

Introduction à la recherche scientifique

Master Informatique 2017-18

Dominique Py

1

Motivation

- ▣ C'est la recherche qui fait progresser les domaines scientifiques.
- ▣ La recherche se pratique dans les laboratoires mais aussi dans les entreprises.
- ▣ En M2 informatique, vous choisirez entre un stage « professionnel » et un stage « recherche ».

2

Objectifs de l'UE

- ▣ Donner un aperçu de ce qu'est la recherche en informatique.
 - ▣ Faire un peu de théorie, un peu de pratique.
 - ▣ Aborder les deux domaines de recherche du LIUM
 - Ingénierie pédagogique (EIAH)
 - Parole et traduction
- ☞ faire un choix informé en M2 et plus tard.

3

Contenu

- ▣ Un cours
 - Qu'est-ce que la recherche?
 - En EIAH?
 - En reconnaissance de la parole?
- ▣ Deux TD (EIAH, parole)
- ▣ Deux TP (EIAH, parole)
- ▣ Examen écrit + notes de TP

Intervenants: A. Laurent (parole), D. Py (EIAH)

4

Plan

1. La recherche scientifique
 - Qu'est-ce qu'un problème de recherche?
 - Démarche et méthodes de recherche
2. Les chercheurs
 - Le travail du chercheur
 - Le parcours du chercheur
 - Les produits de la recherche
 - L'évaluation de la recherche
 - Le processus de publication

5

1 – La recherche scientifique

6

La recherche : formuler un problème, apporter une réponse

Formuler un problème

*Connaître l'état de la science
Délimiter le problème*

L'étudier

*Suivre une méthodologie
Utiliser des concepts et outils*

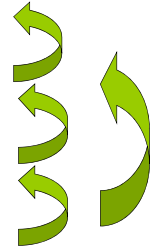
Proposer une réponse

*Évaluer ses apports et limites
Identifier les nouvelles questions*

7

Les étapes d'un travail de recherche

1. Identifier le sujet
2. Définir la méthode
3. Appliquer la méthode
4. Analyser les résultats
5. Communiquer



8

Caractéristiques d'une recherche scientifique

□ L'objet

Toute science se caractérise par son objet d'étude

□ La théorie

Une science s'appuie sur un système de connaissances

□ La méthode

La méthode doit être appropriée à l'objet de la recherche

□ L'objectif

Compléter, enrichir le système de connaissances

□ Les moyens

Données, temps, fonds

9

Une activité de recherche est

□ Une activité systématique

Produire un ensemble cohérent de connaissances et les intégrer dans un système de connaissances

□ Centrée sur la réalité

Confronter les théories aux observations
Éviter la spéculation

□ Utilisant des outils adaptés

Outillage théorique et pratique
Méthodes d'expérimentation, d'analyse

□ Essayant de généraliser

Produire des lois, des théories

10

Finalités de la recherche

□ Exploratoire

Étude d'un nouveau phénomène

□ Explicative

Rechercher les causes d'un phénomène

□ Descriptive

Décrire un phénomène ou des objets

□ Prédictive

Prédire des événements ou comportements

□ Ingénierie

Produire un outil

11

Qu'est-ce qu'une recherche intéressante?

□ Elle apporte quelque chose de nouveau

- Des réponses à des questions nouvelles, ou qui n'avaient pas reçu de réponse adéquate.
- Des réponses différentes à des questions déjà traitées.
- Une confirmation de résultats connus par une nouvelle méthode.
- L'application d'une théorie à des objets nouveaux.

□ Elle apporte quelque chose d'utile

- Des connaissances qui aident à mieux comprendre un phénomène, à mieux résoudre un problème.

12

- Recherche fondamentale
Produit de nouvelles connaissances sur le monde

- Les deux coexistent dans la plupart des domaines scientifiques*

- 13

- L'informatique est une science
Son objet : le traitement de l'information
Recherches fondamentales et appliquées
- L'informatique est aussi une technique
Son but : produire des artefacts (logiciels)
Ingénierie et développement

14

15

16

POUR MA PART, LA VIE DE CHERCHEUR...
C'EST UN MYSTÈRE. MAIS COMMENT PEUT
BIEN SE PASSER LEUR JOURNÉE ?

