

**İTÜ**  
**LİSANSÜSTÜ DERS KATALOG FORMU**

(graduate Course Catalogue Form)

Dersin Adı		Course Name		
Hesaplamalı Bilim ve Mühendislik'te Veri Yönetimi		Data Management in Computational Science and Engineering		
Kodu (Code)	Yarıyıl (Semester)	Kredisi (Local Credits)	AKTS Kredisi (ECTS Credits)	Ders Seviyesi (Course Level)
HBM8XXE	Güz/Bahar (Fall/Spring)	3+0 (Kredisiz)	7.5	Yüksek Lisans/Doktora (Master/Ph.D.)
<b>Lisansüstü Program (Graduate Program)</b>	Hesaplamalı Bilim ve Mühendislik Yüksek Lisans ve Doktora Programı (Computational Science and Engineering Masters and PhD Programme)			
<b>Dersin Türü (Course Type)</b>	Hazırlık (Preparatory)	<b>Dersin Dili (Course Language)</b>	İngilizce/Türkçe (English/Turkish)	
<b>Dersin İçeriği (Course Description)</b>	Bilimsel veri yönetimine giriş, Verinin temelleri, Verinin tanımlanması, Veri gösterimi, İlişkisel veri sistemlerine giriş, Taşınabilir veri tipleri, Ara değer kestirimi, Veri dönüşümü ve analizi, Büyük veri setlerinin işlenmesi ve yönetimi, Bilimsel veri ve ontolojiler			
<u>30-60 kelime arası</u>	Introduction to scientific data management, Fundamentals about data, Describing datasets, Data representation, Introduction to relational databases, Portable data formats, Data interpolation, Data transformation and analysis, Processing and managing large datasets, Scientific data and ontologies			
<b>Dersin Amacı (Course Objectives)</b>	<ol style="list-style-type: none"><li>Öğrenciye bilimsel veri ve yönetimi ile ilgili temelleri öğretmek</li><li>Farklı veri tiplerinin gösterimi için kullanılacak veri yapılarını öğretmek</li><li>Büyük veri setlerinin işlenmesi ve yönetilmesi ile ilgili detaylı bilgi vermek</li><li>Veri, önveri ve köken bilgisi gibi veri ile ilişkili kavramları tanıtmak</li><li>Bilgi paylaşımı ve kavramları ile ilgili bilgi kazandırmak</li></ol>			
<u>Maddeler halinde 2-5 adet</u>	<ol style="list-style-type: none"><li>To give the student a basic understanding of scientific data and management</li><li>To give the student a basic understanding of scientific data types to represent data</li><li>To give the student detailed information about processing and managing large datasets</li><li>To introduce the concepts of data, metadata and provenance information</li><li>To gain intermediate level knowledge and skills in the field of knowledge sharing</li></ol>			
<b>Dersin Öğrenme Çıktıları (Course Learning Outcomes)</b>	Bu dersi başarıyla tamamlayan yüksek lisans/doktora öğrencileri aşağıdaki konularda bilgi, beceri ve yetkinlik kazanırlar;			
<u>Maddeler halinde 4-9 adet</u>	<ol style="list-style-type: none"><li>Bilimsel veri ve yönetimi hakkında temel bilgi edinmesi</li><li>Bilimsel veri gösterimi teknikleri hakkında bilgi sahibi olması</li><li>İlişkisel veri tabanlarının tasarımı ve geliştirilmesini uygulayarak öğrenmesi</li><li>Taşınabilir veri tiplerinin temel yapısı ve işleyişi hakkında bilgi</li><li>Bilimsel veri işleme, dönüşümü ve analizi alanlarında bilgi ve tecrübe kazanma</li><li>Büyük veri setlerinin işlenmesi ve yönetimi araçlarının tanınması ve öğrenilmesi</li><li>Bilgi paylaşımı ve ontolojiler ile ilgili temel bilgi</li></ol>			
	M.Sc./Ph.D. students who successfully pass this course gain knowledge, skill and competency in the following subjects;			
	<ol style="list-style-type: none"><li>To gain a background on basic definitions in scientific data and management</li><li>To gain a background on scientific data representation</li><li>Understand design of relational database systems</li><li>To gain a background on design and functioning of portable data formats</li><li>Be able to process, transform and analyze scientific data</li><li>Get to know and learn the various tools/libraries used for processing and managing large datasets</li><li>To gain a background on knowledge sharing and ontologies</li></ol>			

