

Module Description, available in: EN, FR

Data Management

General Information

Number of ECTS Credits

3

Module code

TSM_DataMgmt

Valid for academic year

2019-2020

Last modification

2018-11-02

Responsible of module

Stefan Keller (FHO, stefan.keller@hsr.ch)

Explanations regarding the language definitions for each location:

- Instruction is given in the language defined below for each location/each time the module is held.
- Documentation is available in the languages defined below. Where documents are in several languages, the percentage distribution is shown (100% = all the documentation).
- The examination is available 100% in the languages shown for each location/each time it is held.

	Berne	Lausanne	Lugano	Zurich
Instruction		X F 100%		X E 100%
Documentation			X E 100%	X E 100%
Examination		X F 100%		X E 100%

Module Category

TSM Technical/scientific specialization module

Lessons

2 lecture periods and 1 tutorial period per week

Entry level competences

Prerequisites, previous knowledge

- Relational Models, Relational Algebra
- Normalization
- SQL
- Transaction Processing, Concurrency Control
- Security in relational database systems
- Query optimization
- RDBMS architectures
- Basics in probability theory

Brief course description of module objectives and content

Aims, content, methods

Learning objectives and acquired competencies

- Students understand the use of modern database technologies for processing and managing large and shared data collections.
- In addition to RDBMS, students also know current data structures (data types) and alternative (e.g. non-relational) database systems and know which of these data types and systems to use depending on the situation and type of data available.
- The students know methods and tools to filter out certain information from structured and unstructured data (data mining).
- Students know how to deal with fuzzy (text) information using databases and search engines (information retrieval).
- The students can also apply the acquired knowledge in their own working environment and to solve their specific problems.

Contents of module with emphasis on teaching content

The lecture is divided into four parts:

1. Database Management - New Data Structures and Alternatives to RDBMS (DB): The first part deals with the storage of data and with the non-relational aspects covered in the undergraduate course.
2. Data Warehousing and Decision Support (DW): The second part deals with data warehousing, i.e. data integration and data aggregation.
3. Data mining and data analysis (DM): In the third part, methods for the evaluation and extraction of information from large amounts of data are developed.
4. Information Retrieval (IR): The fourth part deals with finding certain information, e.g. by text search in databases, search engines and web crawlers.

Weighting:

1. DB: 10 - 30%, 2 - 4 weeks
2. DW: ~ 20%, 3 weeks
3. DM: ~ 30%, 4 weeks
4. IR: 20 - 40%, 3 - 5 weeks

Teaching and learning methods

Frontal teaching, exercises, case studies.

Literature

Optional literature suggestion (books):

- DB: Lena Wiese: Advanced Data Management for SQL, NoSQL, Cloud and Distributed Databases. De Gruyter Textbook. 2015. ISBN 978-3-11-044140-6.
- DW/DM: "Data Mining. Concepts and Techniques". Han, Kamber & Pei. 3rd ed. 2011. San Francisco. ISBN 978-0123814791.
- IR: "Modern Information Retrieval". Baeza-Yates & Ribeiro-Neto, New York (2011). ISBN: 9780321416919.
- IR: Introduction to Information Retrieval. C.D. Manning, P. Raghavan, H. Schütze. Cambridge UP, 2008. Classical and web information retrieval systems: algorithms, mathematical foundations and practical issues.
- IR: Information Retrieval in Practice. B. Croft, D. Metzler, T. Strohman. Pearson Education, 2009.

Assessment

Certification requirements

Module does not use certification requirements

Basic principle for exams

As a rule, all the standard final exams for modules and also all repetition exams are to be in written form

Standard final exam for a module and written repetition exam

Kind of exam

written

Duration of exam

120 minutes

Permissible aids

Aids permitted as specified below:

Permissible electronic aids

Scientific calculator (without communication functions).

Other permissible aids

Summary on one A4 page (possibly described on both sides).

Special case: Repetition exam as oral exam

Kind of exam

oral

Duration of exam

30 minutes

Permissible aids

No aids permitted

Description du module, disponible en: EN, FR

Gestion de données

Informations générales

Nombre de crédits ECTS

3

Code du module

TSM_DataMgmt

Valable pour l'année académique

2019-2020

Dernière modification

2018-11-02

Nom du/de la responsable de module

Stefan Keller (FHO, stefan.keller@hsr.ch)

Explication des définitions de langue par lieu :

- Les cours se dérouleront dans la langue définie ci-dessous par lieu/exécution.
- Les documents sont disponibles dans les langues définies ci-dessous. Pour le multilinguisme, voir la répartition en pourcentage (100% = documents complets)
- L'examen est disponible à 100% dans chaque langue sélectionnée pour chaque lieu/exécution.

