MASTER IN
COMPUTATIONAL DATA SCIENCE
Study plan and content of the courses
Curriculum “Data Analytics”

The Curriculum "Data Analytics" focuses mainly on teaching mathematical-statistical techniques for data analysis, together with computer techniques for the creation of data-centric, end-user-oriented applications capable of extracting knowledge to support decision-making and business processes.

Study Plan

First Year

<table>
<thead>
<tr>
<th>1st Semester</th>
<th>Course</th>
<th>CP</th>
<th>Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advanced Data Management Technologies</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence – Methods and Apps</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Programming and Visualisation for Data Advoc</td>
<td>12</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Statistics for Data Science</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2nd Semester</th>
<th>Course</th>
<th>CP</th>
<th>Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Human-Centered Computing</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Information Retrieval</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Machine Learning</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Systems Security</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Optional Course</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

Second Year

<table>
<thead>
<tr>
<th>1st Semester</th>
<th>Course</th>
<th>CP</th>
<th>Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data Curation</td>
<td>12</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Data Integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data Profiling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Capstone Project</td>
<td>6</td>
<td>Idoneità/Eignung</td>
</tr>
<tr>
<td></td>
<td>Optional Course</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>Free Choice*</td>
<td>6</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2nd Semester</th>
<th>Course</th>
<th>CP</th>
<th>Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advanced English for Scientific Communication</td>
<td>4</td>
<td>Idoneità/Eignung</td>
</tr>
<tr>
<td></td>
<td>Free Choice*</td>
<td>6</td>
<td>**</td>
</tr>
<tr>
<td>Course</td>
<td>CP</td>
<td>Exam</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Advanced Topics in Machine Learning</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Agile Software Development</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Computational Linguistics</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Data and Information Modelling</td>
<td>12</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Data Maintenance and Evolution</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Data Scientist Toolbox</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Decision Making and Support Systems</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Development of Data Products</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Enterprise Architecture</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Formal Verification of Software and Systems</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Intelligent Agents</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Introduction to Parallel Computing</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Lean Start-Up and Entrepreneurship</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Management of Temporal and Spatial Data</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Process-Aware Information Systems</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Process Mining</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Real-Time Big Data Processing</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Recommender Systems</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Research Methods and Technology Transfer</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Semantic Technologies and Linked Data</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Simulation and Modelling</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
<tr>
<td>Web and Text Mining</td>
<td>6</td>
<td>Sì/Ja</td>
<td></td>
</tr>
</tbody>
</table>

* The student can freely advance or postpone the Free Choice credit points
** The student can choose courses that foresee both exams with grades and pass/fail tests
Content of the courses of the Curriculum "Data Analytics"

The teaching of the curriculum "Data Analytics" is structured in different thematic areas:
- Managing Data and Data Curation
- Mathematics
- Programming
- Capstone Project
- Artificial Intelligence and Analytics
- Data and People

Below is a description of the thematic areas and their related courses.

Managing Data and Data Curation - Curriculum "Data Analytics"

The aim of these courses is to train students in the technical aspects of data management and maintenance. On the one hand, methods and techniques for the modeling, design, archiving and processing of complex and heterogeneous data, the processes that govern its generation and use, as well as the corresponding IT infrastructures, are explored in depth. On the other hand, we study principles and tools for data sharing, cleaning, maintenance, integration, storage and evolution.

Core Courses

Advanced Data Management Technologies
- Data warehousing and business intelligence
- Multidimensional modelling and OLAP
- NoSQL
- Main memory database systems
- Programming models for data analysis, (e.g., Hadoop)
- RDD-based programming (e.g., Spark)

Data Curation
- Data integration architectures
- Schema mapping
- Data matching
- Heterogeneous and web data
- Data cleaning
- Query processing for data integration
- Detecting patterns and violations
- Detecting dependencies
- Scrubbing and normalization
- Similarity measures
- Duplicate detection
- Summary extraction

Optional Courses

Data Maintenance and Evolution
- Principles of data integration
- Materialised views, self-maintainable views, updating views
- Schema versioning and evolution, rewriting views in the context of evolution
- Evolution in the context of data warehouses
- Database degradation and restructuring
- Data provenance

Enterprise Architecture
- Business strategy and goal modelling
- Enterprise ontologies
- Relating data modelling and process modelling in organisational contexts
- Service modelling, management and governance
- Problem-solving with enterprise models
- Modelling of business applications and computational infrastructure to support business operations