<b>Kaynaklar</b> (References) <i>En önemli 5 adedini belirtiniz</i>	<ol style="list-style-type: none"> <li>1. Shoshani, A. ve Rotem, D., 2009. Scientific Data Management: Challenges, Technology, and Deployment, Chapman and Hall/CRC, 590 pp</li> <li>2. Allemang, D., ve Hendler, J., 2011. Semantic Web for the Working Ontologist, Second Edition: Effective Modeling in RDFS and OWL, Morgan Kaufmann</li> <li>3. Wes McKinney, 2013. Python for Data Analysis, O'Reilly, 470 pp</li> <li>4. Miner, D. ve Shook, A., 2012. MapReduce Design Patterns, O'Reilly, 251 pp.</li> <li>5. Holmes, A., 2012. Hadoop In Practice, Manning Publications, 536 pp.</li> </ol>		
<b>Ödevler ve Projeler</b> (Homework & Projects)	5 Ödev		
<b>Laboratuvar Uygulamaları</b> (Laboratory Work)	5 Homeworks		
<b>Bilgisayar Kullanımı</b> (Computer Use)	Linux İşletim Sistemi, Python, netCDF, HDF, MySQL, Hadoop, RDF, OWL, Protege		
<b>Diğer Uygulamalar</b> (Other Activities)	Linux Operating System, Python, netCDF, HDF, MySQL, Hadoop, RDF, OWL, Protege		
<b>Başarı Değerlendirme Sistemi</b> (Assessment Criteria)	<b>Faaliyetler (Activities)</b>	<b>Adedi* (Quantity)</b>	<b>Değerlendirmedeki Katkısı, % (Effects on Grading, %)</b>
	Yıl İçi Sınavları (Midterm Exams)	1	30
	Kısa Sınavlar (Quizzes)	0	0
	Ödevler (Homework)	5	30
	Projeler (Projects)	0	0
	Dönem Ödevi/Projesi (Term Paper/Project)	0	0
	Laboratuvar Uygulaması (Laboratory Work)	0	0
	Diğer Uygulamalar (Other Activities)	0	0
	Final Sınavı (Final Exam)	1	40

Hafta	Konular	Dersin Çıktıları
1	Bilimsel Veri Yönetimine Giriş (veri yaşam döngüsü, temel kavramlar)	1
2	Verinin Temelleri (biçim, ölçek, tip, yapısal ve yapısal olmayan veri tipleri ve veri modelleri)	1,2
3	Verinin Tanımlanması (önveri tipleri ve standartlar, köken bilgisi, toplanması, araç ve standartlar)	1,2
4	Veri Gösterimi (veri yapıları – noktasal, örgü, örgü tipleri)	1,2
5	İlişkisel Veri Tabanlarına Giriş (ilişkisel veri modeli ve tasarlanması, veri yönetim sistemleri, MySQL)	2,3
6	Taşınabilir Veri Tipleri I (netCDF ve yapısı, özellikler, önveri ve sıkıştırma teknikleri)	1,2,4
7	Taşınabilir Veri Tipleri I (HDF ve yapısı, gruplar, özellikler, sıkıştırma ve netCDF-HDF dönüşümleri)	1,2,4
8	Veri İç Değerlendirme (1/2 ve 3 boyutta veri iç değerlendirme, iç değerlendirme tipleri – Spline, Bilinear, Bicubic, Komşu-Noktalar ve Korunumlu)	2,5
9	Veri Dönüşümü ve Analizi (temel kavramlar – ortalama, fark, RMSE, bağlılık, bilimsel verinin işlenmesi için çeşitli araçlar – CDO, NCO, ESMP vb.)	2,5
10	Büyük Veri Setlerinin İşlenmesi ve Yönetilmesi I (veri yönetim aşamaları ve havuzu, veri yönetim sistemleri – OpenDAP, Thredds vb.)	1,2,6
11	Büyük Veri Setlerinin İşlenmesi ve Yönetilmesi II (MapReduce sistemi ve temel kavramlar, Map ve Reduce operasyonları)	1,2,5,6
12	Büyük Veri Setlerinin İşlenmesi ve Yönetilmesi III (Hadoop, temel elemanlar, Hadoop dağıtık dosya sistemi – HDFS, Hadoop I/O)	1,2,5,6
13	Büyük Veri Setlerinin İşlenmesi ve Yönetilmesi IV (Hadoop ile MapReduce uygulaması geliştirme, Hadoop küme kurulumu)	1,2,5,6
14	Bilimsel Veri ve Ontoloji (semantik web, temel tanımlar, ontoloji tanımlama ve araçlar - Protege, ontoloji dilleri – RDF, QWL, SparQL)	7

