UC3M R&D FOR INNOVATION
big data

IDENTIFICATION OF BIG DATA RESEARCH ACTIVITY, TECHNOLOGIES, PATENTS, INFRASTRUCTURE AND OTHER CAPABILITIES EXISTING IN UC3M
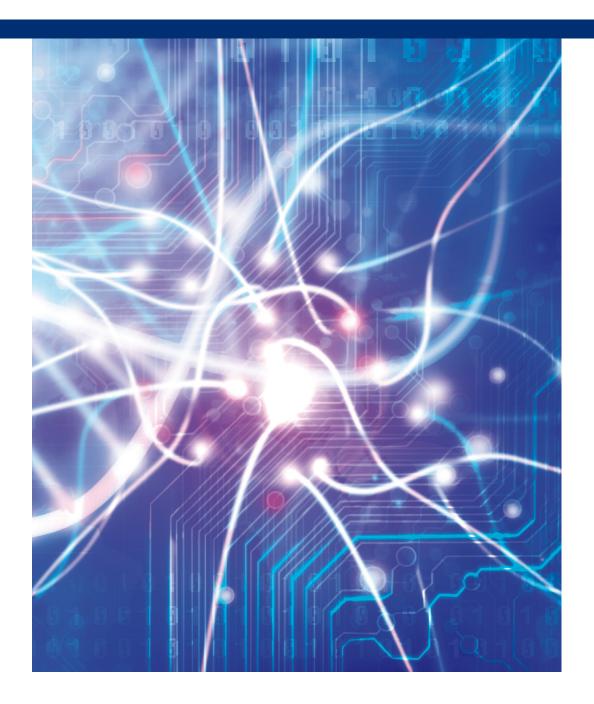
uc3m | **Universidad Carlos III de Madrid**

Vicerrectorado de Política Científica

Servicio de Apoyo al Emprendimiento y la Innovación

The Entrepreneurship and Innovation Support Service (SEI) of the Universidad Carlos III de Madrid wants to present the potential of the university in this "knowledge map" through the research areas developed in the frame of R & D projects, both national and international, patents and other results of UC3M investigators, in Big Data.

The global knowledge obtained, the experience of collaborating with the industry, the existence of infrastructures and proper laboratories and, above all, the multidisciplinary nature of UC3M are characteristics that provide an added value so that our support towards the innovation of institutions, big companies and SMEs has an integral quality.

We invite you to deepen the knowledge of the UC3M and to collaborate in new R & D and innovation projects.

Entrepreneurship and Innovation Support Service
Universidad Carlos III de Madrid

Contact:
comercializacion@uc3m.es

Update date
January 2016

Index

R&D INSTITUTE	DESCRIPTION	OBJECTIVES	LINES OF RESEARCH / RESEARCH PROJECTS	APPLICATION AREA
RESEARCH INSTITUTE UC3M-SANTANDER OF FINANCIAL BIG DATA				
Research Institute UC3M-Santander of Financial Big Data **Director: Daniel Peña Sánchez de Rivera**	The Research Institute UC3M-Santander of Financial Big Data was established in May 2015 as a joint venture between the Universidad Carlos III de Madrid and the Santander Group to promote interdisciplinary research in analyzing Big Data with emphasis on financial applications. The Institute has an international outlook and aims to produce high quality research that has a strong impact on the way Big Data can contribute to make better decisions.	The objectives of the Institute in research, teaching and technology transfer are: • Conducting interdisciplinary basic research in Analytics for Big Data pooling expertise in Statistics, Applied Mathematics, Computer Science and Maching learning and promoting collaboration across different fields. • Conducting applied interdisciplinary research oriented to financial applications. This research may lead to "spin offs" or "start ups" companies. • Offering education in Big Data Analysis through the Master in Big Data and several short courses in different fields of Data Analysis and Data Management. Also, the Institute will provide in-house training for organizations upon request. • Working with private and public partners in providing solutions to their problems in handling, storing and analyzing Big Data sets.	• Analytics in Social Networks • Artificial Intelligence • Bayesian inference in high dimension • Behavioral change and Big Data • Clustering in high dimension • Data Visualization • Deep Learning • Dimension reduction • Factor models in high dimension • Functional data analysis • Image analysis • Multidimensional Time series • Multivariate risk measures • Machine Learning • Outlier detection and robust estimation in large data sets • Optimization In high dimension • Parallel computing • Statistical Learning • Stochastic processes for high dimension	• Finance • Business • Economy • Energy • Industry • Medicine • Genomics • Social Networks • Marketing • Actuarial data • Health • Technological prospective

