



Clustering & Analyse du parcours étudiant

Nirina Andriamanantenasoa

Ingénieur pédagogique
CAPSULE

Delphine Ruggeri

Ingénieure pédagogique
CAPSULE

CAPSULE : Centre d'Appui pour la Pédagogie et SUpport à L'Expérimentation
Faculté des Sciences et Ingénierie





Au départ : demande d'une équipe enseignante pour améliorer un cours jugé insatisfaisant



- cours hybridé sur les Outils en Programmation Informatique (OPI), Polytech Paris
- 55 étudiants de 1ère année (équivalent L3)
- 80 % distanciel sur Moodle ; 20 % présentiel (TP restructuration + TP évaluation)
- 2 phases avec la même structure

Calendrier

Septembre	Octobre	Novembre	Décembre	Janvier
1 J	1 S	1 M	1 J	1 D
2 V	2 D	2 M	2 V	2 L
3 S	3 L	3 J	3 S	3 M
4 D	4 M	4 V	4 D	4 M
5 L	5 M	5 L	5 J	5 V
6 M	6 J	6 D	6 M	6 L
7 M	7 L	7 M	7 V	7 S
8 J	8 S	8 M	8 J	8 D
9 V	9 D	9 M	9 V	9 L
10 S	10 L	10 J	10 S	10 M
11 D	11 M	11 V	11 D	11 M
12 L	12 M	12 S	12 L	12 J
13 M	13 J	13 D	13 M	13 V
14 M	14 V	14 L	14 S	14 D
15 J	15 S	15 M	15 D	15 L
16 V	16 D	16 V	16 L	16 M
17 S	17 L	17 J	17 S	17 M
18 D	18 M	18 D	18 V	18 J
19 L	19 M	19 S	19 L	19 J
20 M	20 J	20 D	20 M	20 V
21 M	21 L	21 M	21 S	21 D
22 J	22 S	22 M	22 D	22 S
23 V	23 D	23 M	23 L	23 L
24 S	24 L	24 J	24 M	24 M
25 D	25 M	25 V	25 D	25 M
26 L	26 M	26 L	26 J	26 J
27 M	27 J	27 D	27 M	27 V
28 M	28 V	28 L	28 M	28 S
29 J	29 S	29 M	29 J	29 D
30 V	30 D	30 M	30 V	30 L
31 L	31 M	31 S	31 M	31 M

Handwritten notes on the calendar: 'PHASE 1' in purple (Oct 20-21), 'PHASE 2' in green (Nov 14-15), 'VACANCES' in yellow (Nov 3-4, Dec 22-23), 'eval 1' in pink (Nov 8), 'eval 2' in green (Dec 6), 'Tp 1' in pink (Oct 27), 'Tp 2' in green (Nov 12).

Phase 1 : La prise en main

- Utiliser un terminal
- Utiliser un gestionnaire de version git
- Choisir un éditeur de texte

Phase 2 : Les Makefile et le debug

- Concevoir et utiliser un Makefile
- Utilisez un debugger
- Découvrir de nouveaux outils

Pour chaque phase :

- 1 Une partie sur moodle en auto-formation
- 2 Une séance de 2h avec enseignants
- 3 Une séance d'évaluation en temps limité

Evaluation

- Avancement et réussite dans moodle (20%)
- Devoir Maison (20%)
- Résultats des 2 TP d'évaluations (60%)



Problèmes évoqués

- Les étudiants arrivent au cours présentiel sans avoir fait la partie à distance.
- Ils se plaignent de ne pas tout comprendre.
- Le forum d'échange sur Moodle n'est jamais utilisé.



L'équipe projet

CAPSULE : équipe pluridisciplinaire d'appui à la pédagogie, en lien avec la recherche
➡ enseignants chercheurs, membres de sa direction et en charge de son pilotage.

Ingénieurs pédagogiques



Gestion de projet, approche qualitative, synthèse, préconisations pédagogiques

Informaticiens, administrateurs de Moodle



Collecte des traces

Data analyst



Analyse des traces

Chercheurs



Choix du protocole d'analyse des traces



Choix méthodologique

Définir le protocole

Quelles traces?
RGPD?
Fiabilisation des données ?
Sript ?

1 Analyse des traces sur Moodle

Protocole évolutif...

