Vers des modèles d'analytique des apprentissages avec le numérique

Camila Morais Canellas[1,2][0000−0002−7226−0931] - 2ème année

[1] Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
[2] Kelis conseil et développement

Résumé Dans un contexte de production de documents pédagogiques via une démarche d'ingénierie dirigée par les modèles, nous étudions les opportunités et contraintes liées à la définition de modèles d'analytique des apprentissages avec le numérique. Un des avantages du contexte est la connaissance systématique de la structure et de la sémantique de tous documents créés et, par conséquent, la possibilité d'explorer cette connaissance afin d'enrichir les indicateurs proposés et/ou de faciliter leur mise en place. Un deuxième avantage est la possibilité, dans le modèle qui sera proposé, de permettre l'adaptation des indicateurs selon les contextes d'usage et rôles des parties prenantes.

Mots-clés : ingénierie dirigée par les modèles · chaînes éditoriales · analytique des apprentissages avec le numérique

1 Introduction

Dans un contexte de développement de systèmes complexes, l'ingénierie dirigée par les modèles (IDM) est une pratique qui permet de se concentrer sur un niveau plus abstrait que la programmation classique [2]. Cette pratique permet de décrire à la fois le problème posé et sa solution.

En informatique, la modélisation est très utilisée afin d'essayer de maîtriser cette complexité [5], tant pour produire un logiciel (conception) que pour le valider (test). L'IDM est une forme d'ingénierie générative [2] au sens où elle suit une démarche par laquelle tout ou partie d'une application informatique est générée à partir de modèles. Dans le but d'atteindre un objectif donné, certains aspects de la réalité (ou d'une solution à un problème) sont simplifiés. Cette modélisation peut également servir à séparer les différents besoins fonctionnels et les préoccupations extra-fonctionnelles (telles que fiabilité, performance, etc.)

Dans un contexte où la production de ressources documentaires à but pédagogique s'appuie déjà sur l'IDM, nous nous interrogeons sur la manière par laquelle une telle démarche peut permettre une analyse poussée des traces laissées par les apprenants, tout en suivant également une démarche s'appuyant sur l'IDM afin de définir un nouveau modèle pour le cycle d'analyse des traces.

Un premier enjeu est de déterminer les modèles de génération de traces et donc les modalités d'exploitation de ces traces utiles en fonction des modalités pédagogiques des différents documents produits avec des chaînes éditoriales.

Il s'agit ici de proposer un modèle générique et adaptable et donc de contribuer à l'approche orientée par les modèles, déjà pratiquée par les éditeurs d'outils de production documentaire auxquels nous contribuons. Les enseignants qui utilisent telles solutions sont déjà habitués à une telle approche de séparation fond/forme, qui certes peut être peu intuitive au début, mais qui est souvent appréciée sur le long terme.

Plus précisément, nous nous posons la question suivante : comment inclure la connaissance du document (structure, sémantique) dans le cycle d'analyses de traces (production, analyses et visualisations), issues de la consultation de ressources pédagogiques produites par des outils auteurs générés par une approche d'ingénierie dirigée par les modèles ?

2 Un exemple de chaîne de production documentaire

2.1 Terminologie

Dans notre contexte, des développeurs ont défini des "briques", dites primitives documentaires, qui servent de base pour la création des modèles documentaires. Ceci correspond à avoir un *méta-méta-modèle* dans une démarche d'IDM (un exemple est donné dans la suite).

Ensuite, un outil de modélisation (dit *méta-modèle*) permet la définition de chaînes éditoriales. Dans la pratique, celui qui utilise cet outil, le modélisateur, définit un *modèle de document* en utilisant les primitives documentaires disponibles. L'enjeu de ce méta-modèle est d'abstraire au maximum la technicité des modalités de l'outil conçu.

Plus tard, un auteur (un enseignant, par exemple) utilise ce modèle pour créer un module de son cours. L'outil de conception inclut des algorithmes de génération qui transforment la modélisation en document, permettant à l'enseignant de publier son cours sur plusieurs formats tels que PDF, Open Document, web, présentation synthétique, etc. L'intérêt ici est la séparation du fond et de la forme, ou l'étape de l'écriture de celle de la publication. Le passage (génération) d'un format à l'autre se fait automatiquement à travers des algorithmes de transformation internes à l'outil [1].

Nous nous intéressons notamment au *document* sous son format web et à la manière dont il va être utilisé ensuite par les apprenants, et ce format est le seul qui permette d'envisager un mécanisme de génération de traces hors outil dédié.