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
BIOENGINEERING AND AEROSPACE ENGINEERING				

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
Aerospace Engineering Research Group **IP: Eduardo Ahedo**	The Aerospace Engineering Research Group, includes professors specialized in areas related to Aeronautics and Space sciences, including computational and experimental Aerodynamics, Structural Dynamics, Flight Dynamics, Advanced Manufacturing, Navigation and Control, Plasmas and Propulsion.	• *Turbulent flows* To study turbulent flows, both in simulations and in experiments, it is necessary to store and analyze large datasets of up to 1TB. The analysis of such databases includes statistical analyses, pattern recognition, etc. which must be done by means of parallel codes (decomposition of domains into blocks, and/or task distribution). The group has also capabilities related to the extraction of reduced-order models for complex phenomena and dynamical systems. • *Image Processing and Analysis* There are a number of relevant experimental techniques in aerodynamics that use optical methods and require large-scale image processing and analysis. Existing methodologies are used, and *ad hoc* techniques are also developed for specific applications. The main methodologies used are: · Image processing. Pattern recognition and tracking · Data mining techniques: Principal Component Analysis, Proper Orthogonal Decomposition, Dynamic Mode Decomposition · Compressed sensing	• AFDAR (Advanced Flow Diagnostics for Aeronautical Research) *Funding: European Union (FP7/2007-2013)* • COTURB, Coherent Structures in Wall-Bounden Turbulence *Funding: European Union (H2020)* • Numerical and Experimental Research on the Unsteady Aerodynamics of Flapping Wings. *Funding: National R&D Plan* • Analysis and Optimization of aircraft trajectories under the effects of METeorological uncertainties (OPTMET) *Funding: National R&D* • Experiments over a flapping airfoil with an actuated Trailing Edge Flap *Funding: TU Delft* • Stochastic Optimal Control towards Enhanced Predictability of four-dimensional Trajectories using of Weather Ensemble Prediction Forecasts *Funding: Eurocontrol trough HALA! Research Network (SESAR Joint Undertaking (WP-E) - European Commission FP-7 Program)* • Meteorological Uncertainty Management for Trajectory Based Operations (TBO-MET)	Most of the data on turbulent flows and of unsteady aerodynamics is produced by means of parallel direct numerical simulation codes developed by researchers at the Bioengineering and Aerospace Engineering Department. Databases generated by they colaborate groups (UPM, KIT, UW, etc.) are sometimes used. The images and experimental data analyzed are the results of laboratory experiments and are processed mainly with codes and techniques developed by researchers in the department Many activities are carried out in collaboration with other universities and research International Research Centers (TU Delft, KTH Stockholm etc.). In the field of navigation and air traffic control, large-scale air traffic data, airplane paths (data from FDR, radar), structure of the air space (e.g., sectors, airways, waypoints, airports) are used. Meteorological data in GRIB format is also used.

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
		BIOENGINEERING AND AEROSPACE ENGINEERING		
Aerospace Engineering Research Group **IP:** Eduardo Ahedo		Techniques for reducing noise propagation and the estimation of uncertainty are also developed. • *Air Navigation and control* · Algorithms to calculate optimal trajectories (deterministic and stochastic) · Trajectory prediction algorithms. · Land-side (flight dispatching) and air-side (FMS) trajectory integration/ synchronization · Processing and integration of the atmospheric phenomenology in trajectory calculation · Environmental impact analysis of aviation, including contrails · Air traffic analysis and Air Traffic Management (ATM) performances (safety, capacity, flight effectiveness and environment)	*SESAR Joint Undertaking (Exploratory Research) - European Commission H2020 Program* • Advisory Services in Air Traffic Management *Private funding*	**Application Area:** The applications are found for the most part in branches of aerospace engineering: • The aerodynamic design of flapping wings, wind turbines, fluid devices and propulsion systems are amongst the applications of computational and experimental aerodynamics • The calculation of paths subjected to meteorological uncertainty and the analysis of performance indicators of the ATM management system stand out among the applications of navigation and air control

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Computer, Communications and Systems Architecture (ARCOS) **IP: Jesús Carretero**	The Computer Architecture, Communications and Systems group (ARCOS) has as its goal to research and develop hardware and software systems in the areas of real-time and embedded systems, high-performance computing, high productivity computing (Cloud and Grid), storage systems, and dependable systems. Their research activities are oriented to tools and methods for software development, high performance computing and data exploitation and management.	• High performance computing: · Scalable massive data management · Cloud and grid computing · Parallel file systems • Distributed and parallel systems: · High performance data retrieval and transmission system · Social network data analysis · Peer to peer systems **Technological Offer:** • Energy efficient data management mechanisms • Development and optimization data intensive applications to gather and process data coming scanners, topographies, X-ray, and other medical devices. ARCOS is specialist on data analysis and representation in quasi-realtime, using a variety of processing architectures such as multicore, GPGPU, Intel Xeon Phi.	• Scalable data management techniques for high-end computing system *Funding: R&D National Plan* • TEALES: New scalable storage techniques for high-performance computing *Funding: R&D National Plan* • Scalable Input/Output techniques for high-performance distributed and parallel computing environments *Funding: R&D National Plan* • FSCLOUDS: Sustainble federated storage clouds for massive data sharing communities. *Funding: R&D National Plan* • Ultra-platform: Towards Sustainable Development by an Ultra-scalable Computational Platform *Funding: UC3M* • NESUS:Network for Sustainable Ultrascale Computing *Funding: UC3M*	• Web • Data base • Social networks **Application Area:** • Smart Cities • Medicine / Health • Intelligent Transport Systems

