via Github montre clairement à l'étudiant s'il est connecté ou non.

Dès que l'étudiant s'est enregistré, les données utiles venant de Github sont rentrées dans la table des utilisateurs Wordpress. Nous pouvons donc combiner cette information avec les données provenant de Moodle pour établir un profil complet de l'étudiant dans notre site Wordpress automatiquement. A noter que Wordpress tente de récupérer aussi d'autres informations venant de Github, dont l'email. Cependant, ce dernier peut être caché dans le compte Github (c'est d'ailleurs le cas par défaut), et n'est donc pas accessible. La plupart du temps, nous n'avons donc que l'email UMONS à disposition. **Nous avons**

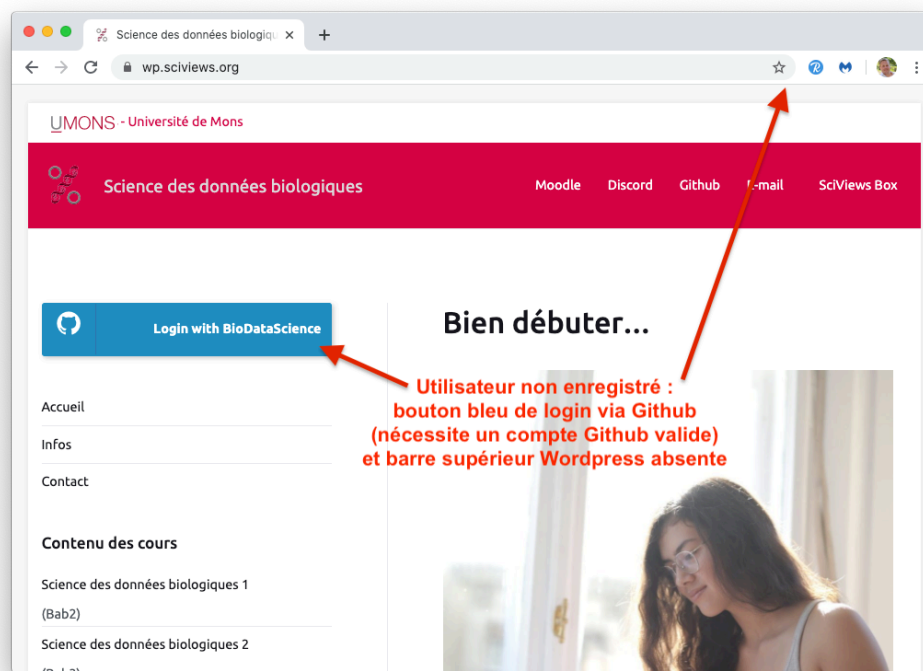


Figure 4.1: Page principale du site lorsque l'utilisateur n'est pas connecté (au travers d'un login BioDataScience dans Github).

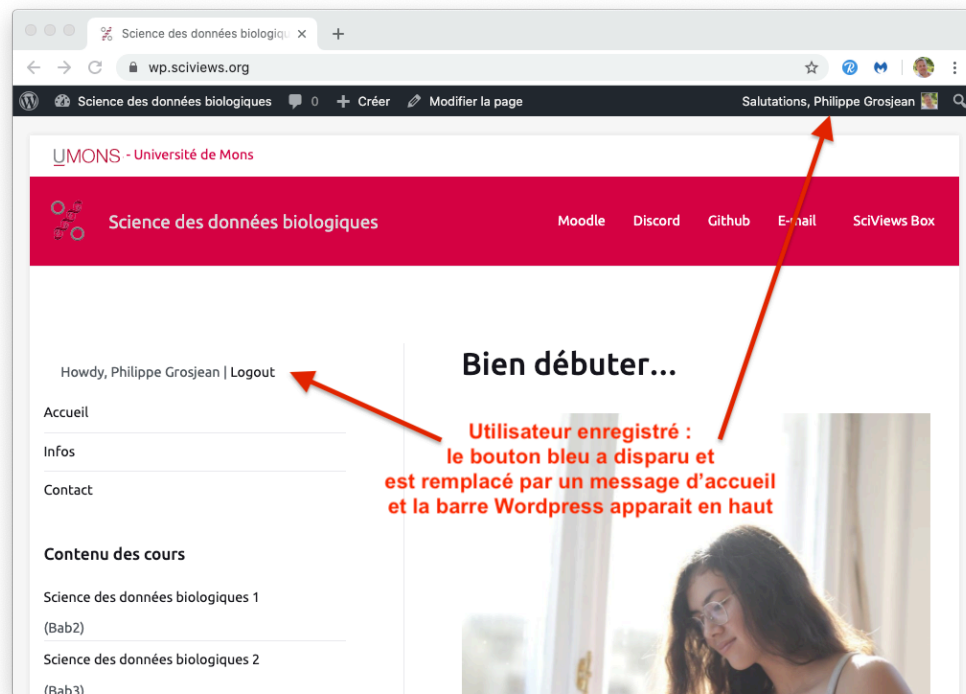


Figure 4.2: La même page principale du site lorsque l'utilisateur est enregistré et connecté via Github.

donc décidé d'utiliser l'email Wordpress, l'email UMONS, ou le login Github de manière interchangeable selon le contexte comme identifiant. Comme l'email du commiteur est automatiquement mis à celui de l'UMONS lorsque RStudio Serveur est accédé via Moodle, l'étudiant se doit d'indiquer le *même* email dans son compte Github, sans quoi, ses commits à partir de la SciViews Box ne pourront pas lui être attribués dans Github. *C'est de sa responsabilité.*



Le changement automatique d'email git dans la SciViews Box (à partir de la version 2020) ne se fait *que* si RStudio est accédé depuis le lien dans Moodle UMONS. Ainsi, un externe peut très bien utiliser n'importe quelle autre adresse mail sans problèmes.

L'adresse mail utilisée est celle renseignée par Moodle, soit `prénom.nom@[student].umons.ac.be`. Cependant, dans de rares cas, l'adresse UMONS utilise des caractères qui posent problèmes ailleurs (Github), comme une apostrophe dans le nom. Ce cas est automatiquement détecté et le système bascule alors sur l'adresse alternative `matricule@umons.ac.be`.

Les informations utilisateurs complémentaires sont fournis manuellement à Wordpress via un tableau CSV et le plugin "Import Users from CSV". Ce tableau provient de Moodle et/ou Github Classroom.

TODO: expliquer la procédure pour extraire la liste des étudiants d'un cours et les pré-enregistrer dans Wordpress avec le bon email.

4.2 Installation Wordpress

Notre nom de domaine `sciviews.org` est géré par `https://gandi.net`. Ce fournisseur propose aussi un hébergement web Wordpress dans un serveur LAMP préconfiguré. Nous utilisons cette option (pour l'instant en test). Le serveur utilise 1 cpu et 256Mo de RAM avec 20Go d'espace de stockage, et un quota de 300.000 pages vues par mois et 60Go par mois. Il peut être augmenté à tous niveaux à n'importe quel moment (plus de cpus ou de ram, plus de débit, et espace disque jusqu'à 1To). Nous verrons comment il se comporte en charge avec nos étudiants et pourrons toujours passer vers un plan supérieur si des lenteurs sont constatées. Le certificat SSL pour avoir une adresse sécurisée en `https://` est inclu dans le pack.

Le serveur est installé initialement avec **PHP 7.3** et une base de données **MySQL 5.7** ne permettant que des requêtes locales (impossible d'accès par l'extérieur par décision du fournisseur). **Wordpress** est à la version **5.4.1**.

Le site est servi à l'adresse `https://wp.sciviews.org`. Les enseignants y sont enregistrés comme éditeurs. En outre, un compte administrateur est aussi créé. Le serveur est accessible par SSH (après activation provisoire de la console dans Gandi), sFTP et git/SSH. Toutes les informations nécessaires à la gestion de l'hébergement sont transmises aux intéressés par voie sécurisée.

4.2.1 Plugins Wordpress

Nous avons installé les extensions suivantes :

- **H5P 1.15.0.** Il faut installer les libraires séparément et les principales ont été rajoutées dès le départ. Les 5 ou 6 widgets H5P que nous avons déjà créés sont facilement importés en les exportant de leur site original au format `.h5p` en cliquant

sur le bouton **Reuse** en bas à gauche du widget. Ensuite dans Wordpress, on va dans le tableau de bord -> H5P Content, on clique **Add New** puis **Téléverser** et on sélectionne le fichier **.h5p**. Une sauvegarde de ces fichiers **.h5p** est également effectué sur p-Cloud dans le dossier **sdd_h5p_backup**. Le plugin est ensuite configuré pour ne pas montrer la bannière H5P en bas des widgets et pour générer tous les événements xAPI possibles (voir 5).

- **SNORDIAN's H5PxAPIkatchu 0.4.2** capture les événements xAPI générés par H5P et les stocke directement dans la base de données MySQL liée à Wordpress. Il y a une interface simpliste pour visualiser les tables, les filtrer et les exporter au format CSV. Nous avons patché cette version pour y rajouter l'enregistrement des événements xAPI directement dans notre base de données MongoDB Atlas, et n'utiliser l'enregistrement local original que comme solution d'appoint au cas où l'enregistrement en MongoDB s'avère impossible pour une raison ou une autre (base saturée, indisponible, erreur, ou port 27017 de MongoDB bloqué).
- **OAuth Single Sign On - SSO (OAuth Client) 6.15.2** configuré pour enregistrer et connecter un utilisateur via son compte Github. Pour permettre ceci une application Github nommée **BioDataScience** a été créée sur Github. La configuration de ce plugin est disponible depuis le tableau de bord de Wordpress -> miniOrange -> OAuth.
- **Import Users from CSV 1.0.1** est un plugin permettant de créer des utilisateurs en batch depuis un fichier CSV. Nous pourrions ainsi exporter en bloc les données venant de Moodle ou Github Classroom et créer d'avance les comptes Wordpress avec des informations supplémentaires.
- **Advanced iFrame Pro 2020.3** permet d'insérer un iframe avec plein d'options : sécurisation, adaptation de la taille, passage de paramètres, réécriture d'URL, etc. Nous utilisons ce plugin pour intégrer les bookdown dans le site Wordpress, et ainsi conserver la bannière supérieure de notre site tout en affichant le bookdown. La version standard est installée depuis le site plugins Wordpress, et puis nous avons payé \$19 pour obtenir la version PRO qui propose des fonctionnalités en plus dont nous avons besoin.
- **My Custom Functions 4.42** permet de rajouter des fonctions PHP au site sans devoir créer de plugin Wordpress pour cela et sans "polluer" les autres fichiers (comme **functions.php** du thème). Les fonctions rajoutées créent des variables dans **localStorage** du browser web sur le client pour l'identification de l'étudiant entre bookdown H5P et les applis Shiny/learnr, pour enregistrer les événements H5P xAPI dans MongoDB via H5PxAPIkatchu patché, et pour modifier Wordpress de manière cosmétique (pas d'icône Wordpress dans la barre du haut, logo et lien BioDataScience dans la page de login).
- **UpdraftPlus 1.16.24** comme système de sauvegarde de l'ensemble (fichiers Wordpress + base MySQL). Le backup se fait en local pour l'instant dans **lamp0/web/vhosts/wp.sciviews.org/htdocs/wp-content/updraft**, mais il est recopié manuellement par sFTP sur notre disque p-Cloud dans le dossier **sdd_site_backup**. TODO: il nous faudra à terme une solution de backup automatisé sur p-Cloud ou OneDrive UMONS.
- Quelques autres plugins aident à gérer le site Wordpress : **Broken Link Checker 1.11.12**, **Health Check & Troubleshooting 1.4.4** & **WP-Optimize 3.0.19**.

Tous ces plugins sont distribués sur le site plugins de Wordpress, sont à jour et activement maintenus, avec un grand nombre d'utilisateurs. Nous avons évité les plugins plus

“confidentiels” pour lesquels nous ne sommes pas certains de la pérennité à plus long terme.

4.2.2 Thème du site

Pour le thème, nous avons choisi Zakra à jour, activement maintenu et totalement personnalisable. En jouant sur les options uniquement, nous avons pu façonner le site de telle manière qu’il reprenne une identité et les couleurs UMONS, tout en restant le plus simple, intuitif et facile d’accès possible.

