

Que peut apporter la fouille de données aux Learning Analytics ?

Armelle Brun

Equipe KIWI – Laboratoire LORIA
Institut des Sciences du Digital, Management et Cognition - IDMC

Armelle.Brun@loria.fr

Les Learning Analytics

▪ Objectifs

- Collecter, mesurer, analyser, synthétiser les données sur
 - les étudiants,
 - les enseignants
 - leurs contextes respectifs
- Comprendre et évaluer processus d'apprentissage
- Améliorer
 - Informations
 - Recommandations étudiants/enseignants

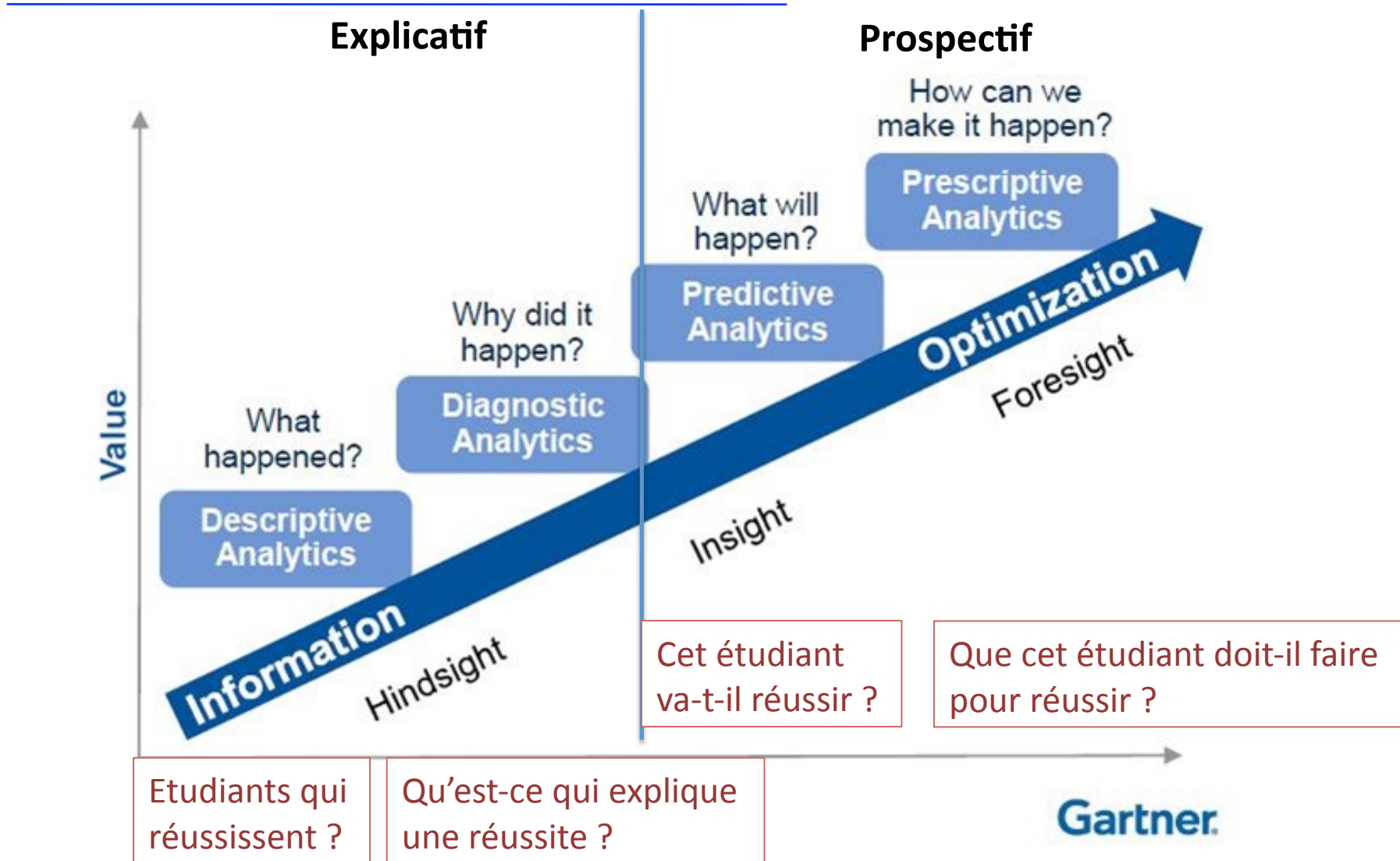
▪ Dimensions abordées

- Réussite étudiants
 - Flexibilité, personnalisation
- Rationnalisation, efficacité du système
- Éthique