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
		COMPUTER SCIENCE DEPARTMENT		
Evolutionary Computation and Neural Networks (EVANNAI) **IP: Pedro Isasi**	Evolutionary Computation and Neural Networks (EVANNAI) is formed by a team that carries out basic and applied research in the fields of evolutionary computation and data mining/machine learning. The group uses evolutionary computation techniques (genetic algorithms, evolution strategies, particle swarms, ...) to solve optimization problems and to deal with typical tasks of data mining, such as classification, time series prediction, clustering, relevant variable selection, etc.	• Data Mining/Machine Learning. **Technological Offer:** • Artificial intelligence software for data optimization, prediction and analysis	• MEMENTO: Platform for Global Availability of Big Data in the Cloud *Funding: R&D National Plan* • Strategic Action in Learning with Biological Inspiration *Funding: UC3M* • CEC-MP (Evolutionary Computation for Classification of Data Mining). *Funding:R&D Regional Plan* • CibMin (Bioinspired Computation for Data Mining). *Funding: R&D Regional Plan* • OE-FIN (Evolutionary Optimisation for Financial Assets Management). *Funding: R&D Regional Plan*	**Application Area:** • Data Structures • Telecommunication systems optimization • Robotics. • Economy and Finance. • Brain Computer Interface • Games

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Group of Advanced Databases (LABDA) **IP: Paloma Martínez**	The Advanced Databases Group, LABDA, is composed by a multidisciplinary team of PhDs and professionals in the field of information technologies. This group is part of Computer Science and Engineering Department of Universidad Carlos III de Madrid. We participate in R & D projects related to human language technologies, information retrieval and extraction in several application domains, question answering systems, usability and accessibility of user interfaces.	• Information Extraction and Retrieval in big data environments in biomedical and health domains with emphasis in pharmacovigilance • Machine learning –based and dicitionary –based algorithms for Named Entities Recognition (NER) and relation Extraction • Aplication of Natural Language Processing (NLP) to Electronic Health Records analysis • Application of NLP to social media monitoring • Methodological frameworks for developing accessible web applications • Accessible User Interfaces • Accessibility in elearning environments • Accessibility in eGovernment. **Technological offer:** • Designing and developing solutions for the treatment and exploitation of unstructured information (text) • The group employs sound methodologies for software development and the latest technologies in databases, data storing, data mining as well as linguistic technology for the processing of unstructured information and information access	• TrendMiner: Large-scale, Cross-lingual Trend Mining and Summarisation of Real-time Media Streams (http://www.trendminer-project.eu/) *Funding. Comisión Europea* • MULTIMEDICA: Extracción de Información Multilingüe en Sanidad y su aplicación a documentación divulgativa y científica *Funding. Ministerio Ciencia e Innovación* • BUSCAMEDIA: Hacia una adaptación semántica de medios digitales multired-multiterminal *Funding: CDTI* • ISSE: Interoperabilidad basada en Semántica para la Sanidad Electrónica *Funding: Ministerio de Industria*	• Web and social media • Databases • Medline **Application Area:** • Medicine/Health • Document Management and Knowledge Management • Treatment and exploitation of structured and unstructured information

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
<u>Applied Artificial Intelligence Group (GIAA)</u> **IP: José Manuel Molina, Jesús García Herrero**	The Applied Artificial Intelligence Group (GIAA) is formed by a team of physicists, telecommunications and information engineers who are nationally renowned due to their ability to solve engineering problems by incorporating the most novel artificial intelligence techniques: automatic learning, evolutionary computation, data analysis, multi-objective optimization, fuzzy systems and intelligent agents. The group has a long work history providing consultancy to companies and developing customized solutions for prediction, optimization, data fusion and signal and image processing.	• Automatic Learning and Data Mining Techniques • Agents and Multi-agent Systems: web, information retrieval, electronic commerce, sensor management • Contextual Information and Data Fusion Systems **Technological offer:** • Optimization, Prediction and Data Analysis Software • Extraction of behaviors from the intelligent data analysis • Design of Data Mining tools using intelligent data retrieval techniques	• European Concerted Research Action designated as COST Action IC1406: High-Performance Modelling and Simulation for Big Data Applications (cHiPSet) *Funding: European Cooperation in the field of Scientific and Technical Research. COST IC1406* • Integration of fusion and interpretation techniques to develop services based on estimates of activity in smart spaces *Funding: R&D National Plan* • Activity estimation techniques for services in smart spaces *Funding: R&D National Plan* • MEDIAN: Medical citizenship trough Data Analytics *Funding: UC3M* • Strategic Action in Data Fusion, Ambient Intelligence and Privacy *Funding: UC3M* • da-MEDIS: Big Data-analytics Medical Decision and Integrated Health Care Support System *Funding: UC3M* • Process Design and Implementation of Data Fusion SIGINT program *Private funding*	• Web • Databases **Application Area:** • Security and Defense • Transport • Medicine / Health