4.3 Configuration serveur et git

Simple hosting de Gandi.net permet d’accéder aux fichiers sur le serveur de trois façons :

- Par sFTP (TODO: détails)
- Par Console SSH (TODO: détails). La console n’est pas active en permanence. Il faut l’activer dans le panneau de configuration du simple hosting sur Gandi.net. Cette activation reste valable pour 2 heures seulement par sécurité.
- Via git. C’est le plus simple pour développer et synchroniser son site. C’est d’ailleurs aussi la seule façon de faire pour ajouter des dépendances à PHP via composer (et nous aurons besoin de rajouter une dépendance par rapport à mongodb pour enregistrer les événements H5P xAPI dans notre base de données).

4.3.0.1 Sécurisation du site Wordpress

- Modification du fichier `.htaccess` dans `htdocs`. Afin d’éviter des attaques de type pingback xmlrpc sur le site Wordpress, il faut rajouter la ligne suivante dans le fichier `.htaccess` :

```
RewriteRule ^xmlrpc\.php$ "http:\/\/0\.\0\.\0\.\0\/" [R=301,L]
```

- Utilisation des plugins “Health Check & Troubleshooting” et “WP-Optimize” pour diagnostiquer et remédier aux problèmes potentiels. Un seul problème reste impossible à résoudre pour l’instant, il s’agit de “A plugin has prevented updates by disabling `wp_version_check()`.” qui peut avoir de nombreuses origines, et dans le cas de notre site, nous n’avons pas pu détecter la cause. Même en désactivant tous les plugins et le thème, le message reste.
- Nous avons aussi utilisé l’outil de sécurité et d’audit de Chrome dans les outils développeurs pour détecter les failles éventuelles et les erreurs côté client (Javascript, CSS, ...), et nous assurer que le site est sain. Globalement, il n’y a pas de failles majeures, mais certains aspects pourraient être améliorés pour avoir un site plus rapide, par exemple. Nous examinerons ces points plus tard.

4.3.0.2 Utilisation de git

Il faut git et un client SSH installé (pas de problèmes pour le Mac ou Linux). Les infos pour utiliser git sont dans la page de gestion du site `wp.sciviews.org` sur Gandi.net. Dans un premier temps, l’identification se fait via un mot de passe. Voici comment cloner le dépôt git à partir d’un répertoire racine de votre choix, mais nous aurons absolument besoin que ce répertoire racine soit le même que celui où les dépôts Github de BioDataScience-Course sont également placés (sur le Mac, il est facile de s’y placer en tapant `cd` suivi d’une espace, et puis en allant faire un drag&drop du dossier depuis le Finder vers le terminal, puis d’exécuter cette commande).

```
# Placez-vous dans le bon dossier
#cd <dossier racine des dépôts BioDataScience-Course>
git clone git+ssh://3693753@git.sd3.gpaas.net/wp.sciviews.org.git
```

Accepter (yes) pour le message relatif à l'authenticité, puis rentrer le mot de passe d'administration du site. **Vérifier que les fichiers du dépôt git ont bien été rajoutés dans le dossier avant de continuer.** Avant de pouvoir effectuer un push via Github Desktop, il faut avoir configuré git correctement et avoir une clé SSH.

```
git config --global user.name <username>
git config --global user.email <user@email.com>
# Générer une clé SSH pour l'authentification
ssh-keygen -t rsa -b 4096 -C '<user@email.com>'
```

Une fois la clé SSH générée, copier le contenu de `~/.ssh/id_rsa.pub` d'une part vers "SSH Keys" dans Github -> Settings -> SSH and GPG Keys (pour être aussi en ordre vis-à-vis de Github), et d'autre part, dans le panneau principal Simple hosting de Gandi.net, dans la section sécurité, cliquer sur "Ajouter une clé SSH". Puis créer une nouvelle clé et y coller le contenu.

Essayez une petite modification anodine dans votre dépôt git en local, puis commit et push. Ça doit fonctionner. Attention que les données qui sont poussées vers le serveur ne sont pas encore publiées à ce stade ! Il faut **déployer** ces changements en faisant :

```
ssh 3693753@git.sd3.gpaas.net 'deploy wp.sciviews.org.git'
```

Le **deploy** ajoute des nouveaux fichiers et remplace les fichiers existants sur le site (à condition que le droit en écriture sur les fichiers soit OK). Par contre, il n'élimine pas de fichiers. On risque donc d'accumuler des fichiers devenus inutiles au fur et à mesure de l'édition des blogdown. Si cela ne devrait pas impacter l'affichage des pages, cela utilise inutilement de l'espace disque sur le serveur. Donc, de temps en temps, il faut procéder à un nettoyage complet avec **clean** suivi d'un redéploiement.

```
ssh 3693753@git.sd3.gpaas.net 'clean wp.sciviews.org.git'
```

Attention qu'un fichier .gitignore correct est indispensable ici dans le dépôt. Les fichiers concernés par le **.gitignore** ne seront pas effacés. Tout le reste oui (faire toujours une sauvegarde complète avant) !.

4.3.1 Ajout de mongodb pour PHP

Le serveur sur Gandi.net n'autorise qu'une seule manière de rajouter des dépendances pour PHP : via git + ssh en utilisant **composer**. Et il faut créer les fichiers **composer.json** et **composer.lock** sur une machine locale avant de pouvoir faire un push sur le serveur. Cela signifie d'avoir PHP et la dépendance mongodb pour PHP installés localement. Ceci n'est pas facile, et il faut les mêmes versions que sur le serveur. Comme le PHP sous MacOS Catalina a la même version, j'ai d'abord essayé sur le Mac... sans succès car la puce de sécurité T2 empêche des modifications indispensables dans les dossiers `\usr\...` pour pouvoir compiler mongodb pour PHP. J'ai essayé aussi avec la SciViews Boxes 2019 (mais version PHP 7.2 trop ancienne) et une VirtualBox Ubuntu Server 20.04, mais version PHP 7.4 trop élevée. Gasp ! Finalement, j'ai pu le faire avec une VirtualBox Xubuntu 19.10. La procédure est détaillée ci-dessous.

- Noter la version X.Y de PHP utilisée sur le serveur (ici 7.3). Il faudra la même en local.

- Aller dans MongoDB Atlas et regarder la version de MongoDB déployée (ici 4.2). Il faudra une version de l'extension mongodb qui supporte à la fois la version de PHP et celle de MongoDB qui gère notre base de données.
- Rechercher la version compatible dans le tableau relatif à PHP ici. Nous pouvons voir qu'il nous faut l'extension 1.6+ ou 1.7+, avec la librairie 1.5 ou 1.6, respectivement, et par ailleurs cette version de l'extension est bien compatible avec la version 7.3 de PHP.
- Installer un (X)Ubuntu d'une version telle qu'il propose la bonne version de PHP. La 18.04LTS utilise PHP 7.2, pas bon. La 20.04LTS utilise la version 7.4 de PHP. Au final, une 19.10 convient, et c'est la version la plus récente avant qu'Ubuntu ne passe à PHP 7.4. Une machine virtuelle préinstallée peut être téléchargée depuis Osboxes. La façon de l'installer dans VirtualBox est expliquée sur le site. Le disque virtuel initial de la machine occupe 5,89Go.
- Configurer la machine virtuelle dans VirtualBox pour avoir un dossier partagé, par exemple 'Desktop' partagé en 'MacDesktop', avec connexion automatique.
- Lancer la machine virtuelle. Login : "oxboxes", mot de passe : "oxboxes.org", mais attention, en clavier configuré QWERTY. Avec un AZERTY, il faut entrer en fait "oxboxes:org" comme mot de passe.
- Aller dans le panneau de configuration et changer le clavier (belge, ou français (Macintosh) par exemple).
- Lancer un terminal et exécuter les lignes suivantes :

```
# Rendre le dossier partagé VirtualBox accessible pour l'utilisateur oxboxes
sudo usermod -aG vboxsf osboxes
# Mettre la machine à jour
sudo apt update
sudo apt upgrade
sudo apt autoremove
sudo apt autoclean
sudo apt clean
# Redémarrer
sudo reboot
```

- Il se peut que la machine virtuelle rebascule en QWERTY. Vérifier et corriger à nouveau. Ensuite, installer PHP et ses dépendances nécessaires :

```
sudo apt install php7.3-cli php7.3-dev php-pear php-mongodb composer
```

Ensuite, nous devrions pouvoir créer les fichiers nécessaires via la commande `composer require mongodb/mongodb`, mais cela ne fonctionne pas lorsque `php-mongodb` est installé via le package Ubuntu. Nous devons le compiler nous-même avec `pecl` sur la machine¹ :

```
# Créer un répertoire dans le dossier partagé et s'y placer
mkdir /media/sf_MacDesktop/php-mongodb
cd /media/sf_MacDesktop/php-mongodb
# Compiler mongodb depuis les sources
sudo pecl install mongodb
```

¹Je suis arrivé à créer `composer.lock` via la commande `composer update` en ayant écrit à la main un `composer.json` qui contient les lignes minimales renseignées sur la page d'aide de Gandi.net, mais j'aboutis à une version trop ancienne du driver, et de toutes façons, c'est un bricolage qui ne me plaît pas !

```
# Créer composer.json et composer.lock (entre autres)
composer require mongodb/mongodb
```

- Copier `composer.json` et `composer.lock` ainsi obtenus à la racine du dépôt git, faire un commit, un push et un deploy. Cette fois-ci mongodb PHP 1.7 semble s'être bien installé sur le serveur lors du deploy.
- Faire un test dans Wordpress. Aller jusqu'à un contenu H5P quelconque sur le site et interagir avec lui. Aller ensuite voir si les données correspondantes ont bien été enregistrées dans notre MongoDB ATLAS, dans la collection `sdd.h5p`. Si pas, vérifier si ces mêmes données sont enregistrées dans H5PxAPIkatchu dans le tableau de bord Wordpress. Si c'est le cas, c'est un problème d'accès à la base de données MongoDB ATLAS, sinon, c'est un problème plus profond : est-ce que l'enregistrement xAPI est bien activé, et est-ce que les extensions bookdown nécessaires ont bien été injectées via un fichier `header.html` (voir configuration du bookdown 6) ? L'option de débogage de H5PxAPIkatchu se serait s'avérer utile ici pour déterminer ce qui ne va pas, voir `@reg(h5p)`.

4.4 Agencement du site

La page principale <https://wp.sciviews.org> sert essentiellement à configurer le site lorsqu'elle est appelée depuis Moodle (enregistrement des données de l'étudiant dans un `localStorage`) et donne les instructions principales pour bien démarrer avec le contenu pédagogique.

Cette page laisse une large place à gauche pour le login Github, quelques liens utiles (infos, contact) et liste les différents bookdown disponibles, un par cours.

Chaque bookdown est dans une sous section du site : `/sdd_umons/`, `/sdd_umons2/`, etc. Une fois la page principale activée depuis le lien Moodle **Cours**, l'étudiant peut aller vers le bookdown qui correspond à son cours et tout devrait fonctionner correctement. La page principale de chaque bookdown effectue de son côté les vérifications nécessaires, affiche un court compte-rendu, et permet d'effacer les données personnelles (compatibilité RGPD) si souhaité, voir 6.

Le site est complété d'une bannière supérieure dont l'aspect rappelle le site UMONS et qui propose des liens vers différents items importants dans le cadre du cours pour obtenir de l'aide (Moodle, Discord, Email, ...), et donne aussi accès à l'explication pour installer le logiciel (SciViews Box). Et c'est tout. Le but est de faire ici simple et efficace...