Management of Temporal and Spatial Data
- Spatial-temporal data
- Array data, time series data
- Streaming data
- Abstract and concrete temporal data models
- Spatial indexes
- Spatial analysis

Process-Aware Information Systems
- Integration of business processes and data
- Capturing process-aware information systems with colored Petri nets
- Qualitative analysis and verification
- Quantitative analysis and simulation
- Enactment of a process-aware information system
- Process intelligence

Data and Information Modelling
- Conceptual modelling for information systems
- Data modelling: languages, techniques, methods
- From data models to information systems
- Business process lifecycle
- Process modelling: languages, techniques, methods
- Linking data and processes
- Design of complex static object structures and role-based modelling
- Design of complex relational properties
- Design of complex enterprise events
- Verification and validation of models
- Patterns and anti-patterns
- Model-based code generation and mapping to different implementation platforms

Semantic Technologies and Linked Data
- Semantic metadata
- Linked data
- The RDF standard
- Semantic application architectures
- Distributed queries
- Adding semantics to relational databases
Mathematics - Curriculum “Data Analytics”

These courses provide the mathematical, statistical and algorithmic skills needed especially in the development of data analysis and intelligent applications.

Core Courses

Statistics for Data Science
- Hypothesis testing and ANOVA
- Test of independence
- Correlation
- Linear and Logistic regression with one and multiple variables
- Time series
- Probabilistic models (EM)

Optional Courses

Simulation and Modelling
- Monte Carlo methods
- Variance Reduction
- Uncertainty Quantification
- Stochastic Analysis
- Markov chains and graphs
- Applications to data management

Programming - Curriculum “Data Analytics”

Data science is at the intersection of mathematics, computing and domain skills and very often data scientists have to implement the tools they need, creating new software solutions or extending existing software. For this reason, a solid foundation in programming techniques for data science is fundamental.

Core Courses

Programming and Visualisation for Data Analytics
- Languages for programming data and data visualisation (Perl, Python, R, Java, Java script)
- Programming for text processing (matching, parsing and regular expressions)
- Interfacing and integrating (API programming, Plug-in development)
- Scripting for data science (e.g., simple shell programming)
- Dataflow programming
- Advanced Programming Paradigms
- Human psychology and perception
- Data and image models
- Visualisation software and tools
- Visual Diagnostics
- Exploratory data analytics
- Discovery methods

Optional Courses

Data Scientist Toolbox
- Tools and practices for data collection and logging
- Tools and practices for data pre-processing
● Tools and practices for data summarisation
● Tools for data storage, integration, and maintenance
● Tools and patterns for data analysis
● Tools for data presentation, visualisation, and interpretation

Development of Data Products
● Data centric software engineering
● Methods and practices for data product development
● Domain engineering for data product
● Managing the software quality of data products
● Methods for testing
● Maintenance of software

Introduction to Parallel Computing
● Introduction to parallel and distributed systems
● Shared memory model
● Distributed memory model
● Principle and design of parallel algorithms
● Selection of parallel algorithms
● Performance Analysis, optimization and tuning

Real-Time Big Data Processing
● Complex event processing
● Stream data mining
● Semantic techniques for Streaming data
● Programming models for streaming data
● CEP programming
● Distributed real-time computation system (e.g., Apache, Flink, Storm)

Capstone Project - Curriculum “Data Analytics”

Data science cannot be taught only on a theoretical level. Students must apply and test their skills on real data, interacting with domain experts. To this end, the students carry out a project on real data taken from concrete application domains, such as bioinformatics, sensors, internet of things, business information systems, tourism and agriculture. The project is carried out individually or in groups, autonomously under the joint supervision of a professor and one or more domain experts.

Core Courses
Capstone Project
● Individual or group project based on real data from a specific application domain in areas such as bioinformatics, internet of things, business information systems, tourism, agriculture.

Optional Courses
The student can choose a second Capstone Project.