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Applied Artificial Intelligence Group (GIAA) IP: José Manuel Molina, Jesús García Herrero			• Design and implementation of Data Fusion Multi-sensor System *Private funding* • Intelligent Data Fusion in the Maritime Domain *Private funding* • Training Agreement "Data fusion" *Private funding*	

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				

Information and Knowledge Retrieval
 	
IP: Juan Llorens
	The Information and Knowledge Retrieval group, a part of the UC3M Reusing Knowledge Group, focuses its areas of interest in the recuperation and reuse of knowledge applied to different domains such as Systems Engineering, e-Health, documental and legal databases and social network analysis.	• Universal knowledge representation
		• Information retrieval
		• Traceability
		• Knowledge Quality management
		• Knowledge and information reuse
		• Knowledge Graphs
		• Data Lifecycle Management (acquisition, curation, reconciliation, exploitation)
		• Big Data programming:
		  · Algorithms and programming techniques for large data processing
		  · Predictive Analytics and data mining: deep learning and relational learning
		  · Real-time, stream and batch processing
		  · Intelligent monitoring and analysis techniques
		  · Natural language processing (entity reconciliation, concept extraction, etc.)
		  · Information and data fusion techniques
		  · Hybrid recommender systems
		• Infrastructure and architecture
		  · Cross-scale data architectures
		  · Large-scale hybrid architectures
		  · Cloud platforms and micro services
		  · Management, Governance and Synchronization +
			• AMASS - Architecture-driven, Multi-concern and Seamless Assurance and Certification of Cyber-Physical Systems
			Funding: H2020-ECSEL-2015-1-RIA-two-stage Nº 692474
			• CRYSTAL – Critical Systems Engineering Acceleration
			Funding: FP7 ARTEMIS nº 332830
			• Automatic generation of an oncology ontology
			• HPC-Europa2 Transnational Access programme. Stay and use of the SARA (Netherlands) infrastructure
			Funding: FP7 nº 228398.
			• Marie Curie Experienced Researcher position in Cloud Computing (RELATE-ITN). Monitoring of quality of service in cloud environments at SEERC (Greece)
			Funding: FP7 nº 264840.
			Funding: R&D National Plan
			• Development of a semantic information retrieval system based on metadata and contents.
			Funding: R&D National Plan
			• Research agreement between Radio Televisión Española (RTVE) and UC3M: "Information management, Big Data, Linked Data and application to the audiovisual sector"
			Funding: RTVE and UC3M
				• Web
				• Structured data and information
				• Social networks
				• Electronical databases
				• Existing databases
				
				Application Area:
				• Processes and management of software development
				• Document management
				• Content management
				• Business Process Management
				• Big Data applications:
				  · Monitoring and social network analysis. Cases study on Github and Twitter
				  · Real-time Recommender systems. Cases study on Tourism, Linkedin and videos

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Information and Knowledge Retrieval IP: Juan Llorens		· Large scale Knowledge management · CAP Theorem and Large data storage techniques and models (NoSQL systems: key-value, document oriented, column-family or graph) · Distributed query languages: Apache Pig, Google Big Query, etc. · Big Data Platforms · On-line and off-line data stream processing frameworks • Exploitation and Data-driven applications · Decision support systems · Complex event reasoning and processing of rare patterns · Sensor data streams: continuous queries and processing · Design-time analysis and prediction of complex systems variables · Smart data environments and applications · Social networks analysis: affective computing and sentiment analysis · Big social and semantic data analysis · (Dynamic) Visualization of large datasets · Sensor networks · Context-aware applications and personalization · Domains: e-Health, e-Government, Bioinformatics, Social networks, Cultural Heritage, Opinion mining, Journalism, Marketing, etc. +		

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Information and Knowledge Retrieval IP: Juan Llorens		**Technological offer:** · Techniques for automatic knowledge discovery and extraction · Systems traceability recovery · Algorithms for measuring knowledge quality		

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Control, Learning and Systems Optimization Laboratory (CAOS) **IP: Araceli Sanchis**	The Control, Learning and Systems Optimization Laboratory (CAOS) is formed by a team of professionals who are experts in the prediction, optimization, control of business services, processes and decisions based on data analysis, using advanced artificial intelligence technologies.	• Intelligent data analysis **Technological offer:** • Prediction of trends and process optimization based on intelligent data analysis • Design of intelligent data retrieval -Data • Mining- systems based on state of the art technologies • Forecasting Time Series by means of machine learning techniques. • Explicit representation of the key processes and the knowledge of complex organizations • Advanced tools for the automatic analysis of market data	Research Projects: • Strategic Action in Artificial Intelligence Applied to Systems Control *Funding: R&D Regional Plan*	Application Area: • Transport • Industry • Video games

