madrimasd.org

suplemento
notiweb

institute
IMdea
software

Science and technology for developing better software

Ciencia y tecnología para desarrollar mejor software

fundación para el conocimiento madri+d

El Instituto IMDEA Software: construyendo la estructura de la sociedad de la información

La informática se ha convertido en la base sobre la que se construye el tejido estructural de nuestra Sociedad

Manuel Carro Liñares
Profesor titular en la UPM y director del Instituto IMDEA Software
Associate professor at UPM and director of the IMDEA Software Institute

Escribir este texto para celebrar los quince años de la aparición de los Institutos IMDEA es una enorme satisfacción. Aprovecharé esta oportunidad para dar unas pinceladas sobre el papel que juega la investigación que realizamos en el Instituto IMDEA Software en la sociedad moderna y para introducir el resto de los artículos que conforman este número especial.

Los ordenadores tienen un papel destacado en nuestras vidas. Muchos de nosotros, si no la mayoría, usamos un ordenador como herramienta principal o al menos nos valemos de ordenadores para realizar nuestro trabajo. Con ellos nos comunicamos con nuestros amigos, compramos y pasamos un rato de ocio viendo películas o jugando. Usamos continuamente nuestros *smartphones*, que son sofisticados ordenadores equipados con un lector de SIM, una *app* para realizar llamadas telefónicas y una cámara. Pero pocos reparan en hasta qué punto los ordenadores, y por lo tanto la informática, juegan un papel esencial en la vida moderna. No es atrevido afirmar que la informática se ha convertido en la base sobre la que se construye el tejido estructural de nuestra sociedad.

Imaginemos, por un momento, que un hechizo hace que todo el software en el mundo se detenga. Todo lo demás funciona: las pantallas están encendidas, las líneas de comunicación están activas… pero, por alguna razón, los programas, simplemente, no se ejecutan. ¿Qué pasaría entonces?

Por supuesto, nuestros portátiles dejarían de funcionar, al igual que nuestros *smartphones*. Los televisores también se pararían: en la actualidad, un televisor es un pequeño ordenador con una gran pantalla. Los automóviles también se verían afectados: un automóvil moderno es un conjunto de muchos pequeños ordenadores y controladores conectados entre sí, ejecutando software sofisticado que recibe señales del entorno y reacciona a ellas actuando sobre los componentes del coche: frenos, luces, motor, ruedas, …

La red eléctrica también sufrirá. La transmisión de energía en las líneas eléctricas está regulada por programas que examinan el consumo, producción y precios, y tratan de encajarlos entre sí. Sin este control, los desajustes entre sus partes pueden causar apagones y problemas en las plantas de producción y distribución. Los hospitales, que almacenan la información de los pacientes en bases de datos, verían desaparecer su capacidad para funcionar normalmente cuando se perdiese acceso a estos datos.

The IMDEA Software Institute: Weaving the Fabric of the Information Society

Computer science has become the foundation on which the fabric of our society is built

I am truly happy to write this text to celebrate the fifteen years of the birth of the IMDEA Institutes. I will take this opportunity to describe, summarily, the role of the research we do at the IMDEA Software Institute and how it impacts modern society, and introduce the rest of the articles that make up this special issue.

Computers have a prominent role in our lives. Many, if not most, of us work primarily with a computer or use computers to get our daily job done. We use them to communicate with our friends, to shop, and to spend some leisure time watching movies or playing games. We continuously tap on the screen of smartphones, which are sophisticated computers equipped with a SIM reader, a phone app, and a camera. But few people realize up to which point computers, and therefore computer science, play an essential role in our society. It is not too bold to state that computer science has become the foundation on which the fabric of our society is built.

Let us imagine, for a moment, that a magical spell causes all the software in the world to stop. Everything else is just right: screens are on, communication lines are up… but, for some reason, programs just do not run. What would happen, then?

Of course, our laptops stop working, as well as our smartphones. TV sets also cease to function: these days, a TV is a small computer with a big screen. Cars will be hit hard: a modern car is mainly a set of many small computers and controllers linked together and executing sophisticated software that receives signals and commands from the environment and reacts to them acting on the physical components of the car: brakes, lights, engine, wheels, …

The power grid will also suffer. The transmission of power in the electricity lines is governed by programs that monitor consumption, production, and prices and try to match them. Without this control, mismatches can cause blackouts and problems in the production and distribution plants. Hospitals, which store the information on their patients in databases, will see their ability to function normally seriously impaired when the access to this data is lost.

This list can go on and on: in a modern society, there are computing devices behind the scenes practically everywhere, doing a silent job. Remaining unnoticed and eventually blending with the background is what one would expect of a truly useful

Esta lista puede seguir casi interminablemente: en una sociedad moderna hay dispositivos de computación trabajando silenciosamente entre bambalinas prácticamente en todas partes. Permanecer desapercibido y confundirse con el entorno es lo que cabría esperar de una tecnología verdaderamente útil: estar a nuestro servicio sin suponer un obstáculo. Esto, sin embargo, es más fácil de decir que de hacer. Esperamos que cuando tocamos una pantalla de un teléfono se realice una transferencia bancaria en el otro lado del mundo en milisegundos. Este simple acto requiere una inmensa cantidad de tecnología que necesitó una enorme cantidad de ingenio para ser creada y se basa en rigurosos fundamentos matemáticos y lógicos. Todo esto, y más, se enmarca dentro de la informática y las tecnologías de la información, disciplinas tan amplias que describirlas de manera breve es una tarea casi imposible.