EST-CE QUE
DES MYSTÈRE
LE CHERCHEUR
EST AU TROUQUET ?

BONJOUR !

C'EST LA ?

C'EST POUR ?

2 - 3000 € - LES MEILLEURS

17

18

Le travail du chercheur en informatique (2)

- ▣ Réaliser des expérimentations
 - Définir le protocole, les données
- ▣ Analyser les résultats expérimentaux
 - Conformité aux attentes, explication
- ▣ Rédiger et publier des articles
 - Diffuser les résultats obtenus
- ▣ Participer à des congrès
 - Présenter son travail, échanger avec les autres chercheurs du domaine

19

Le travail du chercheur en informatique (3)

Mais aussi, selon les cas

- Encadrer des doctorants, des stages
- Organiser des congrès
- Mener des projets de recherche
- Établir des partenariats avec des entreprises
- Vulgariser la recherche auprès du public
- Diriger une équipe, un laboratoire
- Recruter des ingénieurs, des chercheurs
- Enseigner

20

Comment devenir chercheur?

- ▣ Diplôme de master (deux ans)
 - Pas de distinction parcours pro/recherche
- ▣ Doctorat (trois années)
 - Dans un laboratoire ou une entreprise
 - Financé (bourse de thèse)
 - Avec un directeur de thèse et un encadrant
 - Possibilité d'enseigner, de diffuser la recherche
 - Se conclut par la soutenance de thèse
- ▣ Post-doctorat, ATER
 - CDD recherche/enseignement en France ou à l'étranger

21

Le parcours du chercheur

- ▣ Concours de recrutement
 - Chercheur: CNRS, INRIA...
 - Enseignant-chercheur: université, IUT...
- Poste occupé: maître de conférences ou chargé de recherche
- ▣ Soutenance d'une habilitation à diriger des recherches + concours de recrutement
- Poste occupé: professeur ou directeur de recherche

22

Aussi: les parcours dans le privé

- ▣ Doctorat avec bourse CIFRE
 - Le doctorant est embauché par l'entreprise
 - Son directeur de thèse est un universitaire
 - Temps: 50% entreprise, 50% université
 - http://www.anrt.asso.fr/fr/espace_cifre/accueil.jsp
- ▣ Secteurs R&D des entreprises
 - Les entreprises embauchent des docteurs
 - Dispositif « Crédit Impôt Recherche »

23

Exemple : le LIUM

Laboratoire d'Informatique de l'Université du Maine (deux sites: Le Mans et Laval)

26 enseignants-chercheurs, 20 doctorants, 7 post-doctorants, 6 ingénieurs et administratifs.

Chercheurs en informatique (et linguistique)

Deux thèmes de recherche

- Environnements Informatiques d'Apprentissage Humain
- Parole et traduction

24

Les produits de la recherche

- ▣ Publications scientifiques
 - Livres
 - Revues
 - Congrès et colloques
- ▣ Brevets
- ▣ Contrats et collaborations avec des entreprises pour passer à l'échelle industrielle

25

L'évaluation des travaux de recherche

La sélection des publications dans un congrès repose sur la « relecture par les pairs » (*peer review*).

- Elle assure la qualité scientifique des publications.
- Elle garantit l'anonymat des relecteurs (*blind review*).

L'évaluation des chercheurs se fait principalement sur leurs publications.

26

Publication d'un article de recherche dans un journal ou une revue

Les intervenants:

- ▣ Relecteurs (*referees*)
 - Le journal possède un *comité de lecture*.
 - Des relecteurs sont sollicités pour relire un article et donnent séparément leur avis.
- ▣ Éditeur (*editor*)
 - L'éditeur choisit les relecteurs et prend la décision en se basant sur leurs avis.
- ▣ Maison d'édition (*publisher*)
 - Elle publie les articles dans la revue.