Artificial Intelligence and Analytics - Curriculum “Data Analytics”

Through these courses students deepen the most important and widespread techniques for data analysis and knowledge extraction. The main methods of data/process mining, business intelligence, machine learning and automatic reasoning about data and corresponding IT systems are taught.
Core Courses

Artificial Intelligence – Methods and Applications
- AI and data/information processing, historical perspective and state of the art
- Overview of main AI techniques: exact and approximate methods, handling imperfect information, use and model domain knowledge.
- Tools and systems for AI based programming
- Project on AI topics, such as: natural language processing, information extraction, games, automated planning, applications of constraint solving, multiagent systems

Information Retrieval
- Document Indexing
- Vector Space Model
- Web Search
- Text Classification
- Topic Modelling
- Introduction to text mining

Machine Learning
- Concept learning
- Resampling and model selection
- Unsupervised learning
- Supervised learning
- Deep learning
- Reinforcement learning

Optional Courses

Advanced Topics in Machine Learning
- Convolutional Neural Networks (CNN)
- Training Neural Networks
- Understanding and visualising Convolutional Neural Networks
- Deep Reinforcement Learning
- Computer vision
- Image classification

Computational Linguistics
- Basics of human language (phonetics and phonology, morphology, syntax, semantics)
- Speech processing and main technological application in automatic speech recognition and text-to-speech synthesis
- Processing words: finite-state machines, parts-of-speech tagging, sequence language models
- Processing sentences: syntactic structures, syntactic and statistical parsing
- Combining techniques in data modelling, data mining, and knowledge discovery to analyse natural language
- Main applications: machine translation, information extraction, question answering, text analysis, categorization

Formal Verification of Software and Systems
- Introduction to formal verification, its current role in industry, and the role of automated reasoning
- Languages and techniques for software and systems verification
- Algorithms and techniques of state of the art solvers
- Using solvers to verify static and dynamic properties of systems
- Software tools to support formal verification
- Formal verification case studies

Intelligent Agents
- Agents and agent architectures
- Models and algorithms for individual agents
- Multiagent systems
- Communication and cooperation, self-interested agents and teamwork, goal-oriented behaviours
- Verification and reasoning
- Agent-based languages, software platforms and tools

Process Mining
- Event data and the process mining framework
- Process discovery
- Conformance checking
- Enrichments of process models with data and social/organisational aspects
- Operational decision support
- Experimentation on concrete case studies using different process mining tools

Recommender Systems
- Collaborative Filtering
- Content-Based and Semantic-Based Methods
- Graph-Based Approaches
- Context-aware systems
- Conversational Systems
- Decision Making

Web and Text Mining
- Link mining
- Social network analysis
- Document summarization, clustering and categorization
- Entity extraction
- Sentiment analysis
- Ranking models

**Data and People - Curriculum “Data Analytics”**

These courses aim to provide students with a good theoretical knowledge and practical experience of the skills essential for the design and evaluation of interactive data-based products, placing users, and their needs and expectations, at the centre of the process.

**Core Courses**

Advanced English for Scientific Communication
- Writing skills: improvement of academic and scientific written skills through the practice and production of subject-specific texts;
• Spoken skills: improvement of spoken interaction and production through the practice and production of academically and professionally acceptable presentations and other domain-specific speaking activities;
• Development of receptive skills through the exposure to and analysis of various types of written and spoken discourse typical of master’s studies in Computer Science
• Development of grammatical and lexical range and accuracy so that communication is fluent and spontaneous.

Human-Centered Computing
• Interaction design processes: context of use analysis, development, evaluation and empirical research studies
• Design principles, patterns, models, tools
• Data gathering methods spanning the interaction design process (e.g., interviews, probes, think-aloud, eye-tracking systems)
• Study design considerations for evaluation and empirical research in HCC (e.g., exploratory or explanatory study, formative or summative study, small-scale or large-scale study, qualitative or quantitative study)
• Study metrics (e.g., performance metrics, process metrics), data analysis and visualisation: statistical considerations.

Systems Security
• Security by design
• Security coding
• Data security
• Security of distributed systems
• Vulnerabilities and attacks
• Social security

Optional Courses

Agile Software Development
• Software crisis and the origin of the agile software movement
• Different agile software development approaches and key agile practices
• From time-boxed agile methods to continuous flow: lean software development
• Continuous experimentation and continuous software engineering
• Teamwork in agile software development
• Scaling agile: distributed and/or large software development projects using agile methods

Decision Making and Support Systems
• Modelling decisions
• Modelling uncertainty
• Modelling preferences
• Modelling negotiations
• Decision support tools
• Psychology of decision making
• Persuasion