3

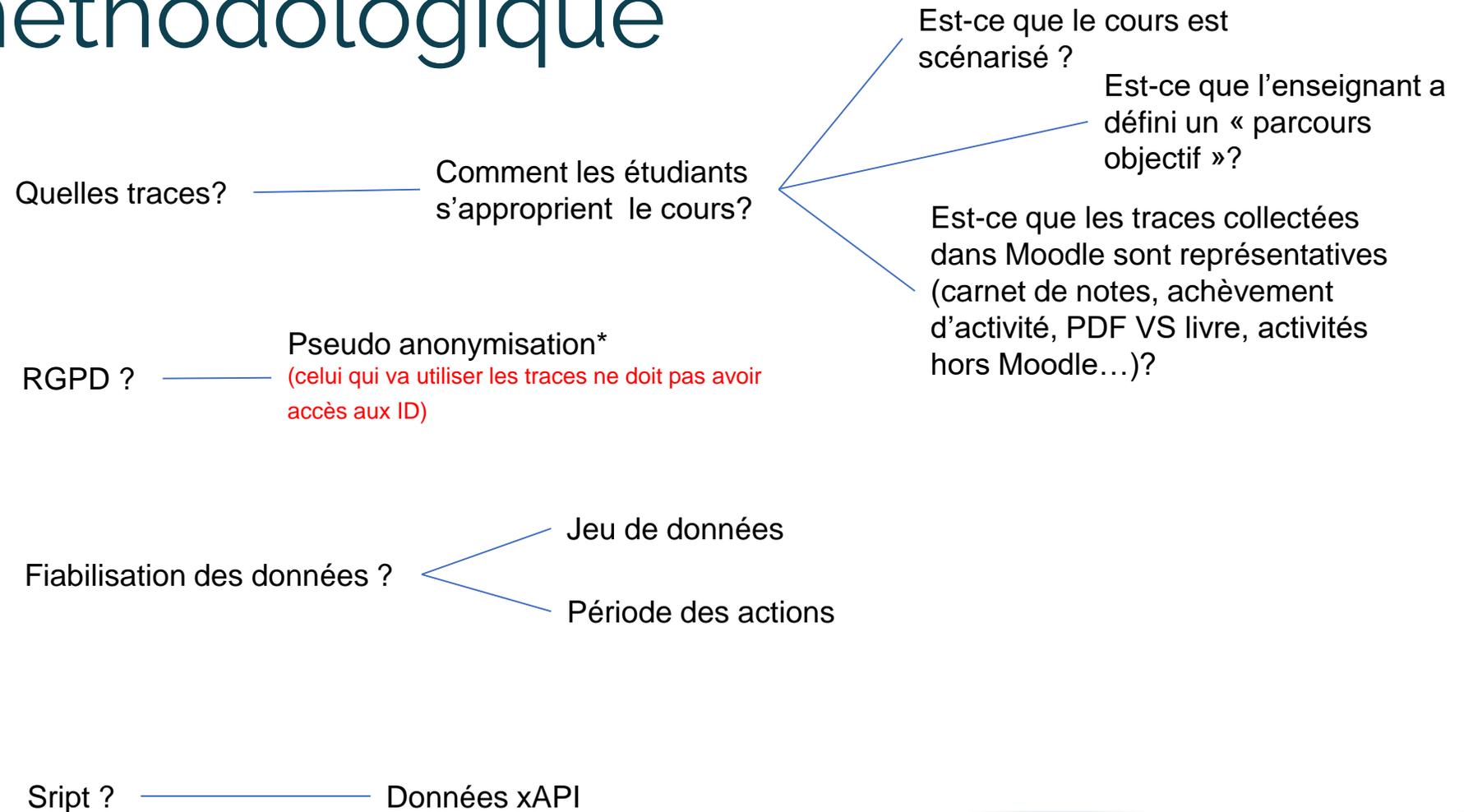
Entretiens qualitatifs

Questionnaires

2



Choix méthodologique



1

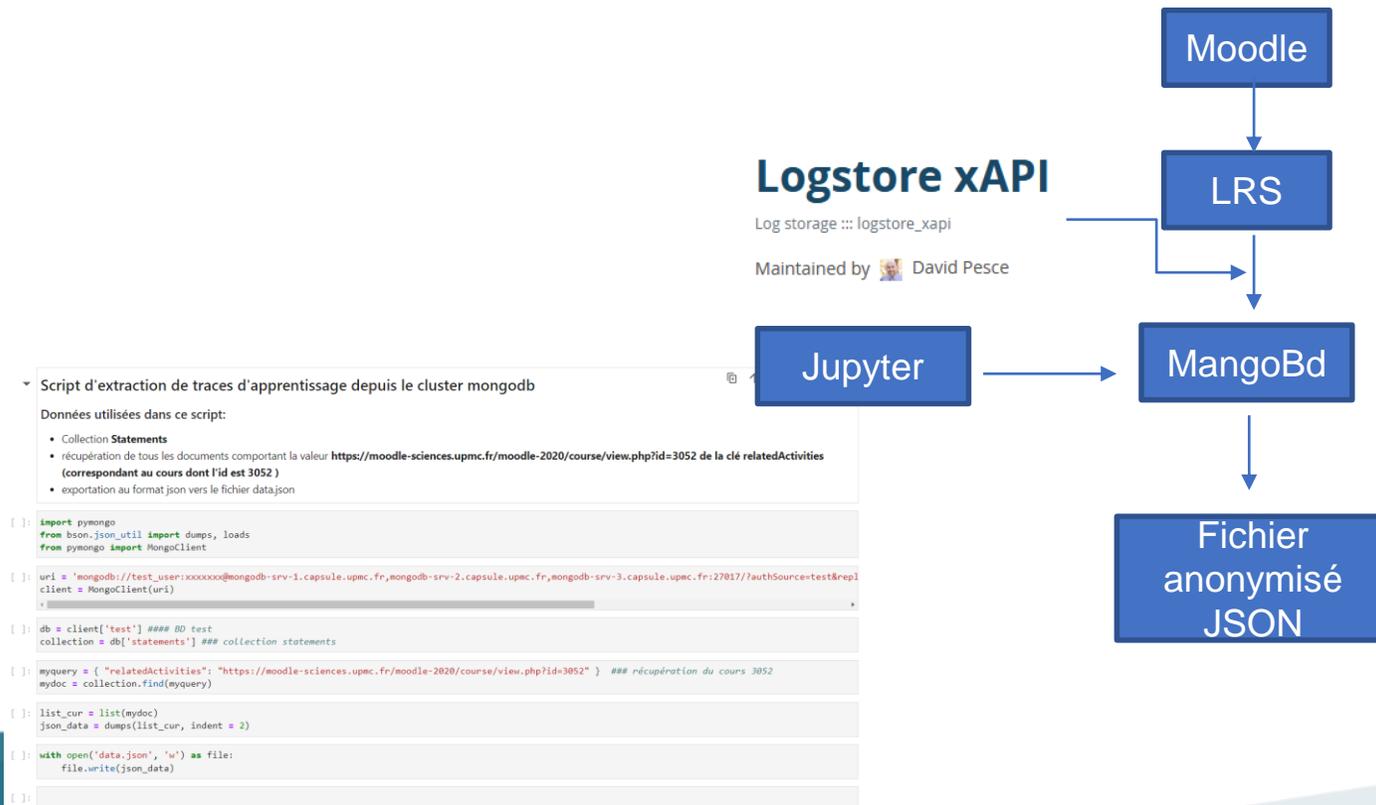
Analyse des traces sur Moodle



Choix méthodologique

Pseudo anonymisation*

(celui qui va utiliser les traces ne doit pas avoir accès aux ID)



Documents

LRS_moodle2024.statements

Documents Aggregations Schema Explain Plan

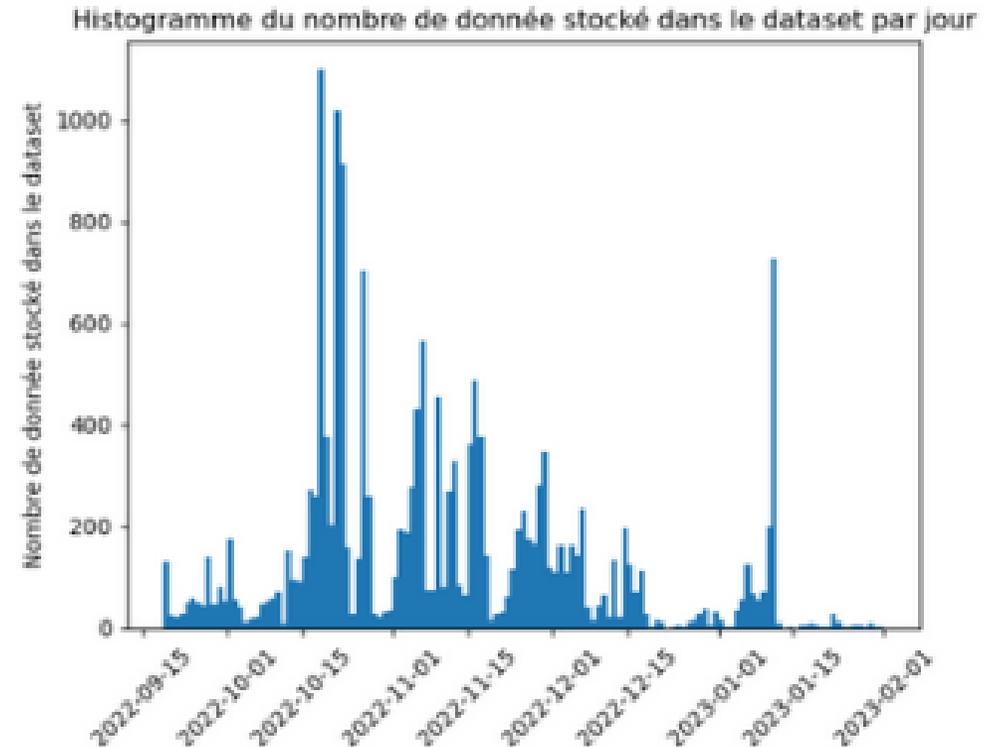
ADD DATA VIEW

```
{
  "_id": ObjectId("666aeefa70b271102edcac30"),
  "hasGeneratedId": true,
  "organisation": ObjectId("666852b0cd7b040a9caf98b2"),
  "Irs_id": ObjectId("66685f0e839230857c5a52d"),
  "client": ObjectId("66685f0e839230857c5a52e"),
  "person": null,
  "active": true,
  "voided": false,
  "timestamp": 2024-05-16T08:38:38.000+00:00,
  "stored": 2024-06-13T13:07:05.730+00:00,
  "hash": "feda9e2ff6c4e8f47053408c6882e3537566359",
  "agents": Array,
  "relatedAgents": Array,
  "registrations": Array,
  "verbs": Array,
  "activities": Array,
  "relatedActivities": Array,
  "statement": Object,
  "metadata": Object
}
```



Analyse de traces Moodle

- Nombre d'étudiants dans le jeu de données : 55
- Période des actions
 - Action la plus récente : 2023-02-01
 - Action la plus ancienne : 2022-09-19





Analyse de traces Moodle

Analyse des consultations des cours et TP :