Les Learning Analytics – les données

- **Existent désormais en grand nombre**
 - Accessibles aisément
- **Domaine récent (2011)**
 - prometteur, nombreux travaux
- **Quelles données ?**
 - Apprentissage
 - interaction des apprenants avec environnement d'apprentissage
 - Connaissances
 - S.I. étudiant, méta-données ressources, cours,
 - Enseignement
 - Interaction des enseignants avec l'environnement
 - Administratives
 - Programme, règlement

Les objectifs de l'analytics



La fouille de données

- **Apprentissage automatique (machine learning)**
- **Aussi appelée extraction d'informations**
 - Relations/associations entre données
 - Répétitions
- **Grands volume de données**
 - Fouille automatique / semi-automatique
- **Fouille de données vs statistiques ?**
- **Concrètement**
 - Segmenter
 - structuration des données, découvertes de communautés
 - Classifier
 - Prédiction d'une communauté/classe d'une entrée
 - Décrire

La fouille de données - Décrire

▪ Fouille motifs

- Identification liens entre données
 - récurrences, associations, corrélations (cachées,)

▪ Quels domaines ?

- Commerce, santé, RH, bancaire, éducation

▪ Quelles données ?

- Transactions d'achats
- Dossiers médicaux

▪ Exemples de motifs

- (pain, confiture, beurre)
 - vs <appareil photo, zoom, trépied> ?
- (aspirine, collyre)
 - vs <aspirine, cétizirine, cortisone> ?
- Identification de « comportements » ou de « trajectoires »

▪ Mesures de qualité

La fouille de données - Décrire

▪ Quel modèle ?