Lean Start-up and Entrepreneurship
• Nature and characteristics of innovative start-ups
• Problem identification and validation with design thinking tools
- Customer development process
- Build-measure-learn loops
- Continuous retrospectives for start-up learning
- Supporting toolkits for start-up processes

Research Methods and Technology Transfer
- Research paradigms in information and computer sciences and research methods
- Quality assessment of research papers
- Literature review
- Dissemination techniques for research results
- Research planning
- Models and methods of technology transfer
Curriculum “Data Management”

The "Data Management" Curriculum is mainly focused on the modeling and management of data and corresponding software architectures at enterprise level, and imparts techniques and methods typical of information technology for the development of information systems and IT infrastructures for data storage and support to the execution of decision-making and business processes.

Study Plan

First Year

<table>
<thead>
<tr>
<th>1st Semester</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>CP</td>
<td>Exam</td>
</tr>
<tr>
<td>Advanced Data Management Technologies</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Algorithms for Data Processing</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>• Artificial Intelligence - Methods and Applications</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Data and Information Modelling</td>
<td>12</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>• Data and Process Modelling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Information Systems Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

2nd Semester

<table>
<thead>
<tr>
<th>2nd Semester</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>CP</td>
<td>Exam</td>
</tr>
<tr>
<td>Enterprise Architecture</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Semantic Technologies and Linked Data</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Systems Security</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Optional Course</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

Second Year

1st Semester

<table>
<thead>
<tr>
<th>1st Semester</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>CP</td>
<td>Exam</td>
</tr>
<tr>
<td>Data Curation</td>
<td>12</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>• Data Integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Data Profiling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capstone Project</td>
<td>6</td>
<td>Idoneità/Eignung</td>
</tr>
<tr>
<td>Optional Course</td>
<td>6</td>
<td>Si/Ja</td>
</tr>
<tr>
<td>Free Choice*</td>
<td>6</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

2nd Semester

<table>
<thead>
<tr>
<th>2nd Semester</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>CP</td>
<td>Exam</td>
</tr>
<tr>
<td>Advanced English for Scientific Communication</td>
<td>4</td>
<td>Idoneità/Eignung</td>
</tr>
<tr>
<td>Free Choice*</td>
<td>6</td>
<td>**</td>
</tr>
</tbody>
</table>
* The student can freely advance or postpone the Free Choice credit points
** The student can choose courses that foresee both exams with grades and pass/fail tests

**Optional Courses - Curriculum “Data Management”**

<table>
<thead>
<tr>
<th>Course</th>
<th>CP</th>
<th>Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Topics in Machine Learning</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Agile Software Development</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Computational Linguistics</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Data Maintenance and Evolution</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Data Scientist Toolbox</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Decision Making and Support Systems</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Development of Data Products</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Formal Verification of Software and Systems</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Human-Centered Computing</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Information Retrieval</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Intelligent Agents</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Introduction to Parallel Computing</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Lean Start-Up and Entrepreneurship</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Management of Temporal and Spatial Data</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Process Mining</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Process-Aware Information Systems</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Programming and Visualisation for Data Analytics</td>
<td>12</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Real-Time Big Data Processing</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Recommender Systems</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Research Methods and Technology Transfer</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Simulation and Modelling</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
<tr>
<td>Web and Text Mining</td>
<td>6</td>
<td>Sì/Ja</td>
</tr>
</tbody>
</table>

**Content of the courses of the Curriculum “Data Management”**

The teaching of the curriculum "Data Management" is structured in different thematic areas:

- Managing Data and Data Curation
- Mathematics
- Programming
- Capstone Project
- Artificial Intelligence and Analytics
- Data and People

Below is a description of the thematic areas and their related courses.
Managing Data and Data Curation - Curriculum “Data Management”

The aim of these courses is to train students in the technical aspects of data management and maintenance. On the one hand, methods and techniques for the modeling, design, archiving and processing of complex and heterogeneous data, the processes that govern its generation and use, as well as the corresponding IT infrastructures, are explored in depth. On the other hand, we study principles and tools for data sharing, cleaning, maintenance, integration, storage and evolution.