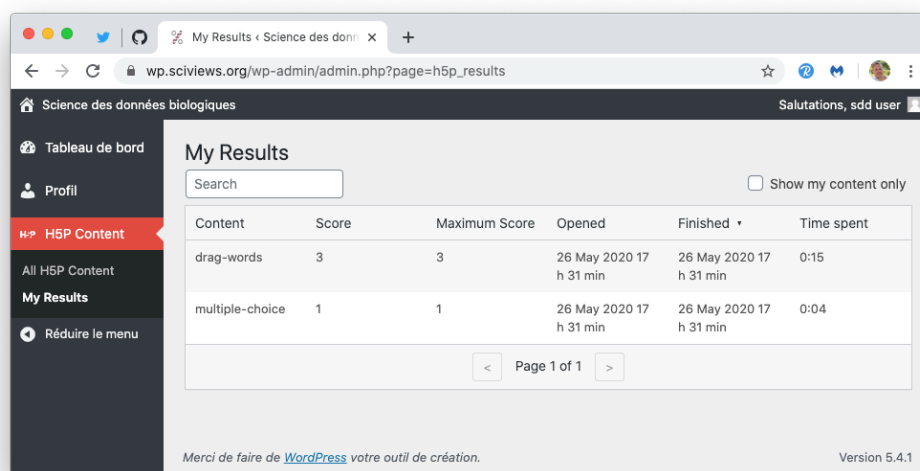
4.4.1 Tableau de bord

Dès qu'un utilisateur s'enregistre sur le site, une barre supérieure noire apparaît avec des options supplémentaires en fonction de son rôle. Un étudiant a évidemment un nombre d'options limitées à l'édition de son profil et la consultation de sa progression H5P. Un éditeur (prof) peut aussi créer et modifier des pages. Enfin, un administrateur a accès à l'ensemble des fonctions pour personnaliser le site.

Pour tous les utilisateurs, un **tableau de bord** est accessible à partir de cette barre noire supérieure. Une fois le compte créé dans Wordpress, un étudiant a la possibilité de s'y connecter via Github et de modifier ses informations dans le tableau de bord, dont son adresse email. S'il le fait, cette adresse n'est alors plus la même que celle de l'UMONS. C'est pourquoi, nous conservons ces deux informations séparément : `email` pour l'email du site Wordpress et `iemail` pour l'email institutionnel tel que renseigné à partir de

Moodle. C’est toutefois l’adresse email enregistrée dans Wordpress qui est employée pour enregistrer les activités H5P, et par extension pour les learnr et les app Shiny. En fait, pour ces derniers, le choix de l’identifiant est le suivant : (1) l’email du compte Wordpress, (2) si vide, l’email UMONS, (3) si vide, le login Github/Wordpress. Si aucune des trois infos n’est disponible, nous n’enregistrons **pas** l’activité learnr et app Shiny, et l’outil de test en début de bookdown indique clairement que la navigation se fait de manière anonyme. Par contre, H5xAPIkatchu enregistre l’activité h5p sous un UUID créé à la volée. Nous pouvons toujours filtrer et éliminer ces données-là plus tard.

Par ailleurs, les événements principaux H5P (début, fin, durée et résultat à la soumission des exercices) sont également enregistrés dans Wordpress directement. Cette information peut paraître redondante avec l’enregistrement H5PxAPIkatchu, mais elle est utile car l’étudiant y a accès directement dans son tableau de bord. Il peut y lire les détails de sa progression dans l’ensemble des exercices H5P de tous les cours, s’il le souhaite (voir capture d’écran). Reste à déterminer l’occupation disque de cette fonction afin de décider si nous la conservons activée sur le long terme !



Content	Score	Maximum Score	Opened	Finished	Time spent
drag-words	3	3	26 May 2020 17 h 31 min	26 May 2020 17 h 31 min	0:15
multiple-choice	1	1	26 May 2020 17 h 31 min	26 May 2020 17 h 31 min	0:04

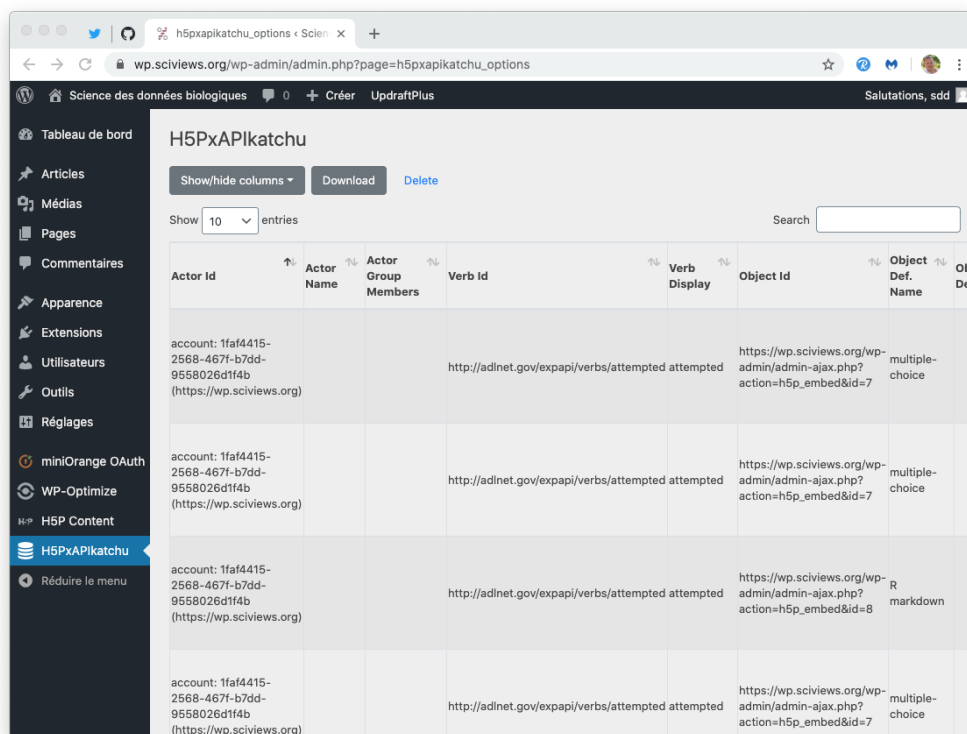
Figure 4.3: Suivi de la progression H5P dans le tableau de bord accessible aux étudiants.

Comme nous pouvons le voir, le tableau de bord de l’étudiant est réduit au minimum. En plus de la partie “H5P Content” où il peut voir et filtrer la liste des applis H5P et voir ses propres résultats H5P, il a accès à son profil où il peut modifier des données utilisées par nos outils. Il me semble que la maîtrise et la visibilité des données enregistrées le concernant font partie des contraintes RGPD pour lesquelles Wordpress assure la compatibilité.

Les enseignants sont enregistrés comme éditeurs sur le site. Dès lors, ils peuvent aussi modifier le contenu du site, et surtout, modifier et créer du contenu H5P avec des éditeurs conviviaux. ainsi, l’ensemble des outils nécessaires pour l’édition du contenu sont regroupés dans ce tableau de bord... à l’exception des bookdowns et des applis Shiny/learnr. Ces dernières sont mise à disposition via un serveur RStudio Connect, totalement différent donc. Par contre, les bookdown doivent, eux, être intégrés dans le site. La façon de réaliser cette intégration est expliquée dans la section consacrée à bookdown (6).

4.4.2 Gestion des événements xAPI

Pour les enseignants, les événements xPI générés par H5P (voir @ref(h(P))) qui n'ont pas pu être enregistrés dans la base de données MongoDB ATLAS dédiée pour une raison ou une autre sont accessibles à partir du tableau de bord de Wordpress dans la section H5PxAPIkatchu. De là, ils peuvent être exportés au format CSV, et réinjectés dans MongoDB à l'aide d'un peu de code R.



The screenshot shows the WordPress admin dashboard for 'Science des données biologiques'. The left sidebar contains various menu items, with 'H5PxAPIkatchu' highlighted. The main content area displays a table titled 'H5PxAPIkatchu' with columns for Actor Id, Actor Name, Actor Group Members, Verb Id, Verb Display, Object Id, Object Def. Name, and Object Description. The table contains four rows of data, each representing an xAPI event. Above the table, there are buttons for 'Show/hide columns', 'Download', and 'Delete', and a search bar. The 'Download' button is highlighted in blue.

Actor Id	Actor Name	Actor Group Members	Verb Id	Verb Display	Object Id	Object Def. Name	Object Description
account:1faf4415-2568-467f-b7dd-9558026d1f4b (https://wp.sciviews.org)			http://adlnet.gov/expapi/verbs/attempted	attempted	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=7	multiple-choice	
account:1faf4415-2568-467f-b7dd-9558026d1f4b (https://wp.sciviews.org)			http://adlnet.gov/expapi/verbs/attempted	attempted	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=7	multiple-choice	
account:1faf4415-2568-467f-b7dd-9558026d1f4b (https://wp.sciviews.org)			http://adlnet.gov/expapi/verbs/attempted	attempted	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=8	R markdown	
account:1faf4415-2568-467f-b7dd-9558026d1f4b (https://wp.sciviews.org)			http://adlnet.gov/expapi/verbs/attempted	attempted	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=7	multiple-choice	

Figure 4.4: Événements xAPI dans H5PxAPIkatchu. En bas du tableau, on peut filtrer les données, et le bouton **Download** permet de récupérer le tableau filtré au format CSV. Le lien **Delete** permet de tout effacer et n'est disponible que pour un administrateur du site.

4.5 Variables contextuelles

Pour le bon fonctionnement de nos outils et l'enregistrement de l'activité des étudiants, nous avons besoin de définir un certain nombre de variables reprises dans le tableau ci-dessous. Toutes ces variables sont obtenues directement à la source². La colonne origine indique d'où nous obtenons ces informations (pour Moodle, il faut que l'étudiant ait lancé https://wp.sciviews.org depuis son compte Moodle et à partir de la page de cours correspondante).

²En 2018-2019 et 2019-2020, nous utilisons le login Github et l'adresse email UMONS pour identifier les événements learnr, mais comme les étudiants devaient rentrer ces infos manuellement, nous avons eu beaucoup de soucis avec des encodages erronés... Donc maintenant, nous récupérons ces infos de manière automatique à la source autant que possible.

Variable	Contenu	Origine
login	Le login Github de l'étudiant	WP via Github
email	Le mail renseigné dans Wordpress	WP via Github si public
displayname	Nom utilisateur Wordpress(/Github)	WP
firstname	Prénom comme indiqué dans Wordpress	WP
lastname	Nom de famille comme indiqué dans Wordpress	WP
iemail	Le mail UMONS (institutionnel) de l'étudiant	Moddle iemail= Adresse de courriel
iid	Numéro de matricule de l'étudiant	Moodle iid = Nom d'utilisateur
ifirstname	Prénom comme indiqué dans Moodle	Moodle ifirstname = Prénom
ilastname	Nom de famille comme indiqué dans Moodle	Moodle ilastname = Nom
institution	Institution (ici UMONS)	Moodle institution = Institution
icourse	Nom de code Moddle du cours (ex.: S-BIOG-006)	Moodle icourse = N° d'identification du cours
iurl	URL du serveur institutionnel (Moodle)	Moodle iurl = URL du serveur
iref	Identifiant unique dans Moodle	Moodle iref = Numéro d'identification
ictitle	Titre du cours (ex. Science des Données I: visualisation)	Moodle ictitle =

A noter que nous croisons différentes sources (Moodle, Github, Wordpress), et les infos peuvent différer entre ces sources. Considérant un étudiant qui a “installé” ses outils correctement, à savoir :

- Son compte est créé au préalable dans Wordpress à partir du tableau des utilisateurs du cours issu de Moodle ou de Github Classroom³,
- Il a créé un compte Github et y a renseigné son adresse mail UMONS,
- Il a au moins une fois lancé la page <https://wp.sciviews.org> depuis le lien Moodle “Cours” dans la section “Ressources en ligne” de son cours,
- Dans la page qui s'ouvre, il s'est connecté dans Wordpress à partir du bouton de login via Github/BioDataScience.

A ce moment-là, tout sera en place pour gérer correctement son identité dans les différents matériels pédagogiques (bookdown, H5P, learnrs, apps Shiny, etc.) Il ne lui restera plus

³Il serait possible de proposer l'enregistrement libre, mais il faut acheter la version pro du plugin d'identification Github... et de plus, tout le monde ayant un compte Github pourrait alors s'enregistrer ! C'est pas ce qu'on souhaite.

alors qu'à installer la SciViews Box pour être pleinement opérationnel. Cela semble être un nombre important d'opérations à réaliser pour arriver à cela, mais l'étudiant est guidé de manière naturelle : il va dans Moodle, se connecte et sélectionne le cours dans la liste (il a déjà l'habitude). A partir de là, il clique naturellement sur "Cours", d'où la page qui l'enregistre et lui propose de créer un compte Github. Une fois que c'est fait, il peut cliquer sur le gros bouton bleu "Github/BioDataScience" pour se connecter dans le site "Science des Données Biologiques", et il sélectionne le manuel correspondant à son cours dans la liste. A ce stade, c'est fait. Il est enregistré !



