

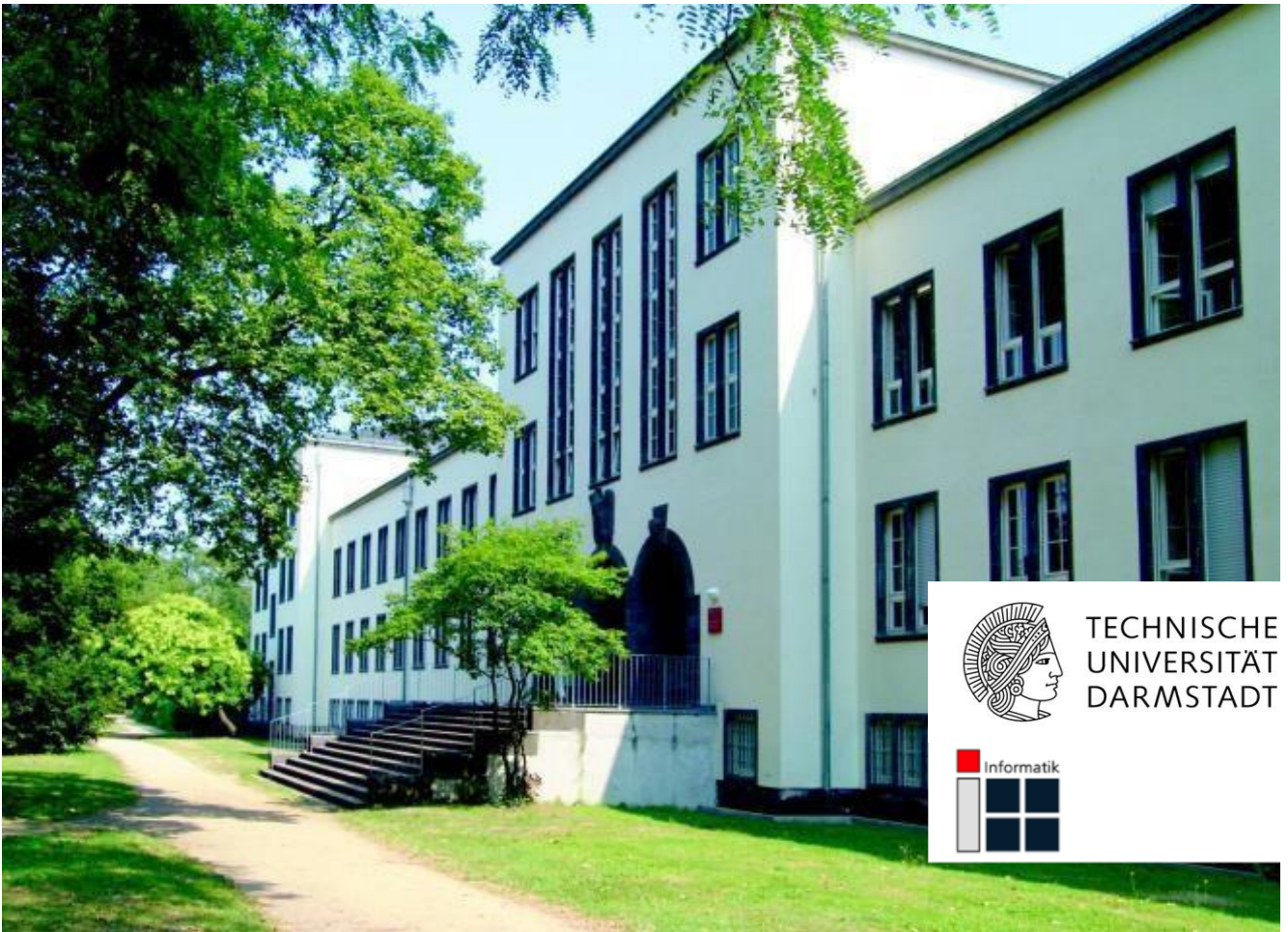
---

---

# Module Handbook

## Data and Discourse Studies – Computer Science

Technische Universität Darmstadt  
Department of Computer Science



## **Module Handbook Data and Discourse Studies – Computer Science**

Technische Universität Darmstadt

Fachbereich Informatik

Hochschulstr. 10

64289 Darmstadt

**Redaktion**

Dipl.-Inform. Tim Neubacher

**Stand: 13.09.2018**

<b>Module description</b>					
<b>Module</b> Information Management					
<b>Module No.</b>	<b>Credit points</b>	<b>Hours required</b>	<b>Self-study</b>	<b>Module duration</b>	<b>Recurrence frequency</b>
20-00-0015	5 CP	150 h	105 h	1 Semester	Summer semester
<b>Language</b> German			<b>Person responsible for the module</b> Vice Chairperson, Academic Affairs		
<b>1</b>	<b>Courses in the module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0015-iv	Information Management	5	Integrated course	3
<b>2</b>	<p><b>Content</b></p> <p>Basic concepts of information management:</p> <p>Concepts of information systems Information storage/retrieval, searching, browsing, declarative access and access via explicit navigation</p> <p>Quality characteristics: Consistency, scalability, availability, reliability</p> <p>Data Modeling:</p> <p>Conceptual data models (ER / UML structure diagrams) Conceptual design Operational Models (relational model) Mapping the conceptual to the operational model</p> <p>Relational model:</p> <p>Operators Relational algebra Relational calculi Impact on query languages based on relational algebra and relational calculus Design theory and normalization</p> <p>Query languages:</p> <p>SQL (in detail) QBE, Xpath (overview) Storage media: RAID, SSD Buffering and caching</p>				

	<p>Implementation of relational operators:  Implementation algorithms  Cost functions</p> <p>Query optimization:  Heuristic query optimization  Cost-based query optimization</p> <p>Transaction processing:  Flat transactions</p> <p>Concurrency control and correctness criteria:  Serializability, recoverability, ACA, strictness  Insulation grade  Lock-based scheduling, 2PL  Multiversions to control concurrency  Optimistic scheduling  Logging  Intermediate statuses (checkpointing)</p> <p>Current trends in Information Management:  Main memory databases  Column-based data management  NoSQL</p>