Core Courses

Advanced Data Management Technologies
- Data warehousing and business intelligence
- Multidimensional modelling and OLAP
- NoSQL
- Main memory database systems
- Programming models for data analysis, (e.g., Hadoop)
- RDD-based programming (e.g., Spark)

Data Curation
- Data integration architectures
- Schema mapping
- Data matching
- Heterogeneous and web data
- Data cleaning
- Query processing for data integration
- Detecting patterns and violations
- Detecting dependencies
- Scrubbing and normalization
- Similarity measures
- Duplicate detection
- Summary extraction

Enterprise Architecture
- Business strategy and goal modelling
- Enterprise ontologies
- Relating data modelling and process modelling in organisational contexts
- Service modelling, management and governance
- Problem-solving with enterprise models
- Modelling of business applications and computational infrastructure to support business operations

Data and Information Modelling
- Conceptual modelling for information systems
- Data modelling: languages, techniques, methods
- From data models to information systems
- Process modelling: languages, techniques, methods
- Process analysis and simulation
- Linking data and processes
- Design of complex static object structures and role-based modelling
- Design of complex relational properties
- Design of complex enterprise events
- Verification and validation of models
- Patterns and anti-patterns
- Model-based code generation and mapping to different implementation platforms

Semantic Technologies and Linked Data
- Semantic metadata
- Linked data
- The RDF standard
- Semantic application architectures
- Distributed queries
- Adding semantics to relational databases

Optional Courses

Data Maintenance and Evolution
- Principles of data integration
- Materialised views, self-maintainable views, updating views
- Schema versioning and evolution, rewriting views in the context of evolution
- Evolution in the context of data warehouses
- Database degradation and restructuring
- Data provenance

Management of Temporal and Spatial Data
- Spatial-temporal data
- Array data, time series data
- Streaming data
- Abstract and concrete temporal data models
- Spatial indexes
- Spatial analysis

Process-Aware Information Systems
- Integration of business processes and data
- Capturing process-aware information systems with colored Petri nets
- Qualitative analysis and verification
- Quantitative analysis and simulation
- Enactment of a process-aware information system
- Process intelligence

Mathematics - Curriculum “Data Management”

These courses provide the mathematical, statistical and algorithmic skills needed especially in the development of data analysis and intelligent applications.

Core Courses

Algorithms for Data Processing
- Graph algorithms and analytics
- Sequence algorithms
- Algorithms for numerical optimization
- Linear programming
- Fundamentals of computational complexity
Heuristic and approximation strategies for solving hard problems

**Optional Courses**

**Simulation and Modelling**
- Monte Carlo methods
- Variance Reduction
- Uncertainty Quantification
- Stochastic Analysis
- Markov chains and graphs
- Applications to data management

**Programming - Curriculum “Data Management”**

Data science is at the intersection of mathematics, computing and domain skills and very often data scientists have to implement the tools they need, creating new software solutions or extending existing software. For this reason, a solid foundation in programming techniques for data science is fundamental.

**Optional Courses**

**Data Scientist Toolbox**
- Tools and practices for data collection and logging
- Tools and practices for data pre-processing
- Tools and practices for data summarisation
- Tools for data storage, integration, and maintenance
- Tools and patterns for data analysis
- Tools for data presentation, visualisation, and interpretation

**Development of Data Products**
- Data centric software engineering
- Methods and practices for data product development
- Domain engineering for data product
- Managing the software quality of data products
- Methods for testing
- Maintenance of software

**Introduction to Parallel Computing**
- Introduction to parallel and distributed systems
- Shared memory model
- Distributed memory model
- Principle and design of parallel algorithms
- Selection of parallel algorithms
- Performance Analysis, optimization and tuning

**Programming and Visualisation for Data Analytics**
- Languages for programming data and data visualisation (Perl, Python, R, Java, Java script)
- Programming for text processing (matching, parsing and regular expressions)
- Interfacing and integrating (API programming, Plug-in development)
- Scripting for data science (e.g., simple shell programming)
- Dataflow programming
- Advanced Programming Paradigms
- Human psychology and perception
- Data and image models
- Visualisation software and tools
- Visual Diagnostics
- Exploratory data analytics
- Discovery methods

Real-Time Big Data Processing
- Complex event processing
- Stream data mining
- Semantic techniques for Streaming data
- Programming models for streaming data
- CEP programming
- Distributed real-time computation system (e.g., Apache, Flink, Storm)

Capstone Project - Curriculum “Data Management”

Data science cannot be taught only on a theoretical level. Students must apply and test their skills on real data, interacting with domain experts. To this end, the students carry out a project on real data taken from concrete application domains, such as bioinformatics, sensors, Internet of things, business information systems, tourism and agriculture. The project is carried out individually or in groups, autonomously under the joint supervision of a professor and one or more domain experts.