27

Le processus (1)

- ▣ L'auteur choisit une revue appropriée et soumet son article suivant les instructions des éditeurs.
- ▣ L'éditeur vérifie le respect des critères élémentaires :
 - thème de l'article en adéquation avec la revue,
 - longueur de l'article,
 - mise en forme.
- ▣ L'éditeur choisit n relecteurs dans le comité de lecture et leur demande de relire l'article.

28

Le processus (2)

- ▣ Les relecteurs lisent l'article (grille d'analyse) et transmettent leur avis à l'éditeur.
 - ▣ L'éditeur prend la décision :
 - acceptation en l'état,
 - acceptation après modifications,
 - rejet,
- puis la transmet à l'auteur avec les commentaires des relecteurs.

29

Le processus (3)

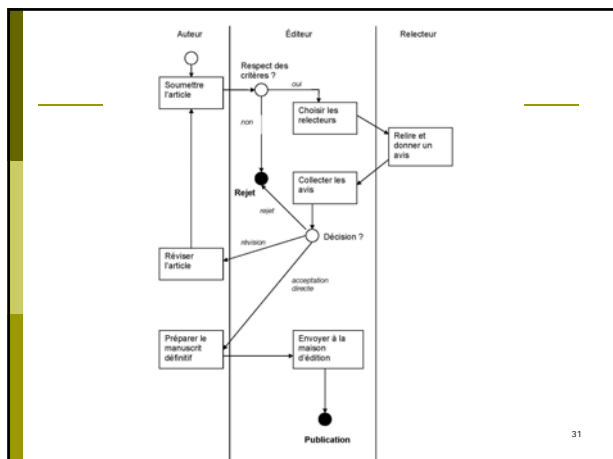
Selon la décision de l'éditeur

- ▣ Acceptation

L'auteur transmet le manuscrit définitif à la revue pour publication.
- ▣ Modifications

L'auteur soumet une nouvelle version modifiée, qui est transmise aux mêmes relecteurs (ou éventuellement, il renonce à la publication).

30



Publication d'un article de recherche dans un colloque ou un congrès

- ▣ Les organisateurs constituent un comité de programme et les relecteurs sont issus de ce comité
- ▣ L'auteur envoie un article avant la date limite.
- ▣ Le comité de programme se répartit les articles
- ▣ Chaque relecteur de l'article donne un avis
- ▣ Le comité de programme se réunit et prend une décision (acceptation ou rejet)
- ▣ La décision est transmise aux auteurs avec les commentaires des relecteurs
- ▣ En cas d'acceptation, l'auteur modifie son article et transmet le manuscrit définitif.

32

Environnements Informatiques d'Apprentissage Humain

Master informatique
2017-18

EIAH

1 - Définitions, caractérisation

Définitions

« **Un EIAH** est un environnement informatique dont la finalité explicite est de susciter et d'accompagner l'apprentissage humain, c'est-à-dire la construction de connaissances chez un apprenant » [Grandbastien 2006]

Définitions

Un EIAH

- ▣ mobilise un ou plusieurs agents humains
 - un ou plusieurs apprenants
 - éventuellement, un ou plusieurs enseignants
- ▣ va du simple logiciel au dispositif complexe
 - plateforme, système en réseau
 - ressources locales ou distribuées



Définitions

L'EIAH est un domaine de recherche qui a pour objet l'apprentissage humain avec et par des dispositifs informatiques

Les travaux en EIAH visent à concevoir des modèles, réaliser des dispositifs informatiques et les expérimenter avec des utilisateurs en conditions réelles

TICE versus EIAH

- ▣ TICE (Technologies de l'Information et de la Communication pour l'Éducation)
 - variantes : *e-learning*, formation en ligne, ...
 - applications de l'informatique à l'éducation
 - point de vue pratique et technologique
- ▣ EIAH
 - étude de l'apprentissage avec et par des dispositifs informatiques
 - point de vue recherche