1ères consultations des pages de Cours de **la phase 1**

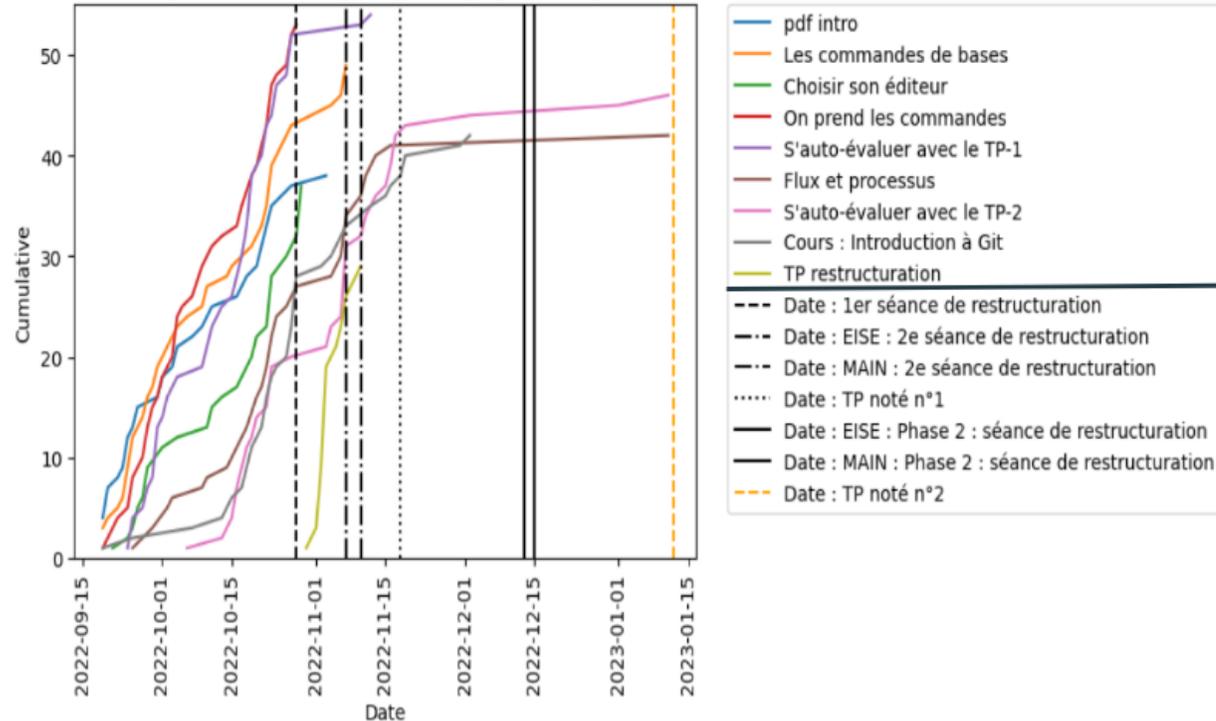
 pdf intro

 Forum

Ressources

 Les commandes de bases

Quantité cumulée des premières consultations des pages de la phase 1 au fil du temps



Remarques

- Page peu consultée : “On prend les commandes” et “Choisir son éditeur”
- Une perte de dynamisme au cours du temps
 - Diminution du nombre d’étudiant qui ont fait le TP-2 par rapport au TP-1



Analyse de traces Moodle

Analyse des consultations des cours et TP :
1ères consultations des pages de Cours de **la phase 2**

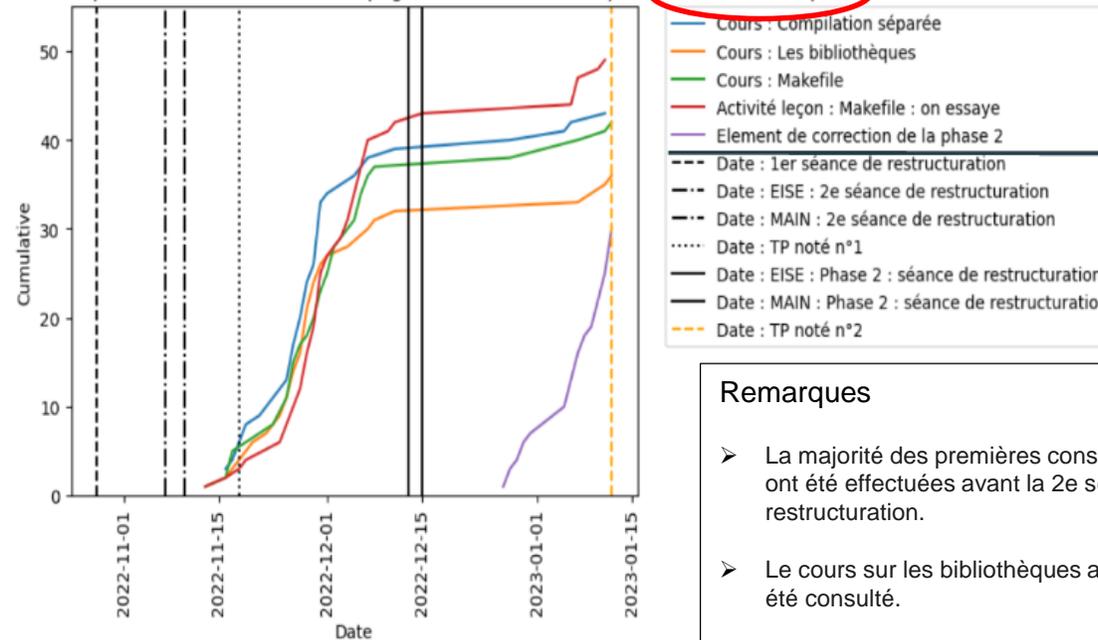
Compilation séparée

-  Cours : Compilation séparée
-  Cours : Les bibliothèques
-  TP : Compilation séparée

Makefile

-  Cours : Makefile
-  Cours : Makefile : on essaye
-  TP : Comprendre un Makefile
-  TP : Écrire un Makefile

Quantité cumulée des premières consultations des pages des 'COURS' de la phase 2 au fil du temps



Remarques

- La majorité des premières consultations ont été effectuées avant la 2e séance de restructuration.
- Le cours sur les bibliothèques a très peu été consulté.
- L'élément de correction de la phase 2 a été consulté par 55% des étudiants.