2.2 Dans la pratique

Afin d'appréhender les étapes et rôles des parties prenantes dans la démarche de conception de documents à des fins pédagogiques via l'ingénierie dirigée par les modèles, nous proposons ici un exemple pratique très simplifié. Cet exemple reprend le vocabulaire du modèle Opale [3] de Scenari, qui est adapté à la création de cours pour l'enseignement supérieur ou secondaire.

3. `https://doc.scenari.software/Opale/fr/`

Étape 1 - Le développeur Dans un premier temps, un développeur définit des primitives documentaires telles que "texte", "multimédia", "quiz", "organisation", etc. Chacune de ces "briques" sert à construire le modèle documentaire. C'est à partir de ces briques (qui peuvent ressembler à un bloc multimédia ou quiz sous Moodle, par exemple) que le modélisateur va travailler, en définissant un modèle de document qui répond aux besoins d'un groupe de métier. Le langage utilisé est donc adapté, il en va de même pour les fonctions disponibles.

Étape 2 - Le modélisateur Un modélisateur utilise ensuite les primitives disponibles pour définir des modèles documentaire. Par exemple, un modèle très simplifié est constitué d'"activités d'apprentissage". Celles-ci ont la possibilité (primitive d'organisation) d'avoir un ou plusieurs :

- "Introduction" (primitive texte)
- "Concept" (primitive texte ou image)
- "Grain de contenu" : constitué à son tour de parties (primitives texte et images) intitulés "information" et "exemple".
- "Conclusion" (primitive texte)
- "Pratique" : constitué à son tour de quiz (primitives quiz).

Un module peut avoir une ou plusieurs "activités d'apprentissage". Une "activité d'apprentissage" doit obligatoirement avoir exactement une "introduction", une "conclusion" et une partie "pratique" à la fin, et inclure entre "introduction" et "conclusion" un ou plusieurs "concept" et/ou "grains de contenu". Une partie "pratique" doit avoir un ou plusieurs quiz.

Le modélisateur définit aussi que, lors de la publication en format web, une page est créée pour chaque "activité d'apprentissage", que celle-ci contienne un ou plusieurs "concepts" et/ou "grains de contenu", un défilement de la page permet d'accéder aux contenus s'ils sont longs. Un menu est également créé et affiché à gauche, permettant aux apprenants de parcourir le module, soit par les pages, soit par des sauts vers les différentes parties internes à la page.

Le modélisateur définit de ce fait la structure (chaque partie et ses constituants possibles/obligatoires) ainsi que la sémantique de base (ce que chaque partie devra contenir) pour chaque document qui se fondera sur ce modèle documentaire. En d'autres termes, les différentes "parties" possibles du document sont pré-établies au moment de la modélisation et le type de contenu de chaque partie est défini par leurs noms même, donc la sémantique sera connue. Ainsi, l'auteur doit choisir le modèle correspondant aux besoins de son métier et cela aura des conséquences également dans la phase d'analyse des traces, notamment en ce qui concerne la sémantique. En d'autres mots, il ne pourra pas, par exemple, définir une autre sémantique pour son cours.

Étape 3 - L'auteur Un auteur, par exemple un enseignant, utilise ce modèle pour créer un cours. Son cours est constitué de quatre "activités d'apprentissage", chacune avec une "introduction", deux "concepts" et quatre "grains de contenu", une "conclusion", suivie d'une activité "pratique" afin de vérifier la

compréhension du contenu théorique. Une fois le cours créé, l'enseignant peut le publier en format web, en le plaçant sur la plate-forme dédiée de son choix.

Étape 4 - L'apprenant Les apprenants accèdent aux contenus publiés, ouvrent les pages dans l'ordre qui leur convient, les font défiler, répondent aux quiz, etc.

3 Un modèle d'analytique des apprentissages avec le numérique

Notre but est d'établir un modèle d'analytique des apprentissages avec le numérique qui s'appuie sur la démarche IDM, c'est-à-dire de définir pour chaque phase de modélisation un méta-méta-modèle, un méta-modèle, un modèle et, enfin, un document (visualisations) concernant les traces d'apprentissage numérique liées à l'usage du document original.