<b>3</b>	<p><b>Qualification objectives / learning outcomes</b>  After successfully taking part, the students will be informed about the basics of Information Management. They will understand the techniques of creating information management systems and be able to use the respective models, algorithms, and languages to independently use or create (parts of) information management systems. They will be able to rate the quality of such systems based on different quality measures.</p>
<b>4</b>	<p><b>Prerequisites for participation</b>  Recommended: Successful participation in the lectures "Funktionale und Objektorientierte Programmierkonzepte" and "Algorithmen und Datenstrukturen" or equivalent knowledge from other courses of study</p>
<b>5</b>	<p><b>Type of examination</b>  Subject examination in written form, 90 min.</p>
<b>6</b>	<p><b>Prerequisites for Credit Points</b>  <b>Passing the module exam (100%)</b></p>
<b>7</b>	<p><b>Grading</b>  Standard</p> <p>In this lecture, it is possible to take into account accompanying achievements, which – according to § 25 (2) of the 4th amendment of the APB and the rules adopted by the Department 20 on October 2, 2012 – can lead to an improvement in grades of up to 1.0.</p>

<b>8</b>	<p><b>Applicability of the module</b></p> <p>B. Sc. Computer Science          B. Sc. Business Informatics          B. Sc. Psychology in IT          Joint B.A. Computer Science          B. Sc. Sports Science and Computer Science          M. Sc. Sports Science and Computer science          Can also be applicable for other programs in the context of interdisciplinary offers.</p>
<b>9</b>	<p><b>Literature</b></p> <p>Will be announced. Examples are:          Haerder, Rahm, "Datenbanksysteme - Konzepte und Techniken der Implementierung", Springer 1999          Elmasri, R., Navathe, S. B.: Fundamentals of Database Systems, 3rd. ed., Redwood City, CA: Benjamin/Cummings          Ullman, J. D.: Principles of Database and Knowledge-Base Systems, Vol. 1 Computer Science</p>
<b>10</b>	<p><b>Comment</b></p>