## COURSE PLAN

Weeks	Topics	Course Outcomes
1	Introduction to Scientific Data Management (data lifecycle, basic concepts)	1
2	Fundamentals about Data (forms, scales, types, structured and unstructured data and models)	1,2
3	Describing Datasets (metadata types and standards, data provenance, collection mechanisms, tools and standards)	1,2
4	Data Representation (data structures – point, gridded etc., grid/discretization types)	1,2
5	Introduction to Relational Databases (The relational data model and design, database management systems, MySQL)	2,3
6	Portable Data Formats I (netCDF, structure, attributes, metadata and compression techniques)	1,2,4
7	Portable Data Formats II (HDF, structure, groups, attributes, compression and conversion between netCDF and HDF)	1,2,4
8	Data Interpolation (Interpolation in 1D/2D and 3D, types – Splines, Bilinear, Bicubic, Nearest-Neighbor and Conservative)	2,5
9	Data Transformation and Analysis (basic concepts - average, bias, RMSE, correlation, tools to process scientific data - CDO, NCO, ESMP etc., )	2,5
10	Processing and Managing Large Datasets I (data management tasks and repositories, data management systems – OpenDAP, Thredds etc.)	1,2,6
11	Processing and Managing Large Datasets II (MapReduce framework and basic concepts, Map and Reduce operations)	1,2,5,6
12	Processing and Managing Large Datasets III (Hadoop ecosystem, key elements, Hadoop Distributed Filesystem - HDFS, Hadoop I/O)	1,2,5,6
13	Processing and Managing Large Datasets IV (Developing a MapReduce application with Hadoop, setting up cluster)	1,2,5,6
14	Scientific Data and Ontologies (science and semantic web, basic definitions and tools – Protege, ontology description languages - RDF, OWL etc.)	7

## Dersin Hesaplamalı Bilim ve Mühendislik Programıyla İlişkisi

	Programın mezuna kazandıracığı bilgi, beceri ve yetkinlikler (programa ait çıktılar)	Katkı Seviyesi		
		1	2	3
i.	Lisans düzeyi yeterliliklerine dayalı olarak, Hesaplamalı Bilim ve Mühendislik Programındaki bilgilerini uzmanlık düzeyinde geliştirebilme ve derinleştirebilme (yeterli bilgi birikimi) ( <i>bilgi</i> ).			X
ii.	Alanının ilişkili olduğu disiplinler arası etkileşimi kavrayabilme ( <i>bilgi</i> ).			
iii.	Alanında edindiği uzmanlık düzeyindeki kuramsal ve uygulamalı bilgileri kullanabilme( <i>beceri</i> ).			
iv.	Alanında edindiği bilgileri farklı disiplin alanlarından gelen bilgilerle bütünleştirerek yorumlayabilme ve yeni bilgiler oluşturabilme ( <i>beceri</i> ).		X	
v.	Alanını ile ilgili karşılaşılan sorunları araştırma yöntemlerini kullanarak çözümlenebilme ( <i>beceri</i> ).			
vi.	Alanını ile ilgili uzmanlık gerektiren bir çalışmayı bağımsız olarak yürütebilme ( <i>Bağımsız Çalışabilme ve Sorumluluk Alabilme Yetkinliği</i> )			
vii.	Alanı ile ilgili uygulamalarda karşılaşılan ve öngörülemeyen karmaşık sorunların çözümü için yeni stratejik yaklaşımlar geliştirebilme ve sorumluluk alarak çözüm üretebilme ( <i>Bağımsız Çalışabilme ve Sorumluluk Alabilme Yetkinliği</i> ).			
viii.	Alanındaki güncel gelişmeleri ve kendi çalışmalarını, nicel ve nitel veriler ile destekleyerek, alanındaki ve alan dışındaki gruplara, yazılı, sözlü ve görsel olarak sistemli biçimde aktarabilme ( <i>İletişim ve Sosyal Yetkinlik</i> ).			
ix.	Bir yabancı dili en az Avrupa Dil Portföyü B2 genel düzeyinde kullanarak sözlü ve yazılı iletişim kurabilmek ( <i>İletişim ve Sosyal Yetkinlik</i> ).		X	
x.	Alanının gerektirdiği düzeyde bilgisayar yazılımı ile birlikte bilişim ve iletişim teknolojilerini ileri düzeyde kullanabilme ( <i>İletişim ve Sosyal Yetkinlik</i> )			X
xi.	Alanı ile ilgili verilerin toplanması, yorumlanması, uygulanması ve duyurulması aşamalarında toplumsal, bilimsel, kültürel ve etik değerleri gözeterek denetleyebilme ve bu değerleri öğretebilme ( <i>Alana Özgü Yetkinlik</i> ).			X
xii.	Alanı ile ilgili konularda strateji, politika ve uygulama planları geliştirebilme ve elde edilen sonuçları, kalite süreçleri çerçevesinde değerlendirebilme ( <i>Alana Özgü Yetkinlik</i> ).			
xiii.	Alanında özümledikleri bilgiyi, problem çözme ve/veya uygulama becerilerini, disiplinlerarası çalışmalarda kullanabilme ( <i>Alana Özgü Yetkinlik</i> ).			
xiv.	Hesaplamalı Bilim ve Mühendislik Programında, kendi çalışmalarını, alanındaki uluslararası platformlarda, yazılı, sözlü ve/veya görsel olarak aktarabilme ( <i>Alana özgü yetkinlik</i> ).			