3.1 Des modèles pour le cycle d'analyse des traces

L'exemple de modèle documentaire ci-dessus nous permet de nous poser des questions relatives aux opportunités que ce contexte pourrait présenter au regard de l'analyse des traces numériques produits par les apprenants. Une première question est : *Peut-on proposer des modèles de collecte, analyse et visualisations prenant en compte l'ensemble de la chaîne de production de documents ?*

Deux particularités de notre contexte sont : 1/ Le besoin d'abstraction du modèle de traces pour tenir compte de tous les modèles documentaires existants ou à créer. 2/ L'absence initiale de données : contrairement à une grande partie des études dans le domaine, notre approche ne se fonde pas sur l'interrogation des données existantes avec pour but d'en tirer des informations pertinentes, mais est liée à une démarche de définition d'observations qui pourraient être utiles et utilisées [6] pour, ensuite, définir les traçages nécessaires pour y arriver.

La définition de notre modèle d'analytique des apprentissages avec le numérique passe alors par l'analyse des étapes de notre exemple (e.g. définition de ce que chaque primitive documentaire doit produire comme trace enrichie, définition des indicateurs à proposer, des algorithmes pour les traiter, des visualisations possibles), mais aussi par la répartition des rôles aux différents acteurs de notre exemple (e.g. qui définit les indicateurs, pour chaque modèle documentaire existant, qui choisit le(s) visualisation(s) possible(s) pour chaque acteur, etc.).

Prenons l'exemple d'un indicateur simple et très utilisé : le nombre de consultations. Si une page d'un cours en ligne a été visionnée par un apprenant, usuellement une trace est créée (ex : "l'apprenant X a consulté la page Y"). Le modèle d'analytique des apprentissages avec le numérique en mode IDM pourrait aller plus loin et définir que chaque page doit être tracée au niveau de ses sous-parties. Nous aurions alors non seulement "l'apprenant X a consulté la sous-partie W de la page Y", mais "l'apprenant X a consulté l'introduction de l'activité d'apprentissage Y".

Il est important de noter dans cet exemple que, non seulement la trace apporte l'information selon une granularité différente (définie), mais également le fait que : 1/ ce qui a été consulté est une "introduction" ; 2/ l'"introduction" en question fait partie d'une telle "activité d'apprentissage". Il devient assez facile de transposer des analyses à un cours créé à partir du même modèle : chaque activité d'apprentissage aurait également une "introduction". En d'autres mots, la sémantique et la structure du document sont de fait présentes dans la trace [4] ce qui pourrait donner lieu à des analyses plus pertinentes et adaptables, selon les besoins des contextes.

La seconde question est relative au choix de l'acteur prenant la décision de tracer la page vs. ses sous-parties (granularité), et l'acteur définissant comment l'information est présentée à ceux qui lui succèdent dans la conception. Par exemple, le modélisateur définirait que les pages et ses sous-parties sont tracées, mais laisserait à l'auteur deux choix sur la manière dont l'information serait présentée : dans un tableau de bord, en forme d'un diagramme à barres pour l'enseignant et d'un point vert/vide à coté de chaque partie dans le menu pour les apprenants, selon qu'ils l'ont visualisée ou pas.

Le modèle de traces doit alors prendre en compte la définition de qui décide quoi dans chaque phase, mais aussi expliciter les contraintes liées aux choix déjà faits et aux choix restants à faire dans les phases qui suivent. Si le modélisateur décide d'inclure un indicateur, un pré-requis est que les traces permettant de le calculer soient bien créées. Ou encore, si le choix de l'apprenant est d'avoir une visualisation particulière, il est nécessaire d'activer une trace de bas niveau à son tour nécessaire au calcul d'indicateur associé (approche *top-down*). Dans ce sens, un outil d'aide à la décision à l'égard du modélisateur pourrait être envisagé, contenant les relations entre traces brutes, indicateurs, visualisations des indicateurs.

3.2 Des indicateurs enrichis

Notre deuxième défi est l'identification de possibilités d'enrichissement des indicateurs couramment utilisés dans les démarches d'analytique des apprentissages avec le numérique grâce aux avantages apportés par les outils étudiés dans notre contexte. *Comment enrichir les indicateurs couramment utilisés dans les démarches d'analytique des apprentissages avec le numérique avec la connaissance fine de la structure et de la sémantique du document ?*

Ainsi, nous souhaitons spécialement comprendre les apports potentiels de ce contexte pour les différentes analyses de données, qu'elles soient analyses descriptives, diagnostiques, prédictives ou prescriptives [4].

En reprenant notre exemple simplifié de document, le fait de savoir que chaque "activité d'apprentissage" se termine par une partie "pratique" composée de quiz visant à vérifier la compréhension de l'activité pourrait nous pousser

4. Dans les faits, cette information pourrait se trouver directement encodée au niveau de la trace ou dans un fichier contenant la structure du document. Dans tous les cas, l'information serait accessible et permettrait les mêmes analyses. L'équilibre technique entre ces deux options reste à définir.