Faut-il *forcer* le login de l'étudiant dans Wordpress ? Autrement dit, il n'aurait accès aux cours que s'il est connecté. Il existe des plugins Wordpress qui permettent de faire cela, et aussi d'exclure les pages qui peuvent être vues par tous. Un de ces plugins est **Force login**. Nous verrons à l'usage...

Toutes les informations de son compte lui sont accessibles dans Wordpress, et il peut librement les modifier dans le tableau de bord. Donc, cela signifie que l'information cruciale pour identifier un étudiant est `iemail`, son adresse email de l'UMONS fournie par Moodle. Toutefois, nous devons recouper avec son `login` Github pour obtenir les données issues de Github pour le même étudiant. Nous avons éventuellement besoin aussi de `email`, son email Wordpress sous lequel les activités H5P sont enregistrées.

Si nous continuons à utiliser Discord, ce serait bien d'avoir le login Discord de l'étudiant aussi. Mais ici, il a le droit d'y poser des questions de manière anonymisée (c'est notre souhait). Néanmoins, il serait utile d'avoir un champ où nous pouvons éventuellement rentrer cette info à la main si souhaité. Une autre colonne `comment` serait aussi utile. Nous pourrions y indiquer par exemple si un étudiant a des difficultés particulières, ou la date de son arrivée pour les étudiants étrangers en passerelle, par exemple.

Chapter 5

H5P

H5P permet de “créer, partager et réutiliser du contenu interactif HTML5 dans un browser Web”, d’après la page d’introduction du site Web. Il s’agit d’un ensemble de contenus HTML5 (html + css + javascript) permettant d’afficher et d’interagir avec des exercices et des présentations sous plusieurs formes. Plus de 40 contenus différents sont disponibles. H5P est distribué sous forme de logiciel libre (gratuit et Open Source) qu’il est possible d’installer sur son propre serveur. De plus, h5p.org propose l’hébergement de contenus gratuits **à des fins de test uniquement**, et offre en complément une formule payante sur h5p.com, mais le tarif n’est pas renseigné.

Dans cette section, nous reprendrons les choix techniques, l’installation, la configuration et le déploiement de H5P dans le cadre de nos cours de science des données à l’UMONS, ainsi que les principaux contenus H5P qui sont disponibles.

5.1 Installation de H5P

H5P ne peut fonctionner seul. Il doit être inclus dans un serveur plus large, soit un LMS (Learning Management System ou plateforme d’apprentissage en ligne comme), soit un CMS (Content Management System, système de gestion de contenu).

L’un des LMS les plus populaires est Moodle qui a été choisi à l’UMONS comme outil institutionnel d’apprentissage en ligne. H5P est compatible avec Moodle (il peut être installé comme plugin). Ce serait donc parfaitement logique de l’utiliser de cette manière. Un test avait été fait en 2019. Si cela fonctionne, nous avons pu constater que le contenu H5P n’est **pas** utilisable ailleurs que dans Moodle lui-même, en tous cas de la façon dont le serveur test était configuré, avec les sécurités nécessaires pour protéger le site. Or, notre souhait est d’intégrer le contenu H5P dans un bookdown qui ne peut être hébergé sur Moodle¹. L’essai qui a été réalisé à l’UMONS en 2019 n’a pas été concrétisé dans le serveur Moodle en production. Or, il nous fallait intégrer du contenu H5P dans nos bookdown, et ce, dès 2020-2021. Donc, nous avons recherché une autre option.

Cette autre option consiste à utiliser soit Wordpress, soit Drupal pour lesquels H5P est également compatible. Après comparaison des possibilités des deux, nous avons opté pour Wordpress, plus largement utilisé.

¹D’un point de vue technique, le contenu H5P s’affiche dans un iframe. Or les possibilités des iframes inter-sites sont limitées par le protocole HTML5 sous-jacent lui-même pour éviter les abus. Il nous faut donc le bookdown (pages statiques) et le contenu H5P (devant être lié à un serveur LMS ou CMS) servis par le *même* site.

H5P permet de récupérer des informations concernant l'activité des étudiants dans les exercices sous forme de requêtes xAPI. xAPI est un standard développé pour permettre l'interaction entre un LMS comme Moodle qui produit des événements xAPI, et un LRS (Learning Record Store) une base de données chargée d'enregistrer les traces de l'apprentissage au format standardisé xAPI. En complément de H5P lui-même et de Wordpress comme serveur principal, il nous faut donc un LRS et un plugin qui fait le pont entre H5P et ce LRS. Enfin, l'information n'est utile que si elle est identifiable comme provenant d'un étudiant en particulier. Le serveur doit donc gérer l'authentification des utilisateurs et transmettre l'identité de l'étudiant au travers de la requête xAPI. Si tout cela paraît très compliqué, voici un petit schéma qui résume l'ensemble (à noter que nous avons choisi l'authentification via Github parce que nos étudiants ont également d'autres activités sur Github/Github Classroom, et il est dès lors plus facile de gérer une identité commune à toutes ces activités, même si l'identité UMONS est, hors contexte, la plus logique ici) :

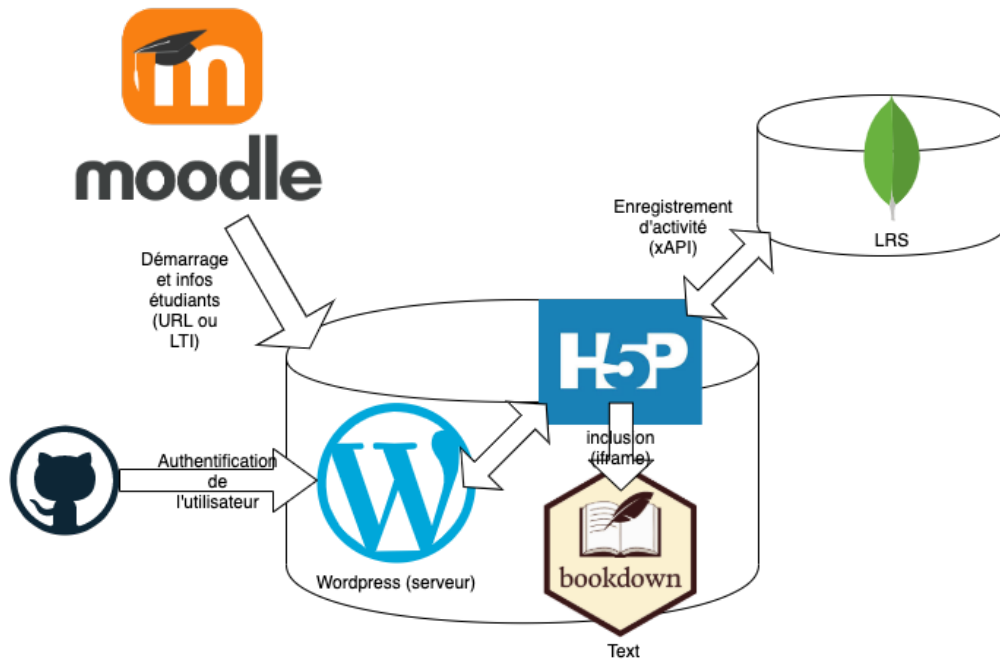


Figure 5.1: H5P et bookdown sont intégrés dans le même serveur Wordpress, ce qui permet l'affichage en iframe sans contraintes. Les informations de l'utilisateur proviennent d'une authentification Github, plus des données issues de Moodle via URL ou LTI. L'activité des étudiants H5P peut ainsi être enregistrée dans le LRS de notre choix via le protocole standard xAPI avec une identification univoque.

5.1.1 Installation du plugin H5P

Une fois le serveur Wordpress opérationnel, l'installation de H5P lui-même est vraiment très simple.

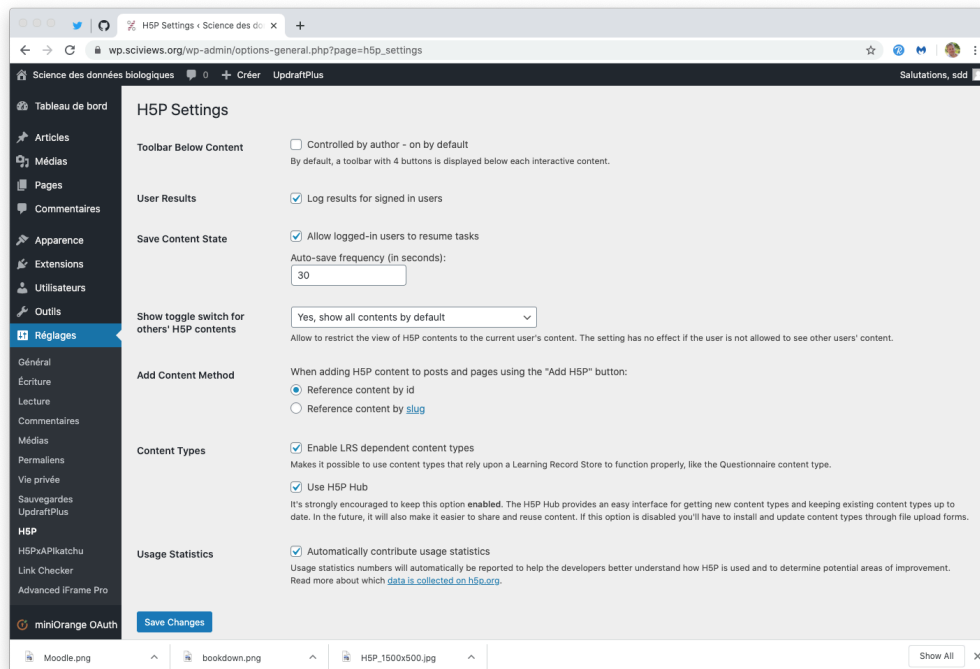
- Connection au serveur Wordpress en temps qu'administrateur,
- Aller dans le tableau de bord -> Extensions -> Ajouter.
- Dans le champ de recherche, introduire "H5P",

- Sélectionner “Interactive content - H5P”. Cliquer “Installer”, puis “Activer”.

C’est fait ! H5P est installé, mais aucune forme de contenu n’est encore disponible. H5P conseille d’installer des **librairies** de contenu. A partir du tableau de bord, cela se fait via “H5P Content -> Libraries”. Cependant, cela n’apparaît pas nécessaire en pratique car lorsqu’on importe du contenu H5P dans le site, les librairies nécessaires sont automatiquement importées également... plus d’explications ci-dessous.

5.1.2 Configuration de H5P

Le panneau de configuration est accessible pour l’administrateur Wordpress depuis le tableau de bord -> Réglages -> H5P. Voici une copie d’écran de la configuration choisie.



- La barre d’outils H5P en dessous du contenu est désactivée par défaut (“Controlled by author” non coché). Cette barre d’outils permet le partage et la réutilisation de contenu. Nous la garderions bien activées, mais le bouton de réemploi doit apparemment être activé pour que l’inclusion dans un bookdown soit possible. Or, si ce bouton est activé, tout le monde peut inclure le contenu dans d’autres sites... et impacter notre quotas d’utilisation de requêtes. Nous ne le souhaitons bien évidemment pas. *Pour récupérer un contenu sous forme de fichier .h5p, il faut réactiver cette option temporairement dans le tableau de bord, et seul l’administrateur du site peut le faire (voir backup H5P plus loin).*
- Les résultats utilisateurs (“Log results...” est coché) permet d’avoir des statistiques de base utilisateur par utilisateur sur les H5P. C’est utile, mais il faudra voir à terme l’impact sur l’espace disque utilisé pour cette fonction.
- La sauvegarde du contenu (“Allow logged-in users to resume tasks”) est activée et utile puisqu’elle permet la mémorisation d’une session à l’autre des sélections dans les contenus H5P pour les utilisateurs enregistrés.