Analyse de traces Moodle

Analyse des consultations des cours et TP :

1ères consultations des pages de Cours de **la phase 2**

Compilation séparée

 Cours : Compilation séparée

 Cours : Les bibliothèques

 TP : Compilation séparée

Makefile

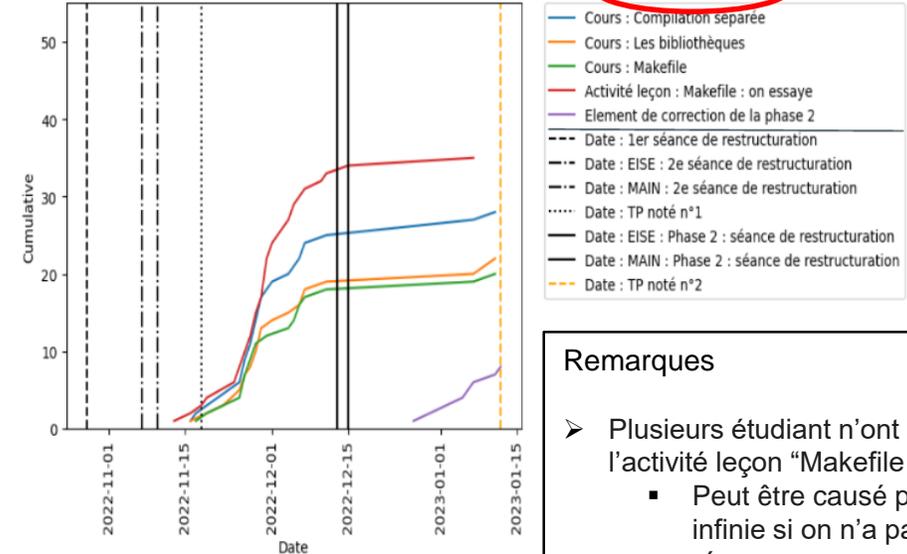
 Cours : Makefile

 Cours : Makefile : on essaye

 TP : Comprendre un Makefile

 TP : Écrire un Makefile

Quantité cumulée des premières consultations des pages des 'COURS' de la phase 2 au fil du temps [cours terminé]



Remarques

- Plusieurs étudiant n'ont pas terminé l'activité leçon "Makefile : On essaye".
 - Peut être causé par une boucle infinie si on n'a pas la bonne réponse.
- Pour les autres cours, il se peut que les étudiants aient fini le cours mais n'ont pas marqué celui-ci comme terminé.



Analyse de traces Moodle

Clustering & Analyse du parcours des étudiants

Clustering : méthode d'analyse de données pour regrouper les apprenants représentant des comportements similaires

Démarche :

- Pour chaque étudiant, on va récupérer son parcours
- Un parcours est constitué de l'ensemble des pages consulté par l'étudiant dans l'ordre de première consultation.

Exemple :

Thomas a consulté les pages du cours dans l'ordre suivant : Son chemin : Pdf Intro ⇒ Cours 1 ⇒ Cours 2 ⇒ Cours 1 ⇒ Activité 1

Son parcours : ["Pdf Intro", "Cours 1", "Cours 2", "Activité 1"]



Analyse de traces Moodle

Clustering & Analyse du parcours des étudiants

Outils et méthodes utilisés :

- Algorithme K-means
- La distance de Levenshtein

La distance de Levenshtein représente le nombre minimum d'opérations nécessaires pour transformer une chaîne de caractères en une autre. Ces opérations comprennent l'insertion, la suppression et la substitution d'un caractère.

Dans notre cas, on va comparer *le parcours des étudiants* avec le *parcours objectif* en utilisant la distance de Levenshtein.

- Si la distance de Levenshtein est faible, cela signifie que le parcours de l'étudiant est assez similaire au parcours objectif.
- Si la distance de Levenshtein est élevée, cela signifie que le parcours de l'étudiant est très différent du parcours objectif.