<b>Module description</b>					
<b>Module</b> Data Mining and Machine Learning					
<b>Module No.</b> 20-00-0052	<b>Credit Points</b> 6 CP	<b>Hours required</b> 180 h	<b>Self-study</b> 120 h	<b>Module duration</b> 1 Semester	<b>Recurrence frequency</b> Usually every winter semester
<b>Language</b> German and English			<b>Person responsible for the module</b> Coordinators Web, Knowledge and Information Processing		
<b>1</b>	<b>Courses in the module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0052-iv	Data Mining and Machine Learning	6	Integrated course	4
<b>2</b>	<p><b>Content</b></p> <p>Due to the rapid development of information technology, there are increasing data volumes available – and they often contain implicit knowledge that, if it actually became known, could be of great economic or scientific significance. Data Mining is a research area that seeks to find potentially useful knowledge in large sets of data, and Machine Learning is one of the key technologies in this area.</p> <p>The lecture provides an introduction to the field of Machine Learning, with a special focus on the aspect of Data Mining. Methods from different paradigms of Machine Learning are presented with references to exemplary applications. Practical exercises with learning algorithms serve to operationalize the knowledge.</p> <ul style="list-style-type: none"> <li>● Introduction (basic terms, learning problems, concepts, examples, representation)</li> <li>● Rule learning <ul style="list-style-type: none"> <li>○ Learning of individual rules (generalization and specialization, structured Hypothesis rooms, Version Spaces)</li> <li>○ Learning of rules-sets (Covering Strategy, evaluation of rules, Pruning, multi-class problems)</li> </ul> </li> <li>● Evaluation and cost-sensitive learning (Accuracy, X-Val, ROC curves, cost-sensitive learning)</li> <li>● Instance-based learning (kNN, IBL, NEAR, RISE)</li> <li>● Decision tree learning (ID3, C4.5, etc.)</li> <li>● Ensemble methods (bias/variance, bagging, randomization, boosting, stacking, ECOcs)</li> <li>● Pre-processing (Feature Subset Selection, discretization, sampling, data cleaning)</li> <li>● Clustering and Learning based on association rules (a priori)</li> </ul>				
<b>3</b>	<p><b>Qualification objectives / learning outcomes</b></p> <p>After successfully taking part, the students will be able to</p> <ul style="list-style-type: none"> <li>● understand and explain basic techniques of Data Mining and Machine Learning</li> <li>● independently use practical Data Mining systems and understand their strengths and weaknesses</li> <li>● critically assess new developments in this field</li> </ul>				

<b>4</b>	<b>Prerequisites for participation</b>
<b>5</b>	<b>Type of examination</b> Subject examination in written and oral form, 60-120 min. / 30 min.
<b>6</b>	<b>Prerequisites for Credit Points</b> Passing the module exam (100%)
<b>7</b>	<b>Grading</b> Standard  In this lecture, it is possible to take into account accompanying achievements, which – according to § 25 (2) of the 4th amendment of the APB and the rules adopted by the Department 20 on October 2, 2012 – can lead to an improvement in grades of up to 1.0.
<b>8</b>	<b>Applicability of the module</b> B. Sc. Computer Science M. Sc. Computer Science M. Sc. IT Security M. Sc. Autonomous Systems M. Sc. Visual Computing M. Sc. Internet and Web-based Systems M. Sc. Business Informatics B. Sc. Psychology in IT Joint B.A. Computer Science B. Sc. Sports Science and Computer Science M. Sc. Sports Science and Computer Science B. Sc. Information Systems Engineering Can also be applicable for other programs in the context of interdisciplinary offers.
<b>9</b>	<b>Literature</b> <ul style="list-style-type: none"> <li>● Mitchell: Machine Learning, McGraw-Hill, 1997</li> <li>● Ian H. Witten and Eibe Frank: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan-Kaufmann, 1999</li> </ul>
<b>10</b>	<b>Comment</b>