- “Show toggle switch for others’ H5P contents” n’est pas très clair pour moi. Je l’ai laissé pour l’instant sur “Yes, show all contents by default”. Est-ce qu’il s’agit de voir le contenu *créé* par d’autres utilisateurs ? Si, oui, c’est ce qu’on veut (les étudiants utilisent le contenu créé par les profs).
- “Add content method” est réglé sur “id”.
- L’option suivante, “Enable LRS dependent content types” est importante et doit être coché pour pouvoir enregistrer l’activité via xAPI (voir plus loin).
- Les deux dernières options sont en relation avec la gestion et le développement H5P. Nous les laissons cochées (“Use H5P Hub” et “Automatically contribute usage statistics”).

5.2 Enregistrement H5P xAPI

Une fois le plugin H5P installé, il nous faut encore permettre l’enregistrement des activités des étudiants. Nous l’avons déjà vu, cette possibilité est prévue dans H5P et le protocole standardisé xAPI est utilisé, ce qui le rend compatible avec les LRS du marché. Dans Wordpress, nous avons deux options :

1. Un enregistrement *local* dans la base de données MySQL de Wordpress directement sur le serveur avec le plugin Wordpress **H5PxAPIkatchu**, voire une modification de ce plugin pour l’enregistrement dans une base de données généraliste, comme MongoDB,
2. Un enregistrement dans un LRS spécialisé xAPI avec le plugin Wordpress **wp-h5p-xapi**.

Nous allons aborder et discuter ces deux options tour-à-tour.

5.2.1 H5PxAPIkatchu

Ce plugin est facile à installer puisqu’il est disponible depuis Wordpress plugins. La procédure est la même que pour l’installation de H5P lui-même (l’administrateur va dans la partie installation d’extension du tableau de bord, rentre “H5PxAPI” dans la zone de recherche, sélectionne l’item et clique “Installer” puis “Activer”). Ce plugin est gratuit et il semble bien maintenu².

La configuration est également extrêmement simple.

Comme nous souhaitons récolter un maximum d’information sur l’activité des étudiants, et voulons récupérer le contenu xAPI intact, nous cochons “Store complete statements” et “Capture everything”. L’option “Debug” n’est pas nécessaire et peut rester décochée.

Une fois cette configuration effectuée, l’activité H5P est enregistrée directement dans la base de données MySQL de Wordpress. Cette base de données n’est pas disponible de l’extérieur dans notre configuration (le fournisseur verrouille les bases de données pour un accès local uniquement par sécurité). Par contre, H5PxAPIkatchu permet de visualiser un tableau avec les données, de le filtrer, et de récupérer le résultat sous forme de fichier CSV avec le bouton “Download” (le lien “Delete” permet d’effacer le contenu et n’est disponible que pour l’administrateur du site).

Et c’est tout ! Pas besoin de LRS, pas besoin de configuration à rallonge... par contre, cela ne convient que pour du *dépannage*. En effet, l’accès aux données n’est pas possible

²J’ai fait une proposition de pull request à l’auteur qui a répondu immédiatement.

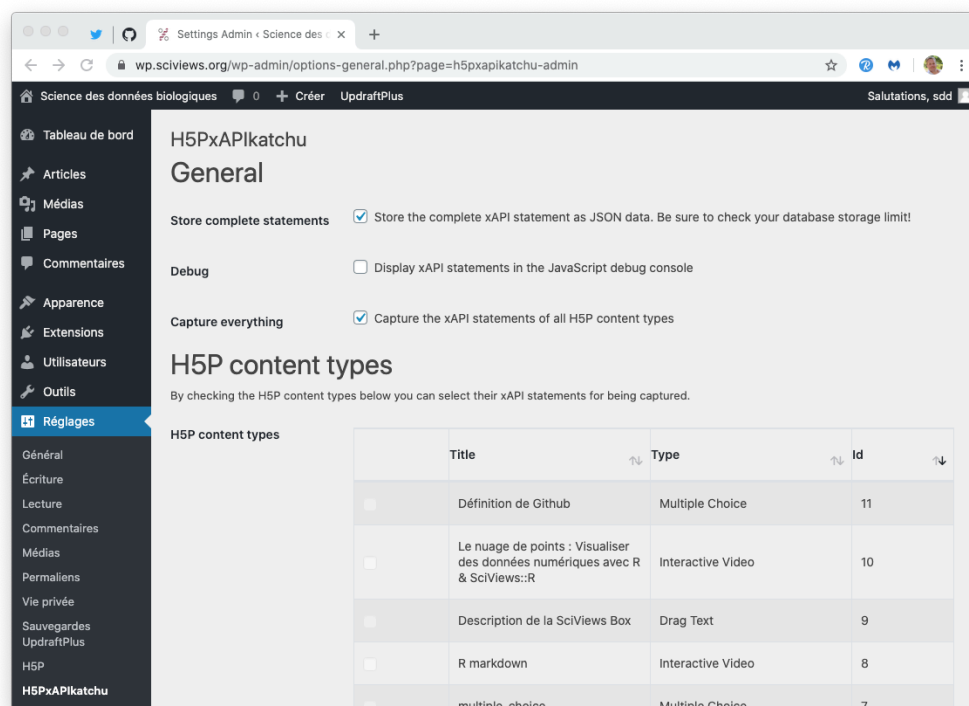


Figure 5.2: Configuration de H5PxAPIkatchu.

pour, par exemple, générer les rapports de progression des étudiants. L'extraction sous forme CSV doit se faire *manuellement* !

Après l'enthousiasme suite à un premier essai d'enregistrement de l'activité réussie lorsque le contenu H5P est utilisé directement dans Wordpress (à partir d'une page Wordpress ou du tableau de bord), nous faisons face à une déception quand nous constatons que cela ne **fonctionne pas dans le bookdown** ! Pourtant, nous avons bien installé notre bookdown dans le *même* serveur que Wordpress/H5P pour permettre une communication fluide... alors, pourquoi ça ne fonctionne toujours pas ? En fait, nous devons nous plonger dans les détails de l'implémentation pour comprendre que ce n'est **pas** H5P lui-même qui effectue l'enregistrement, mais la page dans laquelle le contenu est inclus. Autrement dit, notre brave bookdown en pages statiques va se voir déléguer le travail. Le schéma suivant explique le détail de l'implémentation.

Moralité, pour que notre bookdown enregistre les événements H5P, il faut lui adjoindre le code en Javascript nécessaire pour faire le transfert. Ce code est repris et adapté de H5P lui-même, ainsi que de H5PxAPIkatchu. Ce code est inclus dans le fichier `header.html` qui est inclus dans le bookdown grâce à ceci dans `index.Rmd` :

```
output:
  bookdown::gitbook:
    includes:
      in_header: header.html
```

Tout d'abord, la partie H5P de configuration. La partie **H5PIntegration** est normale-

The screenshot shows the H5P API Katchu admin interface. The left sidebar contains a menu with options like 'Tableau de bord', 'Articles', 'Médias', 'Pages', 'Commentaires', 'Apparence', 'Extensions', 'Utilisateurs', 'Outils', 'Réglages', 'miniOrange OAuth', 'Shortcode Variables', 'WP-Optimize', 'H5P Content', 'Advanced iFrame Pro', and 'H5P API Katchu'. The main content area displays a table of registered xAPI events. The table has columns for Actor Id, Actor Name, Actor Group Members, Verb Id, Verb Display, Object Id, and Object Def. Name. The first row shows an actor named 'Philippe Grosjean' with a verb 'attempted' and an object 'Untitled Single Choice Set'. The second row shows an actor named 'Philippe Grosjean' with a verb 'answered' and an object 'drag-words Bc'. The third row shows an actor named 'Philippe Grosjean' with a verb 'interacted' and an object 'drag-words Bc'.

Actor Id	Actor Name	Actor Group Members	Verb Id	Verb Display	Object Id	Object Def. Name
	Philippe Grosjean		http://adinet.gov/expapi/verbs/attempted	attempted	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=87subContentId=3f31e62e-9c9c-457f-9a7d-ad08060f3558	Untitled Single Choice Set
	Philippe Grosjean		http://adinet.gov/expapi/verbs/answered	answered	https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&id=6	drag-words Bc
	Philippe Grosjean		http://adinet.gov/expapi/verbs/interacted	interacted	https://wp.sciviews.org/wp-admin/admin-ajax.php?	drag-words Bc

Figure 5.3: Contenu xAPI enregistré par H5PxAPlkatchu. L'utilisateur est correctement identifié dans le tableau enregistré.

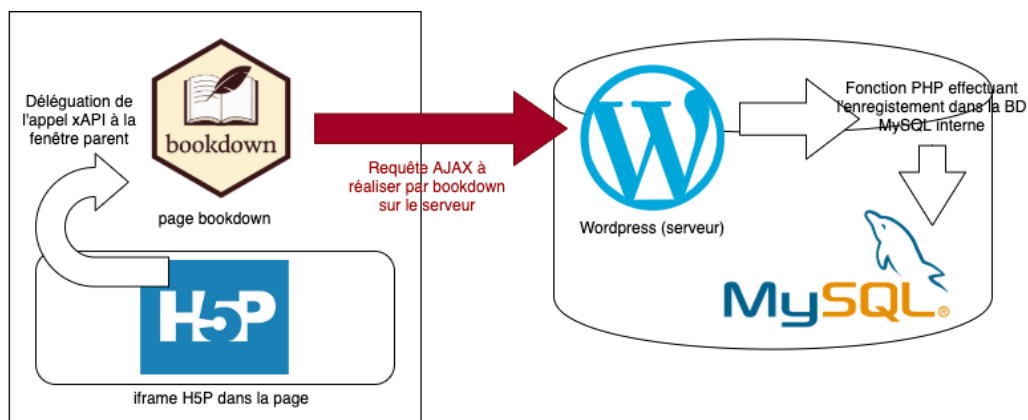


Figure 5.4: Pour des raisons de sécurité, l'enregistrement de l'événement xAPI se fait “server-side” par du code PHP dans H5PxAPlkatchu. Par contre, H5P ne communique jamais lui-même avec le LRS, mais délègue cette tâche à la page parent qui l’héberge (ce qui ne fonctionne que si les deux sont issus du même domaine). Donc, notre bookdown va devoir effectuer le transfert de l'événement xAPI via une requête AJAX sur le serveur (partie en rouge) qui manque actuellement.

ment générée à partir de Wordpress, mais comme nous utilisons des pages statiques dans bookdown, nous le copions tel quel ici. Comme elle contient des informations contextuelles liées au site et à l'installation de H5P, nous devons récupérer cela autrement afin d'avoir une version à jour, mais pour l'instant, ça fonctionne comme cela. Voici ce que cela donne :

```
<!-- Required for H5P xAPI feedback -->
```

```
<script>H5PIntegration = {"baseUrl":"https:\\\\wp.sciviews.org","url":"\\wp-content\\uploads\\h5p\\h5p-content-type\\h5p-content-type-confirmation-dialog.css" href='https://wp.sciviews.org/wp-content/uploads/h5p/h5p-content-type/h5p-content-type-confirmation-dialog.css'></script>
<link rel='stylesheet' id='h5p-core-styles-h5p-confirmation-dialog-css' href='https://wp.sciviews.org/wp-content/uploads/h5p/h5p-core-styles-h5p-confirmation-dialog-css'></link>
<link rel='stylesheet' id='h5p-core-styles-h5p-core-button-css' href='https://wp.sciviews.org/wp-content/uploads/h5p/h5p-core-styles-h5p-core-button-css'></link>
<script src='https://wp.sciviews.org/wp-includes/js/wp-embed.min.js'></script>
```

Et c'est pas tout ! Il nous faut encore le code Javascript H5P et H5PxAPIkatchu :

```
<script>
var wpAJAXurl = "https:\\\\wp.sciviews.org\\wp-admin\\admin-ajax.php";
var debugEnabled = "0";
var captureAllH5pContentTypes = "1";
var h5pContentTypes = [""];
window.H5PxAPIkatchu = {
  captureAllH5pContentTypes: captureAllH5pContentTypes,
  debugEnabled: debugEnabled,
  h5pContentTypes: h5pContentTypes,
  jQuery: jQuery,
  wpAJAXurl: wpAJAXurl
};
</script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/jquery.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-event-dispatcher.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-x-api-event.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-x-api.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-content-type.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-confirmation-dialog.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/h5p-action-bar.js'></script>
<script src='https://wp.sciviews.org/wp-content/plugins/h5p/h5p-php-library/js/request-queue.js'></script>
```

Voilà, avec tout cela, nous recompilons le bookdown, plaçons tout le contenu de \docs sur le serveur... et (après bien des essais et erreurs tout de même)... ça marche !