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
COMPUTER SCIENCE DEPARTMENT				
Planning and Learning Group (PLG) **IP: Daniel Borrajo**	The Planning and Learning Group (PLG) is formed by a consolidated team of experts in developing effective and innovative software solutions for the automation of planning tasks and data analysis by means of advanced artificial intelligence technologies. PLG is one of the leading Spanish groups in automatic learning and task planning technologies. Their technology has won numerous international awards.	• Machine learning • Heuristic optimization • Decision support systems **Technological offer:** • Automatic optimization of the data analysis process (Data mining) • Tools for semantic retrieval from documents (including web pages) • Automation of dynamic planning and decision making processes • Optimization of processes from intelligent data analysis and trend or risk prediction • Design of intelligent data retrieval systems based on state-of-the-art technologies • Automatic data analysis for predicting commercial, financial trends, etc.	• CCI. Complexity, Cost and Change Impact Based on Models *Funding: European Space Agency (ESA).* • GLASS: Management of goals for long-term autonomy in smart cities *Funding: R&D National Plan* • onDroad: Platform for Digital Content Management and Tourism Services for Travelers *Funding: R&D National Plan (AVANZA)* • Knowledge modelling and planning technolgies for automation of data mining analysis *Private funding: Ericsson*	• Documents • Web pages • Sensors **Application Area:** • Industry processes • Networks and logistics systems • Transport • Tourism • Finances • Space • Marketing

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
		COMPUTER SCIENCE DEPARTMENT		
SoftLab **IP: Angel Garcia Crespo**	The SoftLab group has developed its work since 1999. Its works are related to: • The effective integration of systems through concurrent engineering, artificial intelligence and systems information management • Analysis, modeling and design of systems using principles of incremental development • Design and creation of tools by rapid prototyping systems • Design and development of mobile applications in environments of high productivity • Systems integration based on client-server architecture using Internet as a development platform • Development of accessible multimodal interfaces	• Semantic Web Technologies • Linked Data • Business information systems • App development **Technological Offer:** • Corpus generating system in social networks • Tool for analyzing and monitoring in social networks • Analysis system to obtain product recommendations from unstructured data • Analysis system to obtain predictions from unstructured data	• Strategic Action on Semantic Technologies, Ontologies and Artificial Intelligence for the analysis of social networks and open innovation *Funding: UC3M* • Consultancy for the development of a platform based on linked data technologies for the analysis and massive monitoring of information generated through social networks *Private funding* • Extractment 2.0: tweets extraction system based on api streaming *Private funding* • Twittiment: Corpus Management System for Analysis feelings on Twitter *Private funding*	• Web • Data bases • Social networks **Application Area:** • Computing • Recommender. The group has technology that allows to analyze large volumes of data for recommendations or predictions • Marketing. The group has developed a tool to monitor the effect of brands on social networks using Big Data

Universidad Carlos III de Madrid — uc3m

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
<u>Advanced Switching and Communication Systems (ADSCOM)</u> **IP: David Larrabeiti**	The Research Group on Advanced Switching and Communication Technologies (ADSCOM) is part of the Telematics Engineering Department at University Carlos III of Madrid. The Group is has expertise in switching and networking. The main areas of interest of the group are: • Network Architectures • Optical Access Metro and Core Networks • Protocol Design • Deep Packet Inspection	• Application of techniques of data mining and web mining in telematic networks • Design of distributed web crawlers, collection and analysis of social networks data and web platforms data • Processing of large volumes of data with Apache Hadoop and Spark ecosystems • Real-time analysis of large data sets, Spark Streaming, applied to sensor networks (Internet of Things IOT) • Analysis of traffic networks and massive catches of data in lawful interception • Use of the classic Machine Learning and Data Science with R and Python, supervised and unsupervised learning, variable selection and reduction of dimensions techniques and information theory • Information Retrieval and text and unstructured data mining techniques • Big data and privacy: value analysis of information for large companies on the Internet	• TYPES, capture and analysis of data on social networks (Facebook) *Funding: Euroean Union* • Massive analysis of Google Play and Android applications *Private Funding* • Data analysis in social networks, in colaboration with GISC UC3M Group *Funding: National R&D Plan* • INDECT, analysis of large volumes of data in lawful interception of communications *Funding: European Union*	• Social networks **Application Area:** • Communications • Security

+

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
Advanced Switching and Communication Systems (ADSCOM) IP: David Larrabeiti		**Technological Offer:** Integration of Internet information in big data analysis. Practical examples: • Analysis of the working market by combining data from LinkedIn, Infojobs and CareerBuilder • Urban data mining including the capture and analysis of data on real estate portals (idealist) and socio-economic activity (Yelp) and business • Ad-hoc analysis of large datasets • Optimization of clusters and parallelization for the analysis of large volumes of data		