<b>Module description</b>					
<b>Module</b> Data Mining and Machine Learning					
<b>Module No.</b> 20-00-0052	<b>Credit Points</b> 6 CP	<b>Hours required</b> 180 h	<b>Self-study</b> 120 h	<b>Module duration</b> 1 Semester	<b>Recurrence frequency</b> Usually every winter semester
<b>Language</b> German and English			<b>Person responsible for the module</b> Coordinators Web, Knowledge and Information Processing		
<b>1</b>	<b>Courses of the Module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0052-iv	Data Mining and Machine Learning	6	Integrated course	4
<b>2</b>	<p><b>Content</b></p> <p>Due to the rapid development of information technology, there are increasing data volumes available – and they often contain implicit knowledge that, if it actually became known, could be of great economic or scientific significance. Data Mining is a research area that seeks to find potentially useful knowledge in large sets of data, and Machine Learning is one of the key technologies in this area.</p> <p>The lecture provides an introduction to the field of Machine Learning, with a special focus on the aspect of Data Mining. Methods from different paradigms of Machine Learning are presented with references to exemplary applications. Practical exercises with learning algorithms serve to operationalize the knowledge.</p> <ul style="list-style-type: none"> <li>● Introduction (basic terms, learning problems, concepts, examples, representation)</li> <li>● Rule learning <ul style="list-style-type: none"> <li>○ Learning of individual rules (generalization and specialization, structured Hypothesis rooms, Version Spaces)</li> <li>○ Learning of rules-sets (Covering Strategy, evaluation of rules, Pruning, multi-class problems)</li> </ul> </li> <li>● Evaluation and cost-sensitive learning (Accuracy, X-Val, ROC curves, cost-sensitive learning)</li> <li>● Instance-based learning (kNN, IBL, NEAR, RISE)</li> <li>● Decision tree learning (ID3, C4.5, etc.)</li> <li>● Ensemble methods (bias/variance, bagging, randomization, boosting, stacking, ECOCs)</li> <li>● Pre-processing (Feature Subset Selection, discretization, sampling, data cleaning)</li> <li>● Clustering and Learning based on association rules (a priori)</li> </ul>				



3	<p><b>Qualification objectives / learning outcomes</b></p> <p>After successfully taking part, the students will be able to</p> <ul style="list-style-type: none"> <li>● understand and explain basic techniques of Data Mining and Machine Learning</li> <li>● independently use practical Data Mining systems and understand their strengths and weaknesses</li> <li>● critically assess new developments in this field</li> </ul>
4	<p><b>Prerequisites for participation</b></p>
5	<p><b>Type of examination</b></p> <p>Subject examination in written and oral form, 60-120 min. / 30 min.</p>
6	<p><b>Prerequisites for Credit Points</b></p> <p>Passing the module exam (100%)</p>
7	<p><b>Grading</b></p> <p>Standard</p> <p>In this lecture, it is possible to take into account accompanying achievements, which – according to § 25 (2) of the 4th amendment of the APB and the rules adopted by the Department 20 on October 2, 2012 – can lead to an improvement in grades of up to 1.0.</p>
8	<p><b>Applicability of the module</b></p> <p>B. Sc. Computer Science  M. Sc. Computer Science  M. Sc. IT Security  M. Sc. Autonomous Systems  M. Sc. Visual Computing  M. Sc. Internet and Web-based Systems  M. Sc. Business Informatics  B. Sc. Psychology in IT  Joint B.A. Computer Science  B. Sc. Sports Science and Computer Science  M. Sc. Sports Science and Computer Science  B. Sc. Information Systems Engineering</p> <p>Can also be applicable for other programs in the context of interdisciplinary offers.</p>
9	<p><b>Literature</b></p> <ul style="list-style-type: none"> <li>● Mitchell: Machine Learning, McGraw-Hill, 1997</li> <li>● Ian H. Witten and Eibe Frank: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan-Kaufmann, 1999</li> </ul>
10	<p><b>Comment</b></p>