à enrichir un indicateur. Par exemple, l'indicateur reprenant le résultat de la partie "pratique" pourrait être croisé avec celui du nombre total de visualisations des sous-parties de l'activité d'apprentissage en question. Ou, peut-être, avec seulement le nombre de visualisations de la sous-partie "conclusion", etc.

Admettons que le modélisateur ait défini le premier indicateur cité auparavant comme facultatif, au choix de l'auteur de l'activer. Admettons ensuite que les résultats de cette analyse (lors d'autres cours utilisant le même modèle) montrent que l'échec à la partie "pratique" est systématiquement corrélé à une absence de lecture de l'ensemble des sous-parties. Via une remontée de cette information au modélisateur, de manière automatique ou manuelle, cela permettrait de proposer un changement du méta-modèle rendant l'indicateur intégré automatiquement aux cours créés utilisant tel modèle.

4 Perspectives

Après une première étude qualitative avec enseignants et apprenants, et dans une démarche d'IDM, nous concluons en ce moment la proposition d'un modèle d'analytique des apprentissages avec le numérique (composé d'un méta-méta-modèle, un méta-modèle, un modèle, et un document) suffisamment abstrait afin d'être utilisable par les modèles documentaires existants et à créer. Nous choisissons également quelques indicateurs qui nous permettrons à la fois de tester notre modèle et de montrer les avantages de l'enrichissement de ceux-ci.

Un autre moyen de tirer partie des traces enrichies serait d'encourager la mise en pratique de techniques d'apprentissage humain considérées comme les plus utiles. Ainsi, il est établi qu'expliquer avec ses propres mots la raison pour laquelle un fait ou un concept explicitement déclaré est vrai est une technique d'apprentissage particulièrement utile [3]. Nous savons grâce aux traces qu'il vient de consulter une partie "concept", ce qui pourrait permettre d'implémenter des règles du type [quand une structure du type "concept" est détectée, ouvrir avec une interface proposant à l'apprenant de s'arrêter pour produire une explication avant de continuer vers les parties suivantes]. Ce type d'approche serait plus facilement mis en place grâce à la connaissance préalable du document réfléchie dans les traces.

Une fois le modèle d'analytique des apprentissages avec le numérique établi, le plus grand avantage sera l'automatisation des processus de collecte, analyse et visualisation des ces traces pour tout document produit, tout en gardant des options de personnalisation et d'enrichissement. Bien que certains de ces résultats soient atteignables sans une démarche d'IDM, le coût de la mise en place serait plus important pour les auteurs si l'on considère la même granularité et le même enrichissement des traces.

Les résultats de la mise en place de ce modèle seront alors testés auprès des parties prenantes (entretiens, questionnaires), selon les possibilités pragmatiques du terrain. Nous nous servirons d'une série d'indicateurs choisis selon des critères (phase d'analyse, complexité, type, enrichissement, etc.) nous permettant de tester à la fois le modèle proposé en lui-même et l'utilité de celui-ci. Dans la suite,

l'éditeur des solutions Scenari pourra suivre sa démarche actuelle de recueil de besoins (*top-down* et *bottom-up*), afin de choisir quels indicateurs implémenter et proposer pour chaque modèle documentaire et contexte d'usage. Une étude qui vise à identifier les indicateurs utilisés dans des démarches d'analytique des apprentissages avec le numérique est en cours à cet effet.

Références

1. Bachimont, B., Crozat, S. : Instrumentation numérique des documents : pour une séparation fonds/forme. Revue I3 - Information Interaction Intelligence, C **4**(1), 95 (2004)
2. Combemale, B. : Ingénierie Dirigée par les Modèles (IDM) – État de l'art (2008)
3. Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., Willingham, D.T. : Improving students' learning with effective learning techniques : Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, Supplement **14**(1), 4–58 (2013)
4. Gartner, I. : Gartner Says Advanced Analytics Is a Top Business Priority (2014), `https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority`
5. Jézéquel, J.M., Combemale, B., Vojtisek, D. : Ingénierie dirigée par les modèles - Des concepts à la pratique. Ellipses, Paris (2012)
6. Ouali, M.A., Iksal, S., Laforcade, P. : The strategic organization of the observation in a TEL system : Studies and first formalizations. CSEDU 2014 - Proceedings of the 6th International Conference on Computer Supported Education **1**, 317–324 (2014). https://doi.org/10.5220/0004847703170324