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
NETCOM **IP: Arturo Azcorra**	The NETCOM group focus its research activity in networking and internet services.	• Online Social Networks (OSNs) • Online Advertising **Technological Offer:** • Analysis of large scale graphs • Analysis of OSN activity • Influential Users analysis in OSNs • Analysis and monitorization of events in OSNs • Analysis of Online Advertising Fraud • Privacy-preserving solutions in online advertising • Identity matching	• TYPES *Funding: European Union* • RECRED *Funding: European Union* • BRADE *Funding: R&D Regional Plan*	Typically we obtain the data to develop our research as well as our technological solutions. To this end we rely in sophisticated crawling tools to obtain data from the services under study (e.g., OSNs) **Application Area:** • Online Social Networks • Online Advertising

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
WEBTLAB (Web Technologies Lab) **IP: Carlos Delgado Kloos**	The WebTLab is part of the Telematic Applications and Services Group (GAST) at Telematics Engineering Department at UC3M. The activities of the WebTLab laboratory are focused on researching and developing technologies for building Web-based applications.	• Information extraction • Social network analysis • Distributed architectures for time-critical Big Data applications • Traffic data analysis for the purpose of a healthy and efficient driving • Machine learning **Technological solutions:** • Text mining on the Web • Text mining on social networks • Sensor data publishing • T-hoarder framework (Twitter data collection and analysis) • Analysis of efficiency: · Apache Map/reduce · Apache Hive · Apache Storm	• Healthy and Efficient Routes in Massive open-data basEd Smart cities: smart driving and semantic data handling *Funding: Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad* • JUNIPER (FP7) *Funding: European Commission* • Linked Data Network *Funding: R&D National Plan*	• Web • Social networks **Application Area:** • Smart cities • Media • Social networks • Telematics

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
GRADIENT (Educational Technology Lab) **IP: Carlos Delgado Kloos**	The Gradient lab focuses its research activities on educational technology. Its main research lines include, among others, educational data mining, learning analytics and data analysis in online platforms. The Gradient lab is part of the Telematic Applications and Services Group (GAST) at Telematics Engineering Department at Universidad Carlos III de Madrid.	• Inference of intelligent information from raw data in education • Prediction of useful indicators of the learning process • Evaluation of the learning process in educational platforms • Decision making in learning processes based on data • Determination of causes in on-line learning processes • Relationship mining between variables in e-learning platforms • Detection of the structure and grouping of users using clustering techniques • Adaptation and personalization in e-learning platforms **Technological Offer:** • Detectors of cognitive skills, meta-cogitive skills, motivation, emotions, etc. in educational platforms • Adaptive learning systems • Visualization tools of the learning process in educational platforms • Recommenders of educational resources	• SHEILA: Supporting HE to Incorporate Learning Analytics *Funding: European Comission* • RESET: Reformulating Scalable Educational Ecosystems *Funding: R&D National Plan* • SNOLA: Spanish Network of Learning Analytics *Funding: R&D National Plan* • mapaTIC: Improvement of learning based on projects using TICs in Education for Adults *Funding: R&D National Plan* • simLAP: Surgical Simulation platform. *Funding: R&D National Plan* • eMadrid: Research and Development of Educational Technologies in the Madrid Community Region *Funding: R&D Regional Plan* • Analysis, evolution, improvement proposals and development of the adaptive learning and learning analytics systems of the smartick platform *Private funding*	• Educational online platforms, including MOOCs or Learning Management Systems, among others • Activities done by users in any learning platforms connected to Internet **Aplication Area:** • Evaluation of online educational platforms and their learning processes • Actuation and decision making for the learning improvement • Selection of the best practices and tools for learning processes • Adaptive learning

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
TELEMATIC ENGINEERING DEPARTMENT				
Pervasive Computing Lab/ GAST **IP: Andrés Marín Lopez (Pervasive) / Carlos Delgado Kloos (GAST)**	The pervasive computing laboratory (PERVASIVE) of the Telematic Engineering Department contributes to the fields of cyber security, Internet of Things, Crowd Sensing and Smart Cities.	• Social Media variable analysis (crowdfunding Project success, emergency analisys and hot/cold realtime maps) from existing data sets and in real time. Tools for efficient data collection (on sensor curation and pre-processing) and processing • Critical infrastructure management, route planner and incident management for commercial drones **Technological Offer:** • Crowd Sensing platform with on sensor multitenancy management, real time data processing for hot maps and incident management	• INRISCO, Monitoring incidents in intelligent communities: security and mobility *Funding: R&D National Plan*	• Custom Sensors (polution), web scrapping and air route firehose **Application Area:** • Drone navigation control • Management of environmental variables • Analysis of emergencies