EIAH : champ interdisciplinaire

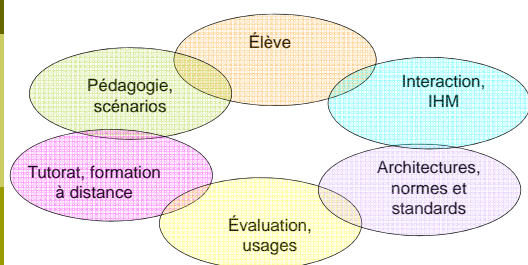
L'EIAH est à l'interface

- ▣ de l'informatique
 - interaction homme-machine (ergonomie, interfaces)
 - intelligence artificielle (connaissances, raisonnement)
 - internet et web
- ▣ et des sciences cognitives
 - sciences de l'éducation (méthodes pédagogiques)
 - didactique (épistémologie, obstacles aux apprentissages)
 - psychologie (apprentissage chez le sujet humain)

Un champ de recherche institutionnel

- ▣ Laboratoires
 - LIUM (Le Mans), LIG (Grenoble), LIP6 (Paris), ...
- ▣ Formations (master, école d'été)
- ▣ Revues scientifiques
 - International Journal of Artificial Intelligence in Education (IJAIED)*, Sciences et Techniques Educatives (STICEF: www.sticef.org), ...
- ▣ Congrès annuels ou biannuels
 - AI&ED*, *Intelligent Tutoring Systems*, *ICALT*, EIAH, ...
- ▣ Associations savantes
 - AI&ED Society*, Association des Technologies de l'Information pour l'Éducation et la Formation (ATIEF <http://www.inrp.fr/atief>)

Thèmes de recherche



Approche scientifique

- ▣ Démarche théorique : conception de modèles décrivant des situations, des dispositifs
- ▣ Démarche d'ingénierie : réalisation de systèmes informatiques opérationnels
- ▣ Démarche empirique : expérimentation, observation de situations réelles d'apprentissage

Productions

Qu'est-ce qu'un résultat en EIAH ?

- ▣ un modèle
- ▣ un logiciel ou un composant logiciel
- ▣ une contribution à une meilleure compréhension des processus d'apprentissage médiatisés par ordinateur

Méthodologies de conception pour les EIAH (1)

- ▣ Les méthodes classiques du génie logiciel (en cascade, en V, en spirale) sont peu adaptées
 - Les fonctionnalités du système sont mises en avant au détriment de l'interaction
 - Principe d'indépendance entre le noyau fonctionnel et l'interface utilisateur
- ▣ Alternatives
 - conception itérative
 - conception participative

Conception itérative

- ▣ Processus de construction itératif par phases
 - affinements progressifs des spécifications
 - évaluations des solutions retenues
 - réalisations, modifications jusqu'à obtention d'un produit satisfaisant
- ▣ Caractéristiques
 - adapté aux problèmes difficiles à spécifier
 - processus de conception ni ascendant, ni descendant
 - développement de solutions partielles, intermédiaires
 - apparition de nouveaux objectifs
 - communication entre concepteurs et utilisateurs

Conception participative

- ▣ Prise en compte des utilisateurs comme partenaires dans la conception
- ▣ Avantages
 - indispensable pour les activités mal identifiées ou peu structurées
 - facilite l'acceptation du logiciel
- ▣ Inconvénients
 - possibilité de contradiction entre les utilisateurs et les autres concepteurs
 - obligation de trouver des compromis

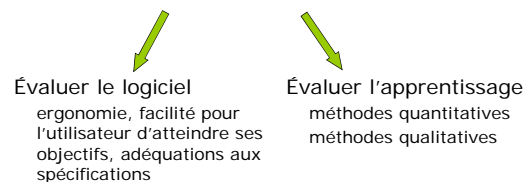
Méthodologies de conception pour les EIAH (2)

Bilan

- ▣ Pas de méthode unique
- ▣ Combiner les avantages de différentes méthodes
- ▣ Évaluations régulières en cours de développement
- ▣ Prise en compte précoce de l'utilisateur

Évaluation des EIAH (1)

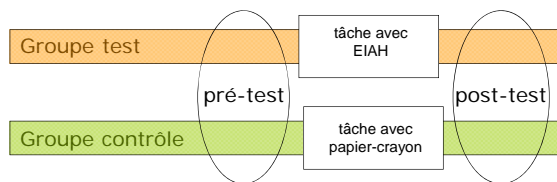
Évalue-t-on le logiciel ou les apprentissages?