Core Courses
Capstone Project
- Individual or group project based on real data from a specific application domain in areas such as bioinformatics, internet of things, business information systems, tourism, agriculture.

Optional Courses
The student can choose a second Capstone Project.

Artificial Intelligence and Analytics - Curriculum “Data Management”

Through these courses students deepen the most important and widespread techniques for data analysis and knowledge extraction. The main methods of data/process mining, business intelligence, machine learning and automatic reasoning about data and corresponding IT systems are taught.

Core Courses
Artificial Intelligence - Methods and Applications
- AI and data/information processing, historical perspective and state of the art
- Overview of main AI techniques: exact and approximate methods, handling imperfect information, use and model domain knowledge.
- Tools and systems for AI based programming
- Project on AI topics, such as: natural language processing, information extraction, games, automated planning, applications of constraint solving, multiagent systems.

Machine Learning
• Concept learning
• Resampling and model selection
• Unsupervised learning
• Supervised learning
• Deep learning
• Reinforcement learning

Optional Courses

Advanced Topics in Machine Learning
• Convolutional Neural Networks (CNN)
• Training Neural Networks
• Understanding and visualising Convolutional Neural Networks
• Deep Reinforcement Learning
• Computer vision
• Image classification

Computational Linguistics
• Basics of human language (phonetics and phonology, morphology, syntax, semantics)
• Speech processing and main technological application in automatic speech recognition and text-to-speech synthesis
• Processing words: finite-state machines, parts-of-speech tagging, sequence language models
• Processing sentences: syntactic structures, syntactic and statistical parsing
• Combining techniques in data modelling, data mining, and knowledge discovery to analyse natural language
• Main applications: machine translation, information extraction, question answering, text analysis, categorization

Formal Verification of Software and Systems
• Introduction to formal verification, its current role in industry, and the role of automated reasoning
• Languages and techniques for software and systems verification
• Algorithms and techniques of state of the art solvers
• Using solvers to verify static and dynamic properties of systems
• Software tools to support formal verification
• Formal verification case studies

Information Retrieval
• Document Indexing
• Vector Space Model
• Web Search
• Text Classification
• Topic Modelling
• Introduction to text mining

Intelligent Agents
• Agents and agent architectures
• Models and algorithms for individual agents
• Multiagent systems
• Communication and cooperation, self-interested agents and teamwork, goal-oriented behaviours
• Verification and reasoning
Agent-based languages, software platforms and tools

Process Mining
- Event data and the process mining framework
- Process discovery
- Conformance checking
- Enrichments of process models with data and social/organisational aspects
- Operational decision support
- Experimentation on concrete case studies using different process mining tools

Recommender Systems
- Collaborative Filtering
- Content-Based and Semantic-Based Methods
- Graph-Based Approaches
- Context-aware systems
- Conversational Systems
- Decision Making

Web and Text Mining
- Link mining
- Social network analysis
- Document summarization, clustering and categorization
- Entity extraction
- Sentiment analysis
- Ranking models

Data and People - Curriculum “Data Management”

These courses aim to provide students with a good theoretical knowledge and practical experience of the skills essential for the design and evaluation of interactive data-based products, placing users, and their needs and expectations, at the centre of the process.

Core Courses

Advanced English for Scientific Communication
- Writing skills: improvement of academic and scientific written skills through the practice and production of subject-specific texts;
- Spoken skills: improvement of spoken interaction and production through the practice and production of academically and professionally acceptable presentations and other domain-specific speaking activities;
- Development of receptive skills through the exposure to and analysis of various types of written and spoken discourse typical of master’s studies in Computer Science
- Development of grammatical and lexical range and accuracy so that communication is fluent and spontaneous.