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
MATHEMATICS DEPARTMENT				
Complex Systems Interdisciplinary Group (GISC) **IP: Rodolfo Cuerno**	The Complex Systems Interdisciplinary Group (GISC) is a renowned research group with national and international visibility, whose research relates with situations or systems in which a large number of agents are in interaction (Complex Systems). GISC dominates the latest methodologies and the most useful tools of the Science of Complexity: statistical mechanics, stochastic processes and other statistical methods, computer simulations, big data analysis, experiments on behavior, etc.	• Statistical mechanics and non-linear dynamics • Evolutionary processes • Modeling in Ecology • Dynamics of socio-economic systems • Game theory and experiments on social behavior. • Social networks **Technological offer:** • Big data analysis • Complex network analysis • Models of social phenomena • Evolutionary models in biology • Analysis of the spanish stock-exchange market • Models of mobility in cities • Experimental study of social behavior	• Bridging the gap: from individual behavior to the socio-technical man (IBSEN) *Funding: FET Open, H2020 Program, UE* • Distributed global financial systems for society (DOLFINS) *Funding: FET Proactive, H2020 Program, UE* • Modeling and simulation of interface dynamics in soft and hard matter *Funding: R&D National Plan* • Mecánica estadística para "Big Data": Adquisición, Análisis y Modelización *Funding: R&D National Plan* • Modeling and simulation of complex systems *Funding: R&D Regional Plan* • Analyzing the influence of Twitter in political elections *Private funding (Twitter Spain S.L)* • Understanding social group formation and evolution in social networks *Private funding (XING AG)*	• Social networks • Data bases • In-house experiments • Citizen Science **Application Area:** • Economics • Biology • Sociology • Anthropology • Urban planning • Physics • Materials science • Nanotechnology • Polítics • Medicine • Public health • Marketing • Smart Cities

Universidad Carlos III de Madrid

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
SIGNAL AND COMMUNICATIONS THEORY DEPARTMENT				
Signal Processing and Learning Group **IP: Antonio Artés**	The Signal Processing and Learning Group focuses on research in the following areas: • Detection and classification of signals and images, including formulations parametric and nonparametric, centralized and distributed • Machine learning and image signal processing, including both design of classifiers and estimators as feature extraction and information for classification and estimation • Advanced techniques of statistical signal processing, including among others nonlinear filtering, Monte Carlo methods, Bayesian methods or signal decomposition • Information theory	• Signal and image detection and classification • Machine learning techniques • Advanced signal processing methods • Information Theory applications **Technological offer:** • Machine Learning and data mining • Data fusion	• Machine Learning for Personalized Medicine *Funding: European Comission* • Computational Inference in High Dimensional Random Complex Systems (COMPREHENSION) *Funding: R&D National Plan* • A new sequential Monte Carlo framework for tracking of nonlinear complex dynamical systems *Funding: Office of Naval Research*	• Databases **Application Area:** • Communications • Medicine • Social sciences

uc3m | Universidad Carlos III de Madrid

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
SIGNAL AND COMMUNICATIONS THEORY DEPARTMENT				
Machine Learning for Data Science (ML4DS) **IP: Jerónimo Arenas**	The Machine Learning for Data Science (MLS4DS) group credits experience of over 15 years in R&D, innovation and technology transfer in the field of machine learning and its applications to analysis and information processing.	• Machine Learning algorithms for Big Data analysis • Kernel methods • Feature selection, extraction, dimensionality reduction • Adaptive learning • Algorithms for topic discovery in document collections • Internet as a Data Source • Machine learning for Neuroimage • Smart Grid • Distributed and energy-aware learning in sensor networks **Technological Offer:** • Smart crawling of Internet domains • Analysis of document corpuses based on classification and topic models • Recommender systems based on Collaborative Filtering • Ad-hoc machine learning designs for Classification, Regression, clustering and novelty detection • Algorithms implementations in Python • D3 visualizations • Solutions for Big Data Machine Learning using Spark and MLLIB	Selection limited to last two years: • Automatic discovery of interpre-table features and metrics for machine learning *Funding: R&D National Plan* • Pilot Project using semantic latents for the watching of the TIC sector and development of Information Society *Funding: R&D National Plan SETSI)* • PRICAM-CM: Smart-Grid Program in the Community of Madrid *Funding: R&D Regional Plan* • Standardization service for the Data Base of domains ".es" *Funding: RED.ES* • Pilot Project for studying the viability of using Internet as a Data Source *Funding: RED.ES* • Recommender System for a mobile platform application *Funding: FUTURE SPACE S.A.*	• Web • Data bases • Text • Audio • Video **Application Area:** • Analysis of multimedia content • Energy • Biomedicine / Health • Sensor networks