Analyse de traces Moodle

Clustering & Analyse du parcours des étudiants

Analyse des clusters

Score		Cluster 0 (7 élèves)	Cluster 1 (11 élèves)	Cluster 2 (13 élèves)	Cluster 3 (2 élèves)	Cluster 4 (11 élèves)	Cluster 5 (11 élèves)
Proportion de cours vu	Moyenne (Écart type)	1	0.96 (0.07)	0.78 (0.17)	0.28 (0.4)	0.66 (0.24)	0.75 (0.21)
Proportion de tp autonomi vu	Moyenne (Écart type)	1	0.97 (0.07)	0.91 (0.14)	0.58 (0.58)	0.71 (0.21)	0.84 (0.13)
Proportion total de page vu	Moyenne (Écart type)	0.96 (0.04)	0.94 (0.05)	0.79 (0.09)	0.41 (0.51)	0.60 (0.17)	0.75 (0.14)
Distance de Levenshtein	Moyenne (Écart type)	12 (1.3)	9.54 (1.7)	11.38 (1.33)	14.5 (2.5)	12.6 (1.56)	9.73 (1.48)
Note TP	Moyenne (Écart type)	72.5 (27.7)	62 (27.7)	49.53 (31.6)	26.88 (26.88)	37.27 (41.1)	27.1 (31)
Note Cours	Moyenne (Écart type)	83.7 (6.9)	82.8 (6.7)	75.73 (9.1)	84.1 (2.1)	72.3 (10.2)	77.5 (9)
Note finale	Moyenne (Écart type)	78.1 (14.77)	72.4 (14.58)	62.63 (17.05)	55.47 (12.38)	54.8 (21.7)	52.3 (17.14)



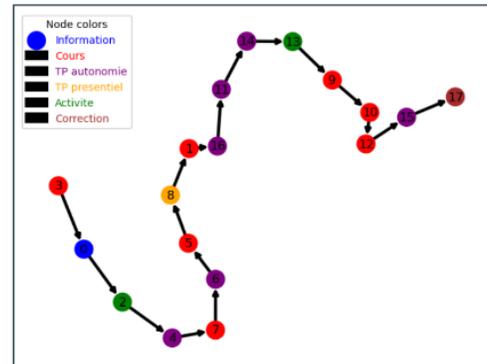
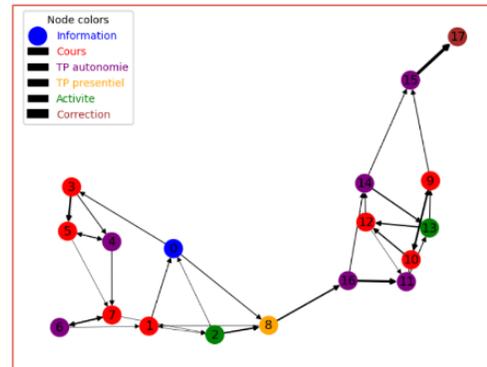
Analyse de traces Moodle

Clustering & Analyse du parcours des étudiants

Cluster 0

Parcours des étudiants d'un point de vue probabiliste

- 0 ⇒ "PDF intro"
- 1 ⇒ "Les commandes de bases"
- 2 ⇒ "Choisir son éditeur"
- 3 ⇒ "On prend les commandes"
- 4 ⇒ "S'auto-évaluer avec le TP-1"
- 5 ⇒ "Flux et processus"
- 6 ⇒ "S'auto-évaluer avec le TP-2"
- 7 ⇒ "Cours : Introduction à Git"
- 8 ⇒ "TP restructuration"
- 9 ⇒ "Cours : Compilation séparée"
- 10 ⇒ "Cours : Les bibliothèques"
- 11 ⇒ "TP : Compilation séparée"
- 12 ⇒ "Cours : Makefile"
- 13 ⇒ "Cours : Makefile : on essaye"
- 14 ⇒ "TP : Comprendre un Makefile"
- 15 ⇒ "TP : Écrire un Makefile"
- 16 ⇒ "TP : Utiliser un débbuger"
- 17 ⇒ "Élément de correction de la phase 2"



Score		Cluster 0 (7 élèves)
Proportion de cours vu	Moyenne (Ecart type)	1
Proportion de tp autonomi vu	Moyenne (Ecart type)	1
Proportion total de page vu	Moyenne (Ecart type)	0.96 (0.04)
Distance de Levenshtein	Moyenne	12 (1.3)
Note finale	Moyenne (Ecart type)	78.1 (14.77)