En dépit de sa simplicité, H5PxAPIkatchu ne convient *pas* tel quel pour l'objectif que nous poursuivons. Nous avons imaginé étendre le plugin pour enregistrer les données dans notre base de données MongoDB Atlas comme première solution. Cela a nécessité du développement PHP et l'utilisation d'une librairie PHP mongodb qu'il faut installer sur le serveur. Une autre solution consiste à utiliser un autre plugin et à passer par un véritable LRS. Ceci est envisagé dans le paragraphe suivant.

Il s'avère que, une fois l'incorporation du driver mongodb dans PHP effectuée (c'est en fait la partie la plus difficile), le code en PHP qu'il faut ajouter à H5PxAPIkatchu pour

effectuer l'enregistrement dans notre base MongoDB est assez succinct. Un fork du dépôt Github H5PxAPIkatchu dans SciViews, nous permet d'inclure le patch.

```
// If fuction h5pxapikatchu_custom_insert_data exists, delegate
// Added in version 0.4.3
if (function_exists('h5pxapikatchu_custom_insert_data')) {
    if ( h5pxapikatchu_custom_insert_data( $actor, $verb, $object, $result, $xapi ) ) {
        wp_die();
    }
}
$ok = Database::insert_data( $actor, $verb, $object, $result, $xapi );
```

Dans `h5pxapikatchu.php`. Ensuite, le dépôt est récupéré au format ZIP, l'ancien plugin est éliminé de Wordpress et cette version est téléversée. Il ne reste plus qu'à ajouter notre fonction personnalisée `h5pxapikatchu_custom_insert_data()` dans "My Custom Function", avec ceci (l'url vers MongoDB Atlas est ici remplacée par) et le tour est joué.

```
function h5pxapikatchu_custom_insert_data( $actor, $verb, $object, $result, $xapi ) {
    $response = false;

    try {
        $manager = new MongoDB\Driver\Manager( 'mongodb://.....' );
        $bulk = new MongoDB\Driver\BulkWrite;
        $doc = [
            'actor' => $actor,
            'verb' => $verb,
            'object' => $object,
            'result' => $result,
            'xapi' => $xapi,
        ];
        $bulk->insert( $doc );
        $result = $manager->executeBulkWrite( 'sdd.h5p', $bulk );
        $response = true;
    } catch (Throwable $e) {
        //echo "Catured Throwable: " . $e->getMessage() . PHP_EOL;
        // Manage error here...
    }
    return $response;
}
```

TODO: Cela fonctionne, mais pour une utilisation plus robuste, il nous faudra passer les paramètres nécessaires à la requête AJAX sur Wordpress de manière dynamique à nos pages bookdown. Nous pourrions utiliser le même mécanisme que celui utilisé pour transmettre les informations de l'utilisateur avec un `localStorage`, ou plutôt, un `sessionStorage` serait mieux.

5.2.2 wp-h5p-xapi et LRS

L'autre plugin Wordpress qui permet d'enregistrer les événements xAPI est `wp-h5p-xapi`. Dans cette configuration, il nous faut un LRS. La plupart des solutions proposées sont

commerciales, payantes et ... chères, sauf une : Watershed. Son utilisation comme LRS est **gratuite** et **sans limitations** ni du nombre d'utilisateurs, ni du nombre d'événements générés par mois. Ça semble presque incroyable, sauf si on comprend que c'est pour eux un produit d'appel vers leurs outils analytics qui effectuent des rapports sur le LRS, et qui sont eux, bien chers (minimum 1600\$/mois) ! Mais nous, du moment qu'on puisse enregistrer nos événements xAPI et les récupérer dans R, on a tout ce qu'il nous faut. Donc, la solution complète envisagée passerait par wp-h5p-xapi et Watershed.

En allant sur leur site, on se rend compte qu'il n'est pas possible de créer un compte directement. Il faut envoyer une demande via un formulaire... et la réponse se fait attendre... indéfiniment ! De plus, autre déconvenue : le plugin Wordpress wp-h5p-xapi n'a pas reçu de commit depuis plus de 2 ans (7 avril 2018) et semble ne plus être conforme aux spécifications de Wordpress (voir ici) de sorte qu'il n'est plus distribué par le site Wordpress plugins depuis le 31 octobre 2019 en raison de violation des lignes directrices de Wordpress... Ce n'est pas très encourageant.

5.2.3 Autres plugins H5P

- Le même auteur de H5PxAPIkatchu propose aussi H5P Resize Pulse qui permet de régler un problème d'affichage apparemment de H5P dans certains containers. Dans les bookdowns, nous ne rencontrons pas de problèmes pour l'instant, mais c'est bon à savoir... Il propose aussi H5P sharing pour obtenir aisément le code `<iframe>...</iframe>` d'un contenu H5P. Nous n'en avons pas besoin car la fonction `learn_down::h5p()` remplit la même fonction dans R.
- H5P CSS Editor adapte le look de l'éditeur H5P à celui du site Wordpress, ... mais ce plugin n'a pas été mis à jour depuis plus de 3 ans. Nous ne l'avons pas testé.
- Gamipress est une solution Open Source et entièrement gratuite de gamification de sites Wordpress (l'utilisateur est encouragé à utiliser le site en gagnant des points, des bagdes, un ranking des utilisateurs peut être affiché, etc.). Ils proposent aussi un plugin qui permet d'obtenir des points avec les activités H5P. Les badges et la gamification de l'apprentissage est une piste très intéressante que nous devrions développer pour stimuler nos étudiants. Donc, nous explorerons cette piste plus tard (voir aussi la section sur les badges dans Moodle 8).

5.3 Gestion des contenus H5P

Chaque contenu H5P est indépendant du reste. Attention que les contenus complexes sont volumineux. Par exemple, les vidéos interactives occupent chacune 2,8Mo sur le serveur lorsque les vidéos sont sur youtube (donc, occupation de place minimale puisque la vidéo elle-même n'est **pas** sur le site Wordpress).

Au départ, aucun contenu n'est réalisable car les **librairies** nécessaires ne sont pas encore installées. Mais il n'est pas nécessaire de les installer car il suffit de téléverser du contenu pour que leur installation se fasse automatiquement. Donc, la procédure pragmatique suivante est retenue pour enrichir le site des librairies dont nous avons besoin :

- Pour un type de contenu pas encore supporté, aller sur h5p.org. Se connecter (créer un compte gratuit, si nécessaire).
- Aller dans "Examples and Downloads" et choisir un exemple de type de contenu souhaité. Cliquer sur "Reuse". Ou alors aller dans "Contents" -> "Create New Content" et créer un contenu du type souhaité. Puis cliquer sur "Reuse" lorsqu'on est satisfait du contenu.

Voici un exemple avec le contenu de type “True/False”.

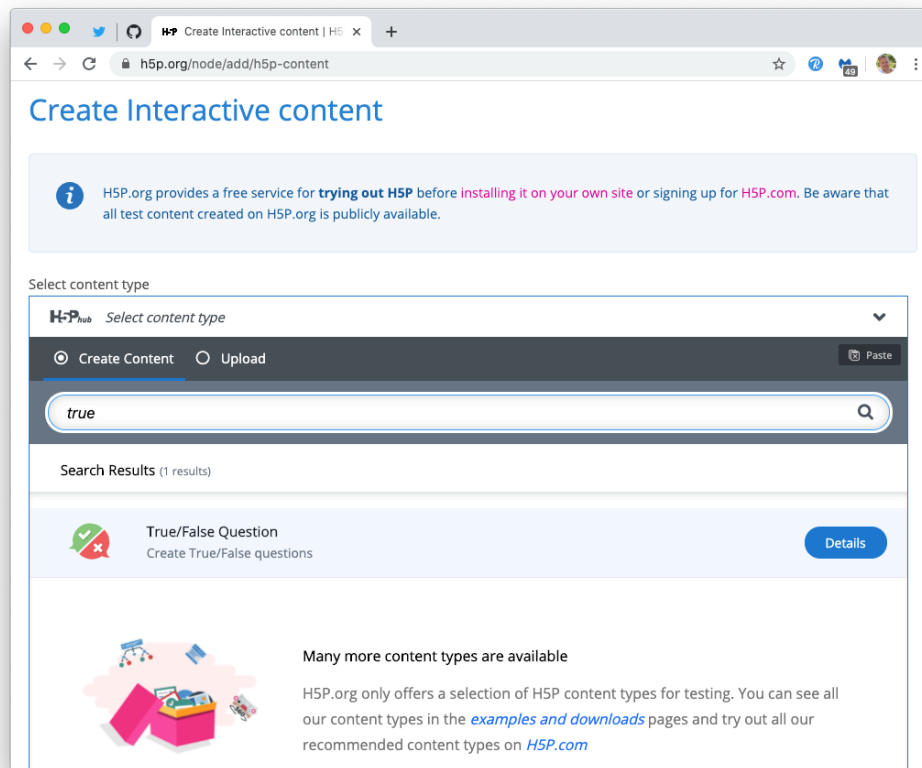


Figure 5.5: Dans son compte sur h5p.org aller dans “Contents” -> “Create New Content”, et choisir le type de contenu que l’on veut créer (ici True/False question).

Ensuite, dans le tableau de bord de Wordpress (pour un utilisateur éditeur), aller dans H5P Content -> Add New. Sélectionner “Téléverser” et indiquer le fichier .h5p que l’on vient de sauver, puis “Utiliser”.

Une fois que c’est fait, non seulement le contenu est téléversé, mais les bibliothèques requises sont également ajoutées. De cette façon, un autre contenu de ce type peut maintenant être créé *directement* depuis la tableau de bord de Wordpress, H5P content -> Add New.

5.3.1 Backup H5P

Le backup du site entier Wordpress est effectué. Par contre, il est utile d’avoir une copie de nos différents contenus H5P au format de fichier .h5p. Pour cela, un administrateur de site doit au préalable temporairement cocher l’option “Toolbar Below Content” dans les paramètres du plugin H5P de Wordpress (tableau de bord), de sorte que le bouton “Reuse” apparaisse en dessus d’un contenu H5P. Pour chaque contenu que l’on souhaite sauvegarder, l’afficher, et cliquer sur son bouton “Reuse” pour récupérer le fichier .h5p. Ensuite, placer les fichiers .h5p sur p-Cloud dans le répertoire `sdd_h5p_backup`.

TODO: une solution de récupération automatisée des fichiers .h5p d’un site devrait être développée à la place, cette procédure ne devrait être utilisée que ponctuellement pour

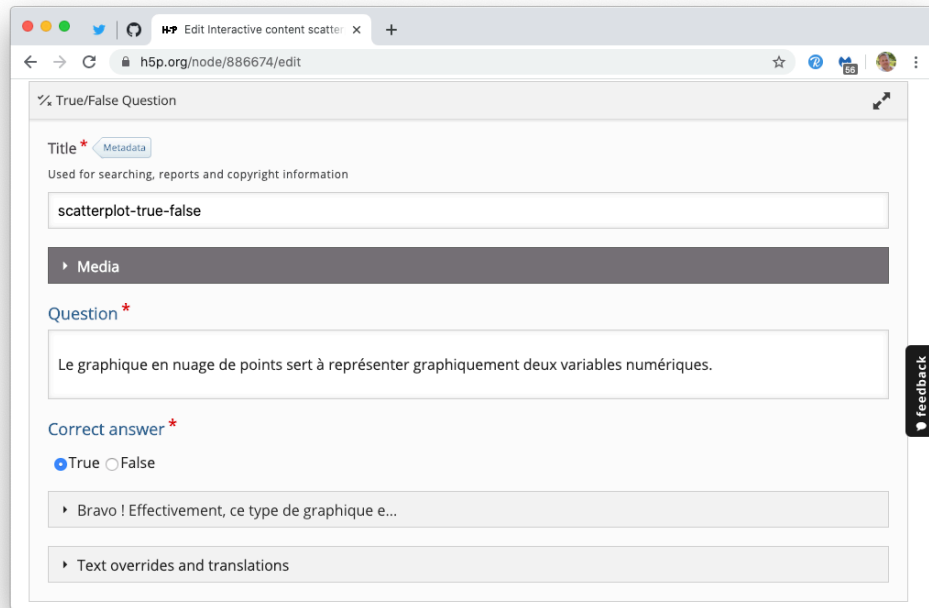


Figure 5.6: Créer le contenu et l'enregistrer).