	Berne	Lausanne	Lugano	Zurich
Leçons		X F 100%		X E 100%
Documentation			X E 100%	X E 100%
Examen		X F 100%		X E 100%

Catégorie de module

TSM approfondissement technico-scientifique

Leçons

2 leçons et 1 leçon de pratique par semaine

Compétences préalables

Connaissances préalables, compétences initiales

- Modèle relationnel, algèbre relationnelle
- Normalisation
- SQL
- Traitement transactionnel, contrôle d'accès simultané
- Sécurité dans les systèmes de bases de données relationnelles
- Optimisation des requêtes
- Architecture des SGBDR
- Bases dans la théorie de la probabilité

Brève description du contenu et des objectifs

Objectifs, contenus, méthodes

Objectifs d'apprentissage, compétences à acquérir

- Les étudiants savent comment utiliser les technologies des bases de données modernes afin de traiter et de gérer des données provenant de grandes bases communes.
- Ils connaissent de nouvelles structures de données (types de données) alternatives aux systèmes de gestion de bases de données relationnels (SGBDR) (non relationnelles notamment) et sont capables de déterminer quels types de données et quel système de base de données sont appropriés en fonction du contexte et du genre de données disponibles.
- Ils sont au courant des méthodes et des outils qui permettent d'extraire des informations provenant de données structurées ou non structurées (Data Mining).
- Ils savent comment manier des informations (sous forme de texte) qui ne sont pas définies précisément au moyen de BD et de moteurs de recherche (Recherche d'informations).
- Ils sont capables de réutiliser les connaissances acquises durant ce cours dans leur propre environnement de travail et de les appliquer afin de résoudre leurs problèmes spécifiques.

Contenu des modules avec pondération du contenu des cours

Le module s'articule en quatre parties (domaines) :

1. Gestion de bases de données – nouvelles structures de bases de données et alternatives aux SGBDR (DB): la première partie traite des aspects concernant le stockage des données, et des aspects NON relationnels des données abordés lors des premières années d'études.
2. Entreposage de données et aide à la décision (DW) : la deuxième partie traite de l'entreposage des données, c'est-à-dire de l'intégration et de l'agrégation des données.
3. Exploration et analyse de données (DM) : la troisième partie traite de l'élaboration de méthodes pour analyser des données et en extraire des informations de grandes quantités de données.
4. Recherche d'informations (IR) : la quatrième partie traite de la recherche d'informations et inclut la recherche contextuelle dans les BD, les moteurs de recherche et les robots Web.

Pondération :

1. DB: 10 ~ 30%, 2 - 4 semaines
2. DW: ~ 20%, 3 semaines
3. DM: ~ 30%, 4 semaines
4. IR: 20 ~ 40%, 3 - 5 semaines

Méthodes d'enseignement et d'apprentissage

Enseignement magistral, exercices, études de cas.

Bibliografie

Suggestion de bibliographie optionnelle (ouvrages):

- DB: Lena Wiese: Advanced Data Management for SQL, NoSQL, Cloud and Distributed Databases. De Gruyter Textbook. 2015. ISBN 978-3-11-044140-6.
- DW/DM: "Data Mining. Concepts and Techniques". (Concepts et techniques de datamining) Han, Kamber & Pei. 3rd ed. 2011. San Francisco. ISBN 978-0123814791.
- IR: "Modern Information Retrieval". (Recherche d'information moderne) Baeza-Yates & Ribeiro-Neto, New York (2011). ISBN: 9780321416919.
- IR: Introduction to Information Retrieval. C.D. Manning, P. Raghavan, H. Schütze. Cambridge UP, 2008. Classical and web information retrieval systems: algorithms, mathematical foundations and practical issues.
- IR: Information Retrieval in Practice. B. Croft, D. Metzler, T. Strohman. Pearson Education, 2009.

Evaluation

Conditions d'admission

Le module n'utilise pas de conditions d'admission.

Principe pour les examens

En règle générale, tous les examens de fin de module réguliers et les examens de rattrapage sont organisés sous la forme écrite

Examen de fin de module régulier et examen écrit de répétition

Type de l'examen

écrit

Durée de l'examen

120 minutes

Aides autorisés

Les aides suivantes sont autorisées:

Aides électroniques autorisées

Calculatrice scientifique (sans fonction de communication).

Autres aides autorisées

Résumé sur une page A4 (probablement recto-verso).

Cas spécial: examen de répétition oral**Type de l'examen**

oral

Durée de l'examen

30 minutes

Aides autorisés

Sans aides