Évaluation des EIAH (2)

Méthode comparative (psychologie cognitive)

- Groupe test et groupe contrôle
- Mesure de la progression entre les tests



Évaluation des EIAH (3)

Méthodes qualitatives (didactique)

Analyse manuelle à partir de

- entretiens individuels ou collectifs
- recueil de verbalisations
- productions écrites des apprenants

Long et complexe mais mieux adapté à des apprentissages de haut niveau

Difficultés

- Capitalisation des acquis (modèles, résultats)
- Reproductibilité des expérimentations
- Transfert industriel, passage à grande échelle
- Vitesse de progression de la technologie >>> celle de la recherche
- Ni algorithmes ni benchmarks

EIAH

2 – Un exemple : Buggy

Buggy, Debuggy, Idebuggy: les erreurs procédurales

(R. Burton et J. Brown, 1978-82)

Hypothèses de base

- Les savoir-faire en arithmétique (addition, soustraction) peuvent être représentés par des procédures organisées en réseau
- Les erreurs ont un caractère systématique
- Elles s'expliquent par la présence de « perturbations » dans les procédures

C'est la première approche cherchant à expliquer les erreurs des élèves

Exemples

Réponse	Explication(s) possible(s)
$\begin{array}{r} 143 \\ -25 \\ \hline 128 \end{array}$	L'élève ne reporte pas la retenue
$\begin{array}{r} 345 \\ -82 \\ \hline 343 \end{array}$	Dans chaque colonne, l'élève soustrait le plus petit chiffre du plus grand, quelle que soit leur place
$\begin{array}{r} 140 \\ -29 \\ \hline 129 \end{array}$	Quand une colonne comporte un zéro en haut, l'élève indique comme résultat le chiffre du bas

Buggy: un modèle exécutable

□ Le modèle

- Un réseau de procédures élémentaires (ajouter 10, reporter la retenue, comparer deux chiffres, passer à la colonne suivante, etc.)
- Toute procédure peut être remplacée par une variante incorrecte

□ L'interaction

- Buggy introduit une perturbation dans le modèle correct et l'exécute sur une série de soustractions
- L'utilisateur doit reconnaître la perturbation puis prédire le résultat sur de nouvelles soustractions

Intérêts et limites

□ Intérêts

- Théorie générative plus puissante que les théories descriptives: se passe d'une description exhaustive des erreurs (catalogue)
- Modèle plus explicatif que les approches statistiques
- Prend en compte la systématique des erreurs

□ Limites

- Ne représente pas le « bruit » (variabilité des erreurs)
- Les savoirs-faire de l'élève sont décrits par rapport à ceux de l'expert (même niveau et granularité)
- Le réseau est figé: ne reflète pas l'évolution de la *structure* des connaissances

Debuggy: outil de diagnostic

- Basé sur le même modèle des erreurs
- Diagnostic à partir d'un ensemble de réponses
- Détecte les combinaisons d'erreurs
- Approche combinatoire et heuristique du diagnostic
 1. Générer des réponses erronées et les comparer aux réponses de l'élève
 2. Retenir les erreurs expliquant certaines réponses
 3. Combiner les erreurs élémentaires et retenir les combinaisons expliquant au mieux les réponses
 4. Introduire des erreurs spécifiques (« *coercions* ») pour traiter le bruit (ex: $6 - 4 = 3$)

Idebuggy: diagnostic en ligne

- Version interactive: diagnostiquer les bugs au fur et à mesure des réponses
- Difficulté: les calculs ne doivent pas ralentir l'interaction
- Avantage: le système peut générer des problèmes discriminants pour lever les ambiguïtés du diagnostic
- Inconvénient: au début, l'espace des hypothèses est vaste (emploi d'heuristiques)

Expérimentations

- Buggy utilisé en formation des maîtres
 - Aide à prendre conscience des aspects logiques et systématiques des erreurs
 - Améliore les capacités de diagnostic humain
- Debuggy utilisé avec des élèves
 - Nombreuses évaluations, semble une réussite
 - L'efficacité du diagnostic dépend de la qualité du jeu de test
- Idebuggy: des expérimentations mais peu documentées