Figure 5.7: Une fois le contenu terminé, il suffit de cliquer sur “Reuse” (flèche rouge) et de cliquer sur “Download as an .h5p file” pour récupérer le fichier .h5p.

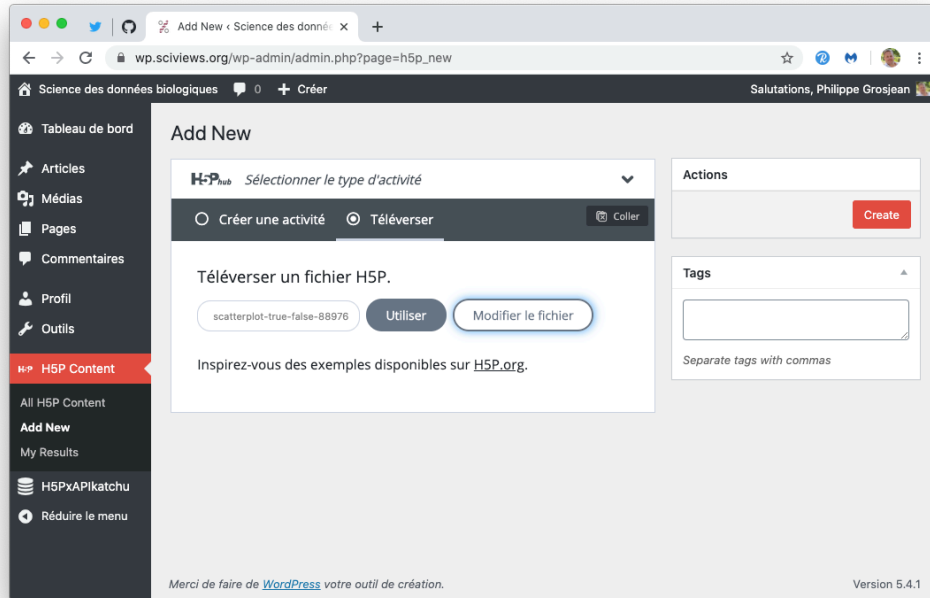


Figure 5.8: Ajout du contenu H5P sur le site Wordpress par téléversement du fichier .h5p.

récupérer un ou un petit nombre de contenus H5P.

5.4 Quelques contenus H5P utiles

H5P propose plus d’une quarantaine de types de contenus différents. Tous ne sont pas utiles dans le contexte des cours de science des données, mais certains le sont particulièrement. Voici une petite sélection.

5.4.1 Apprentissage du code R

Apprendre à coder en R, comme tout apprentissage de langage, nécessite du temps et de la persévérance. L’approche pas à pas, ludique et diversifiée est, en la matière, importante. H5P nous aide notamment avec les contenus suivants :

- **Advanced fill the blanks.** Il est plus facile au début de se focaliser sur des parties du code lorsqu’un exemple est décortiqué. Cependant, l’approche purement passive (l’étudiant suit une explication sur le code) n’est pas très fructueuse. L’étudiant doit être **actif** et forcé à **réfléchir**, tout en l’amenant en douceur à écrire *lui-même* des (parties) de code. “Advanced fill the blanks” est parfait pour cela car il permet un feedback efficace pour aider l’étudiant dans son apprentissage. Ce genre de question pourrait également être envisagée dans une zone de code R d’un learnr, mais c’est plus lourd et à réserver à une phase ultérieure de l’apprentissage.
- S’il faut encore plus faciliter l’exercice, nous pourrions aussi suggérer des fragments de code à replacer au bon endroit, c’est ce que fait Drag the words. C’est en quelque sorte une version ultra-simplifiée de l’exercice de type “remplir les vides” puisque

les solutions sont aussi suggérées sous forme de fragments de code à remplacer au bon endroit.

- Pour varier, nous pourrions aussi proposer des bouts de code et poser ensuite une question simple sur ce code, à répondre par oui ou non, voir True/False question.
- Pour l'apprentissage de la syntaxe R (où sont les variables, les fonctions, les arguments, les chaînes de caractères, etc. dans du code), nous pouvons utiliser Mark the words.

5.4.2 Apprentissage des graphiques

La visualisation des données impose de développer deux compétences en parallèle : (1) la faculté de déterminer quelle est la meilleure représentation graphique pour quelles données, et (2) la capacité à maîtriser techniquement l'outil pour effectivement réaliser le graphique voulu. En plus de cela, il faut aussi que le graphique soit correct (libellé ou unité des axes présents, respect des conventions, pas de surcharges inutiles, etc.)

- Pour montrer les différentes parties importantes d'un graphique ou, par exemple, expliquer les différents éléments d'une boîte de dispersion, l'Image hotspots est idéal. C'est infiniment plus engageant qu'un graphique et une légende statiques.
- Le complément logique pour faire des exercices d'auto-évaluation sur le même thème, c'est le Find multiple hotspots.
- Le Drag and drop propose une alternative intéressante pour varier le type d'exercices consistant à identifier les parties importantes d'un graphique.
- Image juxtaposition permet de comparer deux solutions graphiques différentes d'une manière originale. Il est aussi possible de visualiser une variation sur un thème lié à un même graphique, par exemple, deux palettes de couleurs différentes, ou l'utilisation de la couleur *versus* des symboles. Nous pouvons aussi imaginer avec cet outil comparer un modèle linéaire et un modèle polynomial, par exemple.

5.4.3 Explications générales

- Ici, les Vidéos interactives sont un must. Les étudiants étant immergés dans l'audiovisuel constamment, l'image animée est un canal de communication familier et habituel. Nous devons l'exploiter. Par contre, il est trop facile de regarder passivement une vidéo... et d'avoir l'impression d'avoir compris sans pourtant n'avoir rien assimilé. Parsemer des questions sur ce qui vient d'être visualisé dans la vidéo est un moyen extrêmement puissant pour faire prendre conscience que la lecture est passive. L'étudiant peut alors revenir en arrière et visionner plus attentivement dans le but de pouvoir répondre à la question. Ce faisant, il apprend, comprend et assimile bien mieux le contenu de la vidéo !
- Il ne faut pas négliger non plus le Course presentation. Si les explications deviennent longues et les pages de texte interminables, ce n'est pas très engageant pour l'étudiant. La matière sous forme d'une présentation où des questions d'auto-évaluation sont directement incluses est un moyen très efficace de diversifier, casser la monotonie d'une longue explication textuelle, et plus engageant si la présentation est richement illustrée. Pour un exemple concret, si nous expliquons comment réaliser un test t de Student et en même temps que nous donnons les détails de calcul, cela fait trop. De plus, de part notre expérience, l'étudiant a ensuite du mal à faire la part des choses entre l'application du test dans R et les détails plus techniques du "comment ça marche en interne". Le texte devient aussi très (trop) long.

Une solution est alors de se focaliser sur la résolution du test dans R uniquement sous forme de développement textuel, et réserver une présentation “comment le t test est calculé” pour les détails, et peut-être une autre présentation intitulée “les différents t tests” pour expliquer les différentes variantes.

5.4.4 Autoévaluation

Les learnrs sont des outils de choix que nous appliquons avec succès pour l’autoévaluation des étudiants. Cependant, une autoévaluation plus rapide et plus proche du matériel pédagogique lui-même est encore plus efficace en première ligne (suivi en seconde ligne par des learnrs avec un niveau de difficulté supérieur), et enfin en troisième ligne, les exercices sur données réelles dans un projet RStudio *via* Github Classroom. Nous venons de voir que H5P permet cette autoévaluation directement à l’intérieur des vidéos et des présentations. Il nous manque l’équivalent pour les développements textuels. Les contenus suivants fournissent ce dont nous avons besoin à ce niveau.

- Quizzes permet de rassembler plusieurs questions successives en un contenu unique avec une note finale pour l’ensemble du genre 4/6 questions répondues correctement.
- Pour chaque question du quizz, voire des questions isolées parsemées dans le texte directement, nous avons différents contenus possibles : vrai/faux, choix unique ou Multiple choice.
- Pour une variation plus ludique des questions d’auto-évaluation, nous pouvons envisager utiliser les Flashcards, surtout pour les graphiques.

5.4.5 “Take home message”

Dans une capsule d’apprentissage certains éléments sont cruciaux. Bien définir les objectifs et les prérequis au début de la capsule sont vitaux. En fin de capsule, un résumé des points importants à retenir, le fameux “take home message” est ultra-important également... encore faut-il que l’étudiant assimile et mémorise ce/ces message(s). H5P propose du contenu qui permet d’appréhender ce résumé de manière plus active et ludique.

- Summary se *doit* pratiquement d’être utilisé systématiquement en fin de chaque capsule d’apprentissage pour résumer les points importants. Avec ce contenu H5P, c’est l’étudiant qui établit sa liste de messages importants de manière **active**.

Chapter 6

Bookdown

...

TODO: trouver un moyen simple de mettre à jour les pages statiques book-down des cours dans Wordpress (ainsi que les données pour les rapports de progression des étudiants dans `/users`).

6.1 Intégration dans Wordpress

Pour l’instant, il y a deux versions :

- Les pages affichées telles quelles. Nous avons une présentation bookdown classique. Les informations de l’identité de l’étudiant sont transmises via `localStorage`, mais le bandeau supérieur du site et la barre noire supérieure de gestion Wordpress ne sont pas affichés. Le bookdown n’est pas vraiment intégré dans le site. Cette représentation est utile pour l’utilisation du bookdown dans un contexte hors cours UMONS.
- Les mêmes pages du bookdown, **mais intégrées sous forme d’iframe** dans une page de Wordpress dédiée. Nous utilisons le plugin *Advanced iFrame Pro* qui permet de contrôler la taille de l’iframe, de lui passer des paramètres et encore d’autres choses. Cela donne pas mal, et l’utilisateur conserve la barre d’outils supérieure (accès à Moodle, Discord, etc.) ainsi que la barre noire supérieure de contrôle Wordpress qui permet aussi de vérifier d’un coup d’œil que l’on est connecté, et sous quelle identité nous le sommes. S’il veut lire le contenu du bookdown sans voir la bannière haute, le lecteur peut faire coulisser la page jusque tout en bas et cette bannière viendra se cacher derrière la barre noire supérieure Wordpress (qui, elle, reste toujours visible tant que l’utilisateur est connecté au site).

6.2 Contenu H5P dans bookdown

Nous avons déjà vu que le contenu H5P s’affiche dans un iframe dans un document principal. Il lui transmet les événements xAPI, et c’est donc ce document principal qui doit les traiter. Nous avons ici une impossibilité à le faire si nous sommes dans un cas où H5P est servi par un serveur différent que notre bookdown. En effet, les accès en “Cross-Origin Resource Sharing” ou CORS sont restreints pour éviter des attaques de site évidentes. Ainsi un iframe issu d’un serveur différent de la page qui l’héberge ne peut

communiquer librement avec la page qui le contient si elle provient d'un site différent. Or, nous avons besoin de cette communication pour transmettre les événements xAPI...

La solution que nous avons trouvée est d'ajouter le dossier qui contient nos pages statiques du bookdown à Wordpress directement. Ainsi les deux sont maintenant servis à partir de la même adresse <https://wp.sciviews.org>. Pour l'instant, nous copions simplement ces fichiers dans un sous dossier `sdd-umons` dans `lamp0/web/vhosts/wp.sciviews.org/htdocs`. Ainsi, le bookdown en question est servi par Wordpress à l'adresse <https://wp.sciviews.org/sdd-umons>. Et à présent, la communication des xAPI H5P est possible... à condition d'ajouter le code nécessaire dans bookdown, mais cela, nous l'avons expliqué dans le chapitre consacré à H5P (5).

Il nous faut maintenant inclure facilement les contenus H5P depuis Wordpress dans nos pages bookdown. Ces contenus sont identifiés par un numéro (ID) dans Wordpress. Cet ID apparaît clairement dans la liste des contenus H5P dans le tableau de bord. Il suffit de noter le numéro ID correspondant au contenu que l'on souhaite inclure.