Systems Security
- Security by design
- Security coding
- Data security
- Security of distributed systems
- Vulnerabilities and attacks
- Social security

**Optional Courses**

**Agile Software Development**
- Software crisis and the origin of the agile software movement
- Different agile software development approaches and key agile practices
- From time-boxed agile methods to continuous flow: lean software development
- Continuous experimentation and continuous software engineering
- Teamwork in agile software development
- Scaling agile: distributed and/or large software development projects using agile methods

**Decision Making and Support Systems**
- Modelling decisions
- Modelling uncertainty
- Modelling preferences
- Modelling negotiations
- Decision support tools
- Psychology of decision making
- Persuasion

**Human-Centered Computing**
- Interaction design processes: context of use analysis, development, evaluation and empirical research studies
- Design principles, patterns, models, tools
- Data gathering methods spanning the interaction design process (e.g., interviews, probes, think-aloud, eye-tracking systems)
- Study design considerations for evaluation and empirical research in HCC (e.g., exploratory or explanatory study, formative or summative study, small-scale or large-scale study, qualitative or quantitative study)
- Study metrics (e.g., performance metrics, process metrics), data analysis and visualisation: statistical considerations.

**Lean Start-up and Entrepreneurship**
- Nature and characteristics of innovative start-ups
- Problem identification and validation with design thinking tools
- Customer development process
- Build-measure-learn loops
- Continuous retrospectives for start-up learning
- Supporting toolkits for start-up processes

**Research Methods and Technology Transfer**
- Research paradigms in information and computer sciences and research methods
- Quality assessment of research papers
- Literature review
- Dissemination techniques for research results
- Research planning
- Models and methods of technology transfer
Optional Course

The optional courses of the chosen curriculum allow the student to deepen her/his scientific and technical knowledge in a particular field.

The Faculty Council decides each year which of the optional courses contained in the list should be activated. The activated optional courses are announced in the Manifesto of Studies for the corresponding academic year.

Capstone Project

Capstone projects allow students to apply the scientific and technical knowledge acquired during the study using real data from a specific application domain in areas such as bioinformatics, sensors, Internet of things, business information systems, tourism, agriculture.

Capstone projects are project courses during which the student works autonomously on an individual or group project. The project is supervised by a faculty professor or researcher (hereinafter referred to as a tutor), and a domain expert. The tutor is the person responsible for the course and follows, guides and evaluates the project. The domain expert introduces the student to the data and features of the application domain and provides requirements, guidance and feedback.

The Faculty Council decides each year which capstone projects to activate. The application domains are defined through contacts on the territory and made known through the Manifesto of Studies of the relevant academic year.

Free Choice courses or internship

The student must freely choose courses or an internship for a total of 12 credit points.

The Degree Council of the Master annually approves a list of Free Choice courses. Courses that are not included in this list, or internships, must be approved by the Degree Council, which verifies consistency with the educational path of the student.

For internships, please refer to the General Internship Regulations of the University of Bozen/Bolzano.

Exams taken for Free Choice courses count as a single exam towards the total number of exams taken by the student.

Teaching activities

Various types of teaching activities are planned, which in different ways provide practical and theoretical knowledge in the fields of computer science:

- Lecture: regular meeting with the students during which the teacher explains the theoretical part of the program.

- Exercise: regular meeting with the students accompanying the lecture; during the exercise, the program is reworked, and/or the student, under the supervision of the lecturer, autonomously applies the theoretical knowledge acquired during the lecture.

- Lab: regular meeting with the students accompanying the lecture; during the lab the student, under the supervision of the professor, develops a software product autonomously, immediately applying the theoretical knowledge acquired during the lecture.
• Seminar: presentation in the form of a seminar during which advanced topics are discussed and understanding, communication and teamwork skills are developed. The presentation is carried out by experts in the field and/or students who participate in the seminar, and who deepen a specific topic at the suggestion of the lecturer.

• Capstone Project: activity during which the student works independently on an individual or group project with real data from an application domain, under the supervision of a teacher and a domain expert; in particular, the techniques studied are applied and developed, and planning, communication, user interaction and teamwork skills are acquired.

The amount of teaching is equivalent to 25 hours of student commitment per Credit Point.

These 25 hours are divided into 10 hours in class and 15 hours of individual study in the case of frontal teaching, and 2 hours in class and 23 hours of individual study in the case of Capstone Projects.

The exact subdivision of the hours in class into lectures, exercises, labs or seminars for each course held by internal staff is proposed annually by the lecturer within the Course Presentation Form) and approved with possible modifications by the Degree Council of the Master before the beginning of the academic year; for courses held by external staff the subdivision is decided annually by the Degree Council of the Master.

Lectures must be distributed over the entire semester in a regular and balanced manner.

**Language of Instruction**

The language of all lectures on the Master's degree course is English.