Analyse de traces Moodle

Clustering & Analyse du parcours des étudiants

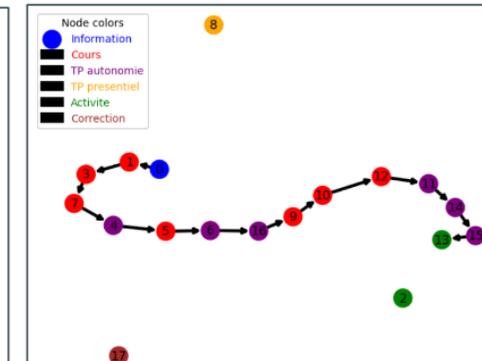
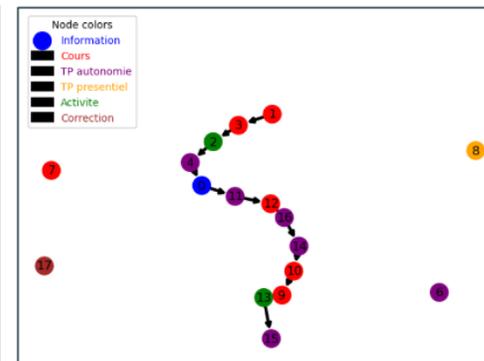
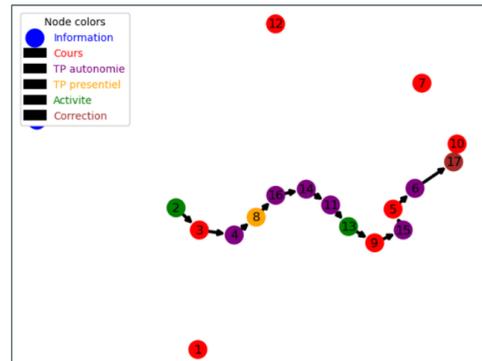
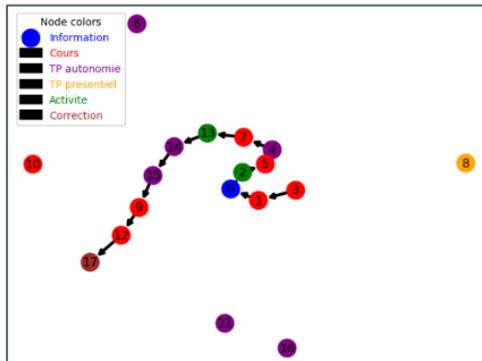
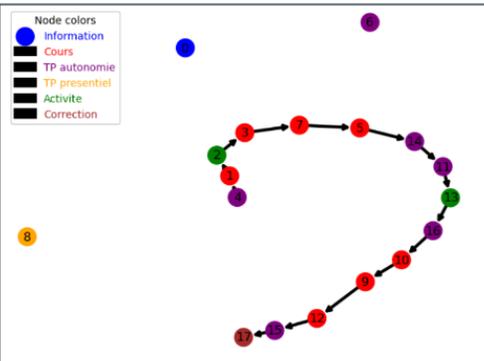
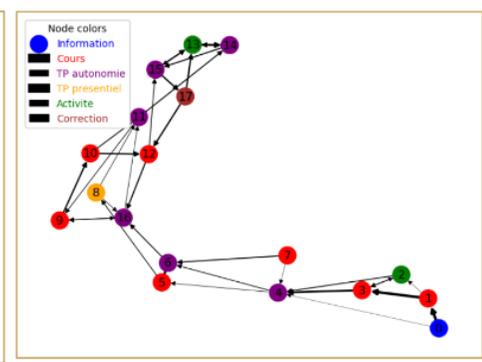
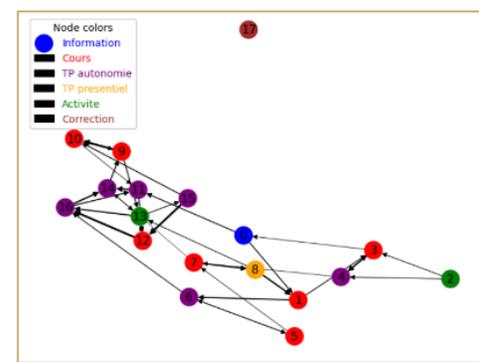
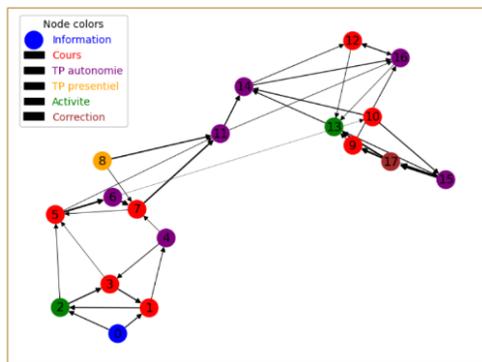
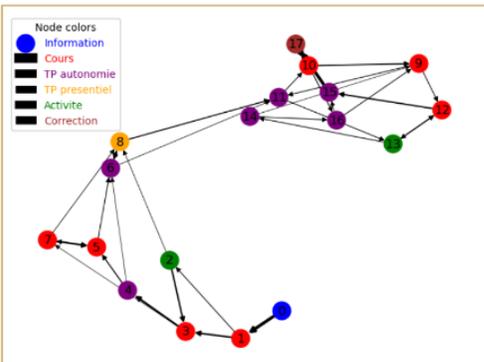
Cluster 1