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
		STATISTICS DEPARTMENT		
Statistics Department **Dirección: Rosa Elvira Lillo Rodríguez**	A multidisciplinary team made up of over 40 PhD's in the areas of Statistics, Econometrics and Operations Research. They use the most advanced methodologies and the most appropriate computational tools to solve modeling, forecast, data analysis and optimization problems. It comprises members of four multidisciplinary Groups: • Operations Research IP: Francisco Javier Nogales Martín • Statistical Modeling and Data Analysis IP: Rosa Elvira Lillo Rodríguez, Daniel Peña Sánchez de Rivera, Ismael Sánchez Rodríguez-Morcillo • Macroeconomic and Financial Prediction and Analysis IP: Antoni Espasa Terrades • Non-parametric and Intensive Computing Techniques. IP: Juan J. Romo Urroz	**Operations Research group:** • Optimization of probabilistic dynamic systems by means of mathematical programming methods • Stochastic programming • Non-linear optimization in very largescale problems • Stochastic combinatorial optimization **Statistical Modeling and Data Analysis group:** • Functional data analysis • Regression analysis • Multivariate analysis • Biostatistics • Bioinformatics • Statistical process control Reliability • Stochastic optimization • Stochastic processes • Time series • Re-sampling techniques - Semi-parametric techniques The **Macroeconomic and Financial Forecast and Analysis group** is particularly oriented towards forecast and diagnosis	**Operations Research Group** • Regularized optimization: new models and methods in the analysis of Big data *Funding: R&D National Plan* **Stastistical Modelling and Data Anaysis** • Advanced statistical methods for complex data *Funding: R&D National Plan* • Statistical decision methods based on knowledge *Funding: R&D National Plan* • Robust methods for statistical analysis of data with complex structure *Funding: R&D National Plan* • Stochastic models for the statistical treatment of complex data and risk measures *Funding: R&D Regional Plan* • Statistical procedures based on functional data and high-dimensional data with applications in finance and biostatistics *Funding: R&D Regional Plan-UC3M*	• Microarrays • Bibliometric Databases • Financial and economic Data bases • Image Databases **Application Area:** • Economy • Energy • Industry • Financial markets • Medicine • Genomics

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
		STATISTICS DEPARTMENT		
Statistics Department Dirección: Rosa Elvira Lillo Rodríguez		**Non-parametric and Intensive Computing Techniques in Statistics Group:** • Functional Data • Time Series • Bootstrapping Techniques • Finance • Genomics **Technological Solutions:** • Analytical and computational decision aid tools for agents in electric energy markets. • Electric energy price and demand forecasting tools based on time series methodologies. • Development of optimization models and methods under uncertainty for financial decision-making. • Optimization of the performance of logistic systems • Design of dynamic protocols for operating complex probabilistic systems (communication networks, sensor networks, production/inventory systems, etc.) while optimizing their performance	**Macroeconomic and Financial Prediction and Analysis Group** • Analysis and prediction of the distribution and consumption of electrical energy *Private funding* • Analysis and Prediction of Inflation. *Funding: R&D Regional Plan* • Predictions and macroeconomic analysis Private funding **Non parametric and Intensive Computing Techniques Group** • Analysis of very high dimension data in economics and business *Funding: R&D National Plan* • Statistical techniques to highly complex data in business and finance *Funding: R&D National Plan* • Non parametric and Intensive Computing Techniques in statistics *Funding: R& Regional Plan*	

+

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
STATISTICS DEPARTMENT				
Statistics Department ———————— **Dirección: Rosa Elvira Lillo Rodríguez**		**Scientific–Technical Services:** • Statistical models and data analysis in highdimensional problems: experience in processing data from Microarrays, Bibliometric Databases, Economic/financial Databases and Image Databases • Analysis with statistical Data Mining techniques including internally developed pattern recognition and classification problem techniques • Implementation of Intensive Computing statistical methods including the design of internal Bootstrap-type re-sampling methodologies with application in time series, or Bayesian techniques with a particular emphasis on the use of Gibbs-type sampling or of MCMC techniques • Macroeconomic forecast methodology • Development and resolution of Operations Research Models for optimizing decisions in different areas of application (networks, logistics, finance, energy, etc.) • Engineering and Medical Systems Reliability Analysis +		

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
STATISTICS DEPARTMENT				
Statistics Department ———— Dirección: Rosa Elvira Lillo Rodríguez		• Comparison of medical treatments or treatment of any other type by means of survival techniques • Stock exchange studies relating to portfolio optimization		

R&D GROUP	DESCRIPTION	LINES OF RESEARCH	RESEARCH PROJECTS	DATA SOURCE & APPLICATION
LIBRARY AND INFORMATION SCIENCE DEPARTMENT				
Technologies applied to information and documentation (TECNODOC) **IP: Antonio Hernández Pérez, José Tomás Nogales Flores**	The Group of Technologies for Information and Documentation (TECNODOC) is a group specialized in auditing and starting up projects related to the implementation of digital information systems, and having extensive knowledge in the creation and maintenance of a digital information system for companies, as well as in the application of technologies in archives, libraries and documentation centers.	• Metadata (Dublin Core, LOM, Premis, Mets **Technological offer:** • Creation of interoperability strategies for information exchange, in terms of accessibility, usability, metadata information architecture, portals, search engines and open archives • Analysis and processing of digital compilations and documents	• Development of a conceptual system recovery by semantic levels representing metadata schemas	• Web • Databases • Documents **Application Area:** • Libraries and documentation centers • Information systems