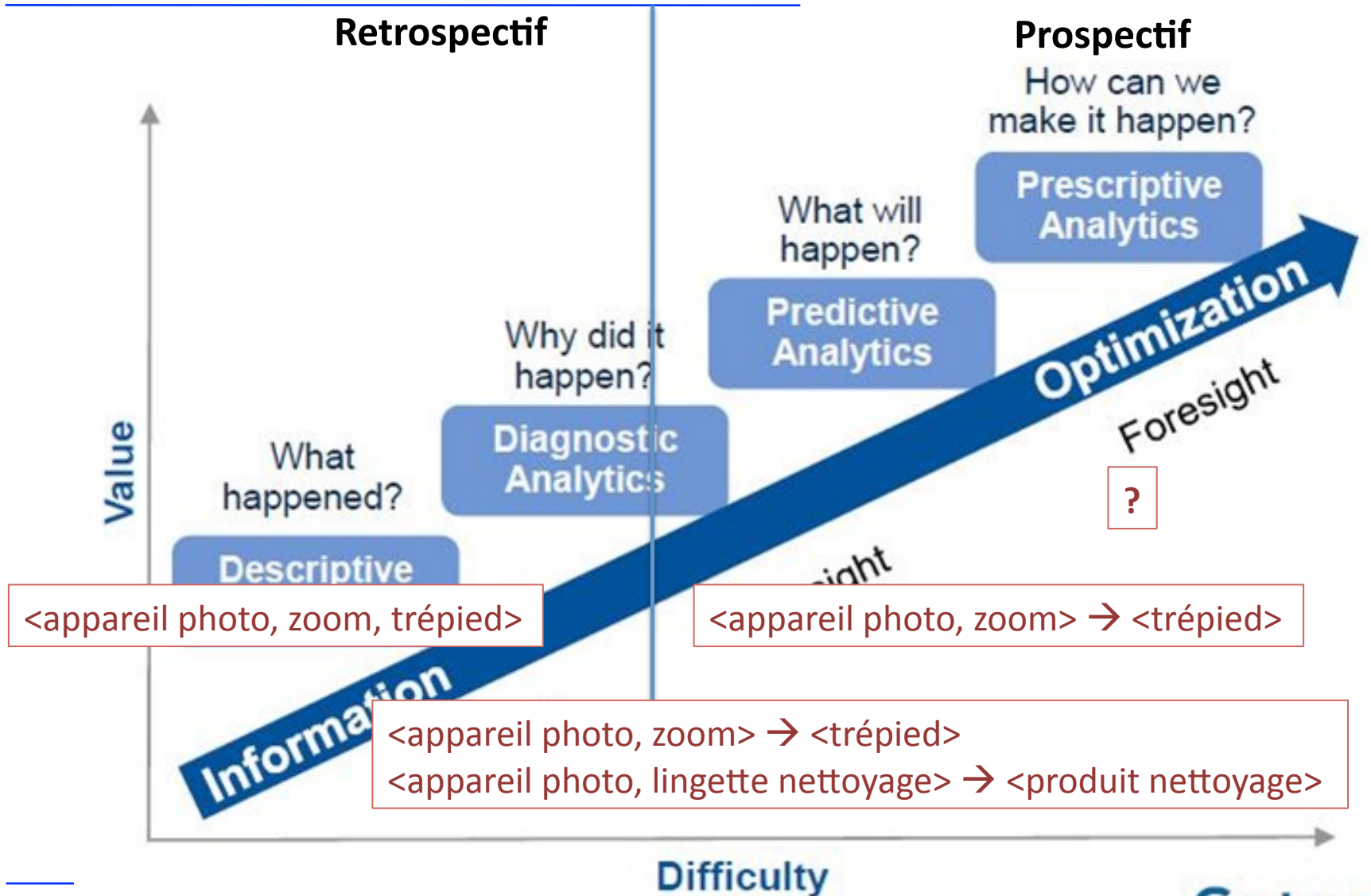
- résumé/synthèse
 - par extraction d'information
- compréhensible pour un humain
 - Sans compréhension des motifs (des modèles) exploitation (pour l'action) limitée
- autres approches ML ?

▪ Autre objectif

- Prise de décision
- Fouille de règles

<appareil photo, zoom, trépied> <appareil photo, zoom> → <trépied>

La fouille de données - l'analytics



La fouille de données

▪ 2 champs de recherche

- Diminuer la complexité de l'algorithme de fouille
 - Temps de recherche
 - Mémoire utilisée
- Augmenter la pertinence de l'ensemble résultat
 - Motifs intéressants
 - Critères d'intérêt/qualité
 - Taille de l'ensemble de motifs

La fouille de données pour les *learning analytics*

- **Analytics « automatique »**
- **Identifier des motifs et faire des prédictions/prescriptions personnalisées**
 - Respectent les caractéristiques de l'apprenant (connaissances habitudes, etc.)
 - Respectent ses objectifs
- **Fouille de**
 - Motifs
 - Règles

- **Prédiction d'abandon**
 - Qui abandonne ?
 - Pourquoi abandonne ?
 - Futur abandon ?
 - Prédiction à quel horizon ?

La fouille de données pour les *learning analytics*

- Source de données unique

- activités des apprenants sur l'environnement d'apprentissage

Activités
étudiants

id1, R12, R13, R65, R43
id7, R12, R13, R99, R43
id9, R55, R13, R65, R43
....