1: Az, 2. Kısmi, 3. Tam

## Relationship between the Course and Computational Science and Engineering Program

	Program Outcomes	Level of Contribution		
		1	2	3
i.	Developing and intensifying knowledge in the Computational Science and Engineering program's area, based upon the competency in the undergraduate level (sufficient knowledge) ( <i>knowledge</i> ).			X
ii.	Grasping the inter-disciplinary interaction related to one's area ( <i>knowledge</i> ).			
iii.	The ability to use the expert-level theoretical and practical knowledge acquired in the area ( <i>skill</i> ).			
iv.	Interpreting and forming new types of knowledge by combining the knowledge from the area and the knowledge from various other disciplines ( <i>skill</i> ).		X	
v.	Solving the problems faced in the area by making use of the research methods ( <i>skill</i> ).			
vi.	The ability to carry out a specialistic study related to one's area independently. ( <i>Competence to work independently and take responsibility</i> ).			
vii.	Developing new strategic approaches to solve the unforeseen and complex problems arising in the practical processes of one's area and coming up with solutions while taking responsibility ( <i>Competence to work independently and take responsibility</i> ).			
viii.	Systematically transferring the current developments in the area and one's own work to other groups in and out of the area; in written, oral and visual forms ( <i>Communication and Social Competency</i> ).			
ix.	Proficiency in a foreign language –at least European Language Portfolio B2 Level- and establishing written and oral communication with that language ( <i>Communication and Social Competency</i> ).		X	
x.	Using the computer software together with the information and communication technologies efficiently and according to the needs of the area ( <i>Communication and Social Competency</i> ).			X
xi.	Paying regard to social, scientific, cultural and ethical values during the collecting, interpreting, practicing and announcing processes of the area related data and the ability to teach these values to others ( <i>Area Specific Competency</i> ).			X
xii.	Developing strategy, policy and application plans concerning the subjects related to the area and the ability to evaluate the end results of these plans within the frame of quality processes ( <i>Area Specific Competency</i> ).			
xiii.	Using the knowledge and the skills for problem solving and/or application (which are processed within the area) in inter-disciplinary studies ( <i>Area Specific Competency</i> ).			
xiv.	In the Computational Science and Engineering program, the ability to present one's own work within the international environments orally, visually and in written forms ( <i>Area Specific Competency</i> ).			

1: Little, 2. Partial, 3. Full

**NOT-1: Ders ile ilgisi olmayan çıktıların boş bırakılması gerekmektedir.**

**NOT-2: Yukarıdaki tabloda işaretlenen katkı seviyeleri tüm programlar için minimum seviyelerdir. Ancak ilgili programın yapısına göre katkı seviyeleri artırılabilir.**