<b>Module description</b>					
<b>Module</b> General Computer Science II					
<b>Module No.</b> 20-00-0290	<b>Credit points</b> 6 CP	<b>Hours required</b> 180 h	<b>Self-study</b> 120 h	<b>Module duration</b> 1 Semester	<b>Recurrence frequency:</b> every other semester
<b>Language</b> German			<b>Person responsible for the module</b> Vice Chairperson, Academic Affairs		
<b>1</b>	<b>Courses of the module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0290-iv	General Computer Science II	0	Integrated course	4
<b>2</b>	<p><b>Content</b> In this course, students will learn the basic algorithms and data structures of Computer Science, based on advanced concepts of the Java programming language.</p> <p>Repetition of basic Java-knowledge:  * Variables, types, classes, program flow  * Inheritance, abstract classes, interfaces  * Arrays and collections</p> <p>Advanced knowledge:  * Graphical User Interfaces  * Input / Output  * Error handling and exceptions</p> <p>Algorithms and data structures:  * Recursion  * Sorting algorithms  * Batches, lists, queues  * Searches  * Trees and graphs</p>				
<b>3</b>	<p><b>Qualification objectives / learning outcomes</b> After attending the event, students will be able to</p> <ul style="list-style-type: none"> <li>- develop extensive programs in Java</li> <li>- to independently use the basic algorithms and data structures of Computer Science</li> <li>- to estimate the advantages and disadvantages of elementary algorithms with regard to complexity and execution time</li> </ul>				
<b>4</b>	<p><b>Prerequisites for participation</b> Recommended:</p> <ul style="list-style-type: none"> <li>- General Informatics I or</li> <li>- basic programming skills</li> <li>- basic knowledge in Computer Science</li> <li>- experience in working with computers</li> </ul>				

5	<p><b>Type of examination</b> Module-accompanying examination:</p> <p>* [20-00-0290-iv] (Subject examination in written and oral form, Standard)</p>
6	<p><b>Prerequisites for Credit Points</b> Passing the module exam (100%)</p>
7	<p><b>Grading</b> Module-accompanying examination</p> <p>* [20-00-0290-iv] (Subject examination in written and oral form, weighting: 100%)</p>
8	<p><b>Applicability of the module</b> B. Sc. Computer Science M. Sc. Computer Science Can also be applicable for other courses of study</p>
9	<p><b>Literature</b> Java lernen mit BlueJ: Eine Einführung in die objektorientierte Programmierung David J. Barnes, Michael Kölling Pearson Studium 4., aktualisierte Auflage, 2009 ISBN-13: 978-3-8689-4001-5</p> <p>Algorithmen in Java Robert Sedgewick Pearson Studium 3. überarbeitete Auflage, 2003 ISBN-13: 978-3-8273-7072-3</p> <p>Einführung in die Programmierung mit Java Robert Sedgewick, Kevin Wayne Pearson Studium 1. Auflage, 2011 ISBN-13: 978-3-8689-4076-3</p>
10	<p><b>Comment</b></p>

<b>Module description</b>					
<b>Module</b> General Computer Science I					
<b>Module No.</b> 20-00-0304	<b>Credit Points</b> 6 CP	<b>Hours required</b> 180 h	<b>Self-study</b> 150 h	<b>Module duration</b> 1 Semester	<b>Recurrence frequency:</b> every other semester
<b>Language</b> German			<b>Person responsible for the module</b> Vice Chairperson, Academic Affairs		
<b>1</b>	<b>Courses of the module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0304-iv	General Computer Science I	0	Integrated course	2
<b>2</b>	<b>Content</b> <ul style="list-style-type: none"> <li>• Brief introduction to Computer Science</li> <li>• Introduction to working with computers</li> <li>• Introduction to programming (KarelJ, Java, or similar)</li> <li>• Binary numeric and information representation</li> <li>• Elementary logical and arithmetic operations</li> <li>• Von Neumann computer architecture</li> <li>• Elementary concepts of operating systems</li> <li>• Basics of computer networks</li> </ul> <p>The lecture will be accompanied by programming exercises.</p>				
<b>3</b>	<b>Qualification objectives / learning outcomes</b> <ul style="list-style-type: none"> <li>• Basic knowledge concerning the most important concepts of Computer Science</li> <li>• Practical use of computers</li> <li>• Basic programming skills</li> <li>•</li> </ul>				
<b>4</b>	<b>Prerequisites for participation</b>				
<b>5</b>	<b>Type of exam</b> Module-accompanying examination  * [20-00-0304-iv] (Subject examination in written and oral form, Standard)				
<b>6</b>	<b>Prerequisites for Credit Points</b> Passing the module exam (100%)				
<b>7</b>	<b>Grading</b> Module-accompanying examination  * [20-00-0304-iv] (Subject examination in written and oral form, weighting: 100%)				