Cluster 2

Cluster 3

Cluster 4

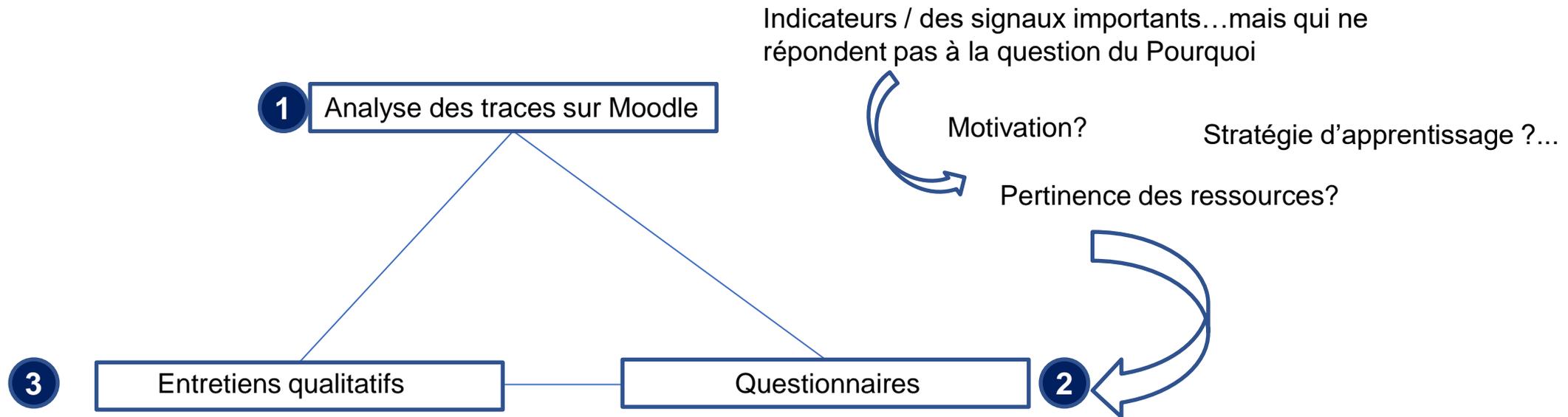
Cluster 5





Choix méthodologique

Protocole évolutif...





Choix méthodologique

Un protocole évolutif...



- Approche qualitative : comprendre les usages et les ressentis des étudiants
- Entretiens anonymisés, enregistrés (enregistreur vocal)
- Ecran d'ordinateur montrant le cours, pour aider les étudiants à se rappeler de leur activité et à donner systématiquement des exemples.

Réponses aux questionnaires de l'EEE pour le cours OPI : 49 réponses.

Questionnement sur 3 thèmes (questions générales ouvertes) :

- La façon dont l'étudiant s'organise pour ce cours
- Les activités du cours (en distanciel comme en présentiel)
- Les évaluations dans le cours



Recoupement des données : problèmes identifiés

- participation moindre au cours du fait de sa modalité à distance
- difficultés à s'organiser pour réaliser les activités du cours
- prise en main du cours à la "dernière minute"
- ceux qui ont le mieux réussi ne sont pas ceux qui ont suivi linéairement le cours
- sentiment d'isolement/décrochage du cours
- documents nouveaux non vus par les étudiants sur Moodle
- manque d'activités pour s'entraîner
- évaluations jugées trop difficiles



Protocole et préconisations : apports de la littérature

- Protocole

Marie-Luce Bourguet, & Yushan Li. (2022). Distance measure between instructor-recommended and learner's learning pathways. *Proceedings of the 15th International Conference on Educational Data Mining*, 729--733. <https://doi.org/10.5281/zenodo.6853087>

- Autodirection

Carré, P. (2010). Chapitre 3. L'autodirection des apprentissages. Dans : Philippe Carré éd., *L'autoformation: Perspectives de recherche* (pp. 117-169). Paris cedex 14: Presses Universitaires de France. <https://doi.org/10.3917/puf.carre.2010.02.0117>

- Présence à distance

Jézégou, A. (2012). La présence en e-learning : modèle théorique et perspectives pour la recherche. *International Journal of E-Learning & Distance Education (IJEDE)*, 26(1). <https://ijede.ca/index.php/jde/article/view/777/1409>

- Alignement pédagogique

Biggs, J. (1996). Enhancing Teaching through Constructive Alignment. *Higher Education*, 32(3), 347–364. <http://www.jstor.org/stable/3448076>



Recommandations proposées

	Piste non retenue	Piste à retenir
Positionnement du cours par rapport aux autres cours : revaloriser le statut du cours		
Créer un syllabus en début de cours Moodle		
Faire du cours OPI un prérequis pour les autres cours de programmation		
Faire coïncider OPI avec les cours de programmation		
Essayer de faire en sorte que les chargés de TP soient les mêmes pour OPI et pour le cours de programmation		
Temps de travail pour ce cours		
Donner des échéances plus resserrées pour éviter l'effet « dernière minute »		
Pour chaque activité, mettre des dates d'achèvement avec une notification si l'activité n'est pas réalisée		
Moyens de favoriser l'autodirection des étudiants		
Créer un planning visuel avec les activités et les échéances (planning transposable les années suivantes) en début de cours Moodle (Genially, H5P peuvent être de bons outils)		
Pour chaque activité, mettre le « temps de travail estimé »		
Partir de la tâche à réaliser pour ensuite donner des supports/ressources qui permettront de la réaliser		



Recommandations proposées, suite

	Piste non retenue	Piste à retenir
Présence à distance		
Prévoir un temps synchrone à distance ou en présentiel, à chaque phase, pour lever les difficultés et expliquer les notions mal comprises : animation par l'enseignant ou par des pairs avec intervention de l'enseignant		
Créer plus d'activités à faire en binôme		
Instaurer un tutorat entre pairs		
Création/amélioration des supports de cours		
Créer des vidéos pas à pas		
Indiquer des ressources externes à consulter		
En cas d'ajout sur Moodle, faire une notification aux étudiants		
Fournir la grille critériée pour le DM « atelier » dès le début de l'activité pour indiquer les attendus		
Programmer seulement 2 tentatives pour les quiz afin d'éviter une boucle infinie		
Créer plus d'activités d'entraînement		
Éléments plébiscités		
Conserver les activités ludiques et les enrichir		
Imaginer un système de jeu avec des points à gagner pour tout le cours : challenge avec les quiz, etc.		
Evaluations		
Rendre plus visible le barème dans le syllabus		
Réduire l'écart de niveau entre les activités réalisées et le TP noté		



Merci pour votre attention

Des questions ?