R12, R13

R13, R65, R43

R13, R65 → R43

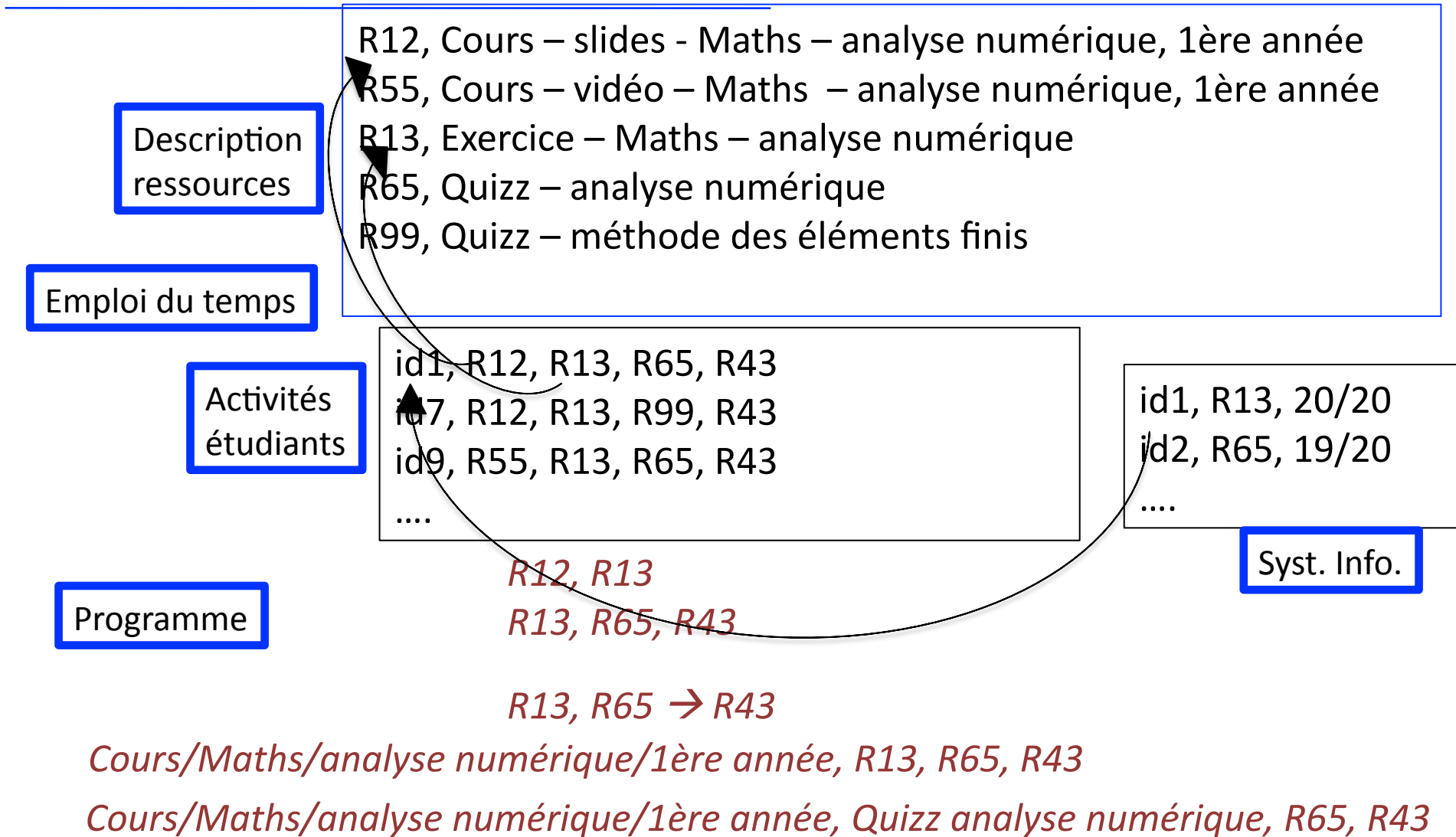
- Une vraie compréhension ?

La fouille de données pour les *learning analytics*

▪ Sources de données multiples

- Meta-données, connaissances externes – description
 - ressources, étudiants, formation, etc.
- Données multi-relationnelles,
 - liées

La fouille de données pour les *learning analytics*



Quelques défis

- **Modélisation distante ?**
 - Effets à longs termes

- **Evénements influenceurs**

- **La précision du temps**
 - Données non ordonnées : quel exemple ?
 - Données ordonnées
 - Données semi-ordonnées

- **Multi-source**
 - Toute source d'information est la bienvenue
 - Quelles sources utiles ?
 - Smart data

- **Emergence / disparition de comportements**
 - Etudiants avec spécificités