---

---

<b>8</b>	<b>Applicability of the module</b> B. Sc. Computer Science M. Sc. Computer Science Can also be applicable for other courses of study
<b>9</b>	<b>Literature</b> Literatur: David J. Barnes und Michael Kölling, Java lernen mit BlueJ: Eine Einführung in die objektorientierte Programmierung, Pearson Studium; 5te Auflage (1. Januar 2013), ISBN 3868949070
<b>10</b>	<b>Comment</b>

<b>Module description</b>					
<b>Module</b> Foundations of Language Technology					
<b>Module No.</b> 20-00-0546	<b>Credit Points</b> 6 CP	<b>Hours required</b> 180 h	<b>Self-study</b> 120 h	<b>Module duration</b> 1 Semester	<b>Recurrence frequency:</b> usually every winter semester
<b>Language</b> German			<b>Person responsible for the module</b> Coordinators Web, Knowledge and Information Processing		
<b>1</b>	<b>Courses of the Module</b>				
	<b>Course No.</b>	<b>Course title</b>	<b>Workload (CP)</b>	<b>Type of lecture</b>	<b>Contact hours per week</b>
	20-00-0546-iv	Foundations of Language Technology	6	Integrated course	4
<b>2</b>	<p><b>Content</b></p> <p>The lecture offers an introduction to the key points of view, problems, methods, and techniques of automatic language technology – based on the example of the programming language Python.</p> <p>Key topics:</p> <ul style="list-style-type: none"> <li>• Language technology / Natural language processing (NLP)</li> <li>• Tokenization</li> <li>• Segmentation</li> <li>• Speech recognition</li> <li>• Corpora</li> <li>• Statistical analysis</li> <li>• Machine learning</li> <li>• Categorization and classification</li> <li>• Information extraction</li> <li>• Introduction to Python</li> <li>• Data structures</li> <li>• Structured programming</li> <li>• Working with files</li> <li>• Use of libraries</li> <li>• Program library NLTK</li> </ul> <p>The lecture is based on the class library NLTK for Python – which provides a powerful toolbox to address the theoretical methods in an exploratory and problem-oriented manner, without requiring extensive programming knowledge.</p>				

3	<p><b>Qualification objectives / learning outcomes</b></p> <p>After attending the lecture, the students will be able to</p> <ul style="list-style-type: none"> <li>● define the basic terminology of automatic language technology,</li> <li>● name and explain essential questions in this area</li> <li>● explain and implement simple Python programs</li> <li>● transfer the discussed methods and techniques to concrete application scenarios of text comprehension</li> <li>● critically assess their knowledge and their limits.</li> </ul>
4	<p><b>Prerequisites for participation</b></p> <p>Recommended:</p>
5	<p><b>Type of exam</b></p> <p>Module-accompanying examination</p> <p>(Subject examination in written and oral form, Standard)</p>
6	<p><b>Prerequisites for Credit Points</b></p> <p>Passing the module exam (100%)</p>
7	<p><b>Grading</b></p> <p>Standard</p>
8	<p><b>Applicability of the module</b></p> <p>B. Sc. Computer Science  M. Sc. Computer Science  M. Sc. IT Security  M. Sc. Internet and Web-based Systems  M. Sc. Business Informatics  B. Sc. Psychology in IT  Joint B.A. Computer Science  B. Sc. Sports Science and Computer Science  M. Sc. Sports Science and Computer Science</p> <p>Can also be applicable for other programs in the context of interdisciplinary offers.</p>
9	<p><b>Literature</b></p> <p>Steven Bird, Ewan Klein, Edward Loper: Natural Language Processing with Python, O'Reilly, 2009. ISBN: 978-0596516499. <a href="http://www.nltk.org/book/">http://www.nltk.org/book/</a></p>
10	<p><b>Comment</b></p>