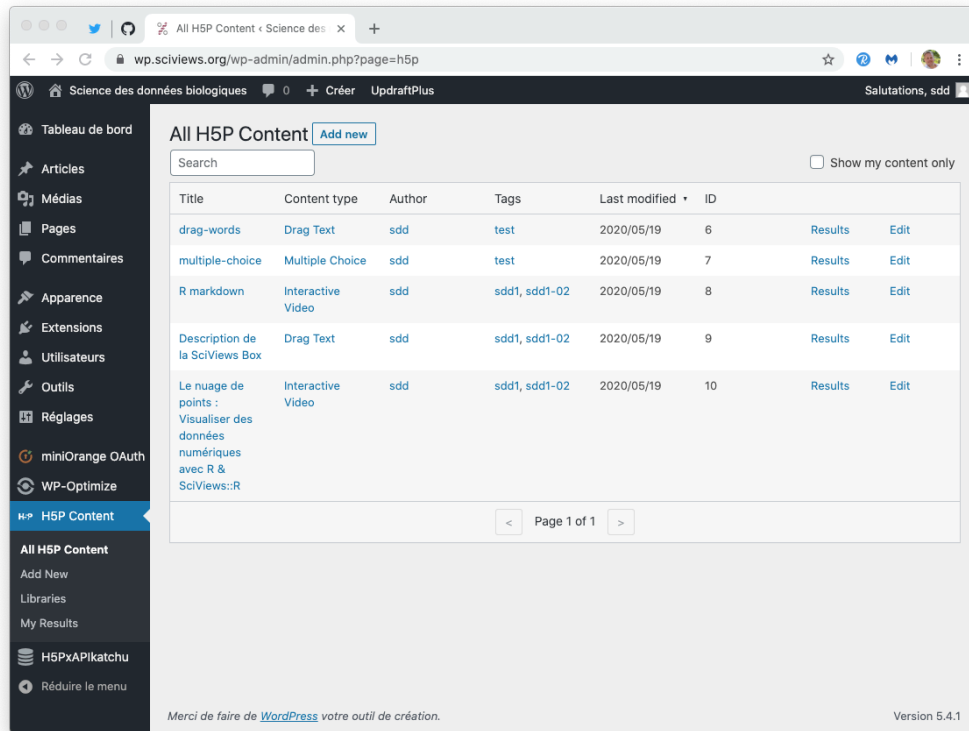


Figure 6.1: H5P Content dans Wordpress. Les widgets sont identifiés par leur ID (ici, 6, 7, 8, 9 & 10).

Nous devons créer un iframe dans la page bookdown, mais nous pouvons déléguer ce travail à une fonction R. A terme, cette fonction se trouvera dans le package `learndown`, mais en voici le code :

```
h5p <- function(id, width = 780, height = 500) {
  glue::glue("\n<iframe src=\"https://wp.sciviews.org/wp-admin/admin-ajax.php?action=h5p_embed&")
```



```
}
```

A partir d'ici, ajouter un contenu H5P est aussi simple que d'ajouter `h5p(id)` dans du code R en ligne où "id" est l'ID que nous avons noté. On peut aussi jouer sur les arguments `width=` et `height=` pour ajuster l'emplacement et la taille du contenu H5P dans la page.

Ainsi programmé notre contenu H5P se comporte comme suit :

- pour un utilisateur enregistré dans Wordpress, les événements xAPI sont enregistrés à son nom. Nous pouvons donc suivre la progression de nos étudiants via les contenus H5P parsemés dans le bookdown également, en plus des applis Shiny et des learnrs. De plus, l'état d'achèvement de l'exercice est aussi enregistré pour un bon nombre de contenus H5P et rétabli lorsque la page est rechargée par le même utilisateur.
- pour un utilisateur non enregistré, les contenus H5P fonctionnent normalement, mais aucun événement xAPI n'est enregistré, et l'état de complétude de l'exercice n'est pas conservé non plus. Si la page est rechargée, les contenus H5P retournent à leur état initial à chaque fois.

Ceci est un inconfort mineur pour un utilisateur anonyme et permet l'utilisation des contenus H5P de manière optimale pour nos étudiants dûment enregistrés, tout en permettant une libre utilisation par d'autres visiteurs sans polluer notre base de données.

Nous avons également tenté de passer des arguments supplémentaires via l'URL qui lance la page bookdown, et qui les transfèrent vers l'iframe H5P, mais sans succès car apparemment Wordpress réécrit l'URL et laisse tomber ces arguments supplémentaires. Mais de toutes façons, ce n'est pas important pour l'instant car nous pouvons tracer nos utilisateurs enregistrés dans Wordpress telle que c'est actuellement implémenté.

Chapter 7

Shiny

...

Chapter 8

Moodle

...

8.1 Gestion des cours

Un espace différent a été créé pour chaque AA (donc, deux pour SDD I et SDD II) par l'administrateur Moodle. Voici la liste des AA, l'intitulé et le code Moodle correspondant (qui ne correspond **pas** au code dans le programme des cours) :

Cours	Code Moodle	Intitulé
SDD I Q1	S-BIOG-006	Science des données I : visualisation
SDD I Q2	S-BIOG-027	Science des données I : inférence
SDD II Q1	S-BIOG-015	Science des données II : modélisation
SDD II Q2	S-BIOG-061	Science des données II : analyse
SDD III Q1	S-BIOG-025	Science des données III : exploration et prédiction
SDD IV all	S-BIOG-043	Science des données IV : pratique
SDD V Q1	S-BIOG-077	Science des données V : recherche reproductible
SDD I Charleroi Q1	S-BIOG-921	Bio-informatique et science des données
SDD II Charleroi Q2	??????????	Bio-informatique et science des données II

8.2 Badges

Les badges sont très utiles car ils matérialisent la progression des étudiants. C'est à la fois une façon ludique d'apprendre, et un moyen efficace de pousser un étudiant à travailler. Les badges sont générés dans Moodle, mais bien sûr, ils sont liés à des activités Moodle. Avec nos applications externes, nous ne pouvons fournir des badges que manuellement, ce qui devient très vite laborieux. TODO: trouver un moyen d'automatiser l'octroi de badges. Moodle peut se connecter à <https://badgr.io>, et nous pouvons peut-être travailler via son interface ? A creuser !

Pour créer des badges, nous pouvons utiliser une application en ligne. Il suffit d'indiquer son adresse email, et de concevoir son badge. La même interface peut d'ailleurs être utilisée pour concevoir visuellement des hex-stickers (à faible résolution toutefois puisque c'est des fichiers .png de 200x200 pixels). Voici un exemple de badge que nous pourrions utiliser dans nos cours :



Figure 8.1: Exemple de badge

Les images correspondant à nos différents badges sont sur p-Cloud, dans le sous-dossier `sdd_badges`.

8.3 Notes

TODO: il est possible d'envoyer les notes à Moodle via une URL qui renvoie un fichier XML... Explorer cette piste pour créer et mettre à jour les notes Moodle depuis nos applis...

Chapter 9

Vidéos

...

Chapter 10

Rapport de progression

Un des gros problèmes des cours en classes inversées est d'arriver à **stimuler** les étudiants à faire le travail chez eux. Livrés à eux-mêmes, seulement un pourcentage trop faible d'entre eux va réaliser le travail demandé à domicile. Évidemment, c'est également préjudiciable au travail en classe, si ce dernier part du principe que la matière hors présentiel est censée être vue. Et si des travaux de groupe (conseillé) sont prévus, ceux qui n'ont rien fait apparaîtront comme des boulets pour les autres.



En parallèle au déploiement du matériel d'e-learning hors présentiel, nous devons mettre également en place des outils pour permettre à l'étudiant de suivre sa progression d'apprentissage, et aussi pour le stimuler dans cet auto-apprentissage.

Les outils qui peuvent être mis en place sont :

- *Pénalisation si le travail n'est pas fait.* Il faut arriver à déterminer ici objectivement si le travail est fait ou pas. Le plus simple est d'interroger les étudiants en début de chaque classe en présentiel, **mais cela prend du temps et n'est pas très constructif** ! Cette approche n'est donc pas conseillée.

La vidéo suivante explique d'ailleurs pourquoi le coercitif ne fonctionne pas bien et pourquoi il faut plutôt **valoriser le travail effectué** plutôt que de punir s'il n'est pas fait. Il faut aussi **une récompense immédiate**, et favoriser l'**émulation individuelle** (se dépasser soi-même), et **de groupe** (se comparer à la progression générale de la classe).

Les aides et encouragements au travail que nous avons mis en place de manière utile sont :

- L'*émulation personnelle* au travers de point reçu si l'on fait tous les exercices (les badges, et la "gamification" en général de l'apprentissage sont également à considérer sérieusement dans ce contexte),
- La *valorisation d'un comportement ou d'un résultat positif au delà de la moyenne* à l'aide de points bonus (on est toujours dans la gratification et la gamification ici),
- l'*entre-aide*, faire expliquer les points difficiles par un collègue qui les ont compris est extrêmement efficace. En plus, cela permet de mitiger un aspect difficile à gérer : la différence de vitesse d'apprentissage entre les "bons" et les "mauvais" étudiants (non, il n'y a pas de bons et de mauvais étudiants, nous le savons, il y a juste des étudiants avec des facultés et des motivations différentes... mais certains sont quand même beaucoup plus lents et **énervants** que d'autres). Au lieu de se tourner

les pouces, demander à ceux qui ont fini en avance d'expliquer aux autres permet d'utiliser efficacement le temps et les ressources humaine en classe.

- Les *travaux de groupes*, les *challenges* (par exemple, les étudiants doivent reproduire un graphique difficile), les *compétitions* (une course contre le temps et/ou entre plusieurs groupes pour résoudre un problème) participent efficacement à dynamiser le travail en classe et à varier les activités.
- L'utilisation des *réseaux sociaux* pour communiquer. Le forum de Moodle est peu ou pas utilisé du tout. Les issues de Github, une fois que les étudiants ont appris à les utiliser, s'avèrent utiles car elles sont disponibles près du dépôt lui-même qui contient les exercices. Mais ce qui fonctionne le mieux (testé en période de confinement Covid-19), c'est *Discord* parce que beaucoup d'étudiants connaissent et utilisent déjà pour les jeux. Le ton des échanges y est d'ailleurs plus décontracté (parfois nettement plus), mais cela reste un outil efficace pour aider un étudiant qui est bloqué dans sa progression.
- Les étudiants sollicitant une *aide plus directe* apprécie de pouvoir poser ses questions par email. Afin de pouvoir être plus réactif, nous avons mis en place une adresse `sdd@sciviews.org` pour laquelle tous les enseignants du cours ont accès. Ainsi, les interventions peuvent être partagées et plus rapides.
- Pour obtenir un *retour plus direct* par rapport aux activités, nous mettons progressivement en place le **rapport de progression**. Les actions de l'étudiant dans les learnrs (et dans un second temps également H5P et les apps Shiny) sont enregistrées dans une base de données. Le rapport de progression est une app Shiny qui collecte, nettoie, prépare, et présente graphiquement la progression de l'étudiant. L'*émulation de groupe* est obtenue en comparant la progression de l'étudiant à la progression générale de la classe.

Afin d'avoir un retour immédiat, nous souhaitons rendre ce rapport de progression dynamique de sorte qu'il montre l'évolution en temps réel. Il est en effet important que l'étudiant puisse visualiser sur son rapport qu'il vient d'avancer en achevant à la seconde un exercice ! De plus, nous souhaitons associer cela à des récompenses sous forme de **badges** lorsque l'étudiant franchit une étape dans son apprentissage (aussi, des badges spéciaux pourraient être associés aux points bonus). La gestion des badges a été abordée dans la section dédiée à Moodle (voir 8).

L'app Shiny de rapport de progression est développée dans le dépôt Github privé `sdd_progress-report`. Comme il contient du code sensible, il n'est pas disponible de manière publique. Par contre, les enseignants du cours y ont accès. Le fichier `README.md` du dépôt détaille son fonctionnement et contient du code pour le tester.

Cette app Shiny est ensuite intégrée dans Moodle via un lien URL qui ouvre l'app dans une fenêtre interne et lui passe des paramètres issus de Moodle (identité de l'étudiant, cours correspondant) afin de déterminer quel rapport il faut présenter.

Appendix A

Installation

TODO: installation des outils logiciels nécessaires pour nos cours...

Références