En este momento debería ser intuitivamente claro que garantizar que un sistema de esta complejidad funcione a la perfección, siendo además confiable y escalable y manteniendo dentro de límites razonables los costes económicos de desarrollarlo y mantenerlo, es un enorme desafío. Existe una gran variedad de funcionalidad que queremos asegurar en un sistema informático. ¿Están mis datos seguros o pueden ser robados? ¿Son mis comunicaciones realmente privadas, o es posible que un tercero malintencionado tenga conocimiento de mis conversaciones? ¿Cuánta energía necesita un programa dado para funcionar, y, por lo tanto, cuánto tiempo puede funcionar con una cantidad de energía dada? ¿Qué garantías tenemos de que los programas en un automóvil no se detendrán, o comenzarán a funcionar erráticamente debido a un ataque informático?

Estas son algunas de las preguntas que la investigación del Instituto IMDEA Software intenta responder. El Instituto IMDEA Software es una institución de investigación avanzada creada bajo los auspicios de la Comunidad de Madrid, como parte de una serie de siete Institutos de investigación de primer nivel mundial en varias áreas de alta relevancia práctica para la sociedad. Estos avances se ponen a disposición tanto de la comunidad científica a través de los canales académicos habituales como del ecosistema industrial a través de contratos de investigación y proyectos conjuntos, y finalmente llegan a la

technology: be at our service without being a hurdle. This, however, is easier said than done. We expect that when we touch a smartphone screen, a bank transfer is made across the globe within milliseconds. This simple act requires an immense amount of technology whose creation required huge doses of ingenuity, and is based on rigorous mathematical and logical foundations. All of this, and more, falls within what we call computer science and information technologies – two disciplines so wide that describing them in a short and meaningful way is a daunting task.

At this point it should be intuitively clear that ensuring that a system of this complexity works seamlessly while being reliable and scalable, and keeping the economic costs of developing and maintaining it within reasonable bounds, is extremely challenging. There is a wide variety of functionality that we want such a computer system to have. Is my data safe, or can it be stolen? Are my communications really private, or is it possible for a malicious third party to be aware of my conversations? How much energy does a given program need to run, and therefore for how long can it run with a given energy storage? Can we have reasonable guarantees that the programs in a car will not halt to a freeze or start malfunctioning due to a hacker attack?

These are some of the questions that the research at the IMDEA Software Institute tries to answer. The IMDEA Software Institute is an advanced research institution created under the umbrella of the Regional Government of Madrid, as part of a series of seven Institutes performing world-class research in several areas of high practical relevance for the society. These advances are made available both to the scientific community through the usual academic channels and to the industrial ecosystem through research contracts and joint projects, and they eventually trickle down to the

tutos que realizan investigación con nivel de calidad internacional en varias áreas de alta relevancia práctica para la sociedad. Los avances conseguidos llegan tanto a la comunidad científica a través de los canales académicos habituales como al ecosistema industrial a través de contratos de investigación y proyectos conjuntos, y a la sociedad como mejoras en los productos que utilizamos, quizás sin darnos cuenta, todos los días.

Entre los trabajos que realiza el Instituto, de los que ofrecemos una muestra en las próximas páginas, podemos citar los avances en seguridad informática, privacidad, criptografía, computación segura, análisis del consumo de energía, seguridad del software aeroespacial y de móviles, sistemas distribuidos, blockchain, lenguajes de programación avanzados, técnicas para probar la corrección de los programas, optimización de la ejecución de programas, y el uso del aprendizaje automático como soporte de estas áreas. Nuestro trabajo impacta en la sociedad no creando o mejorando directamente los programas, aplicaciones y dispositivos que usamos todos los días, sino avanzando en la ciencia y tecnología que hace posible que los ingenieros de software tengan mejores herramientas con las que generar, con más facilidad y rapidez, aplicaciones que sean más seguras y eficientes.

La relevancia y pertinencia de esta investigación está atestiguada por los reconocimientos académicos recibidos a varios niveles por y el interés industrial generado. En el momento de redactar este texto, cuatro investigadores del Instituto son receptores de cuatro ayudas Starting y Consolidator del Consejo Europeo de Investigación (ERC por sus siglas en inglés). Las ayudas ERC son las más prestigiosas concedidas en Europa por un organismo público. En términos relativos, el Instituto IMDEA Software posee el 27% de todas las ayudas ERC de las categorías *Starting* y *Consolidator* que están activas en este momento en España en el área de Tecnologías de la Información. Para una institución del tamaño y antigüedad de IMDEA Software, este es un logro verdaderamente notable.

Desde su creación, el Instituto ha recibido cerca de 150 subvenciones, becas y contratos, que han atraído al Instituto fondos por valor de cerca de 30M€ (aproximadamente, una media de un 30% de su presupuesto durante este periodo, con un nivel de cofinanciación que empezó siendo más bajo en los inicios y que ha llegado a cerca del 50% en los últimos años). Dichos fondos provienen de compañías privadas y de agencias regionales, nacionales, europeas e internacionales. Entre las primeras, el Instituto ha tenido contratos con empresas como Google, Amazon, Facebook, GMV, INDRA, BBVA, Microsoft Research e Intel, y también con otras menos conocidas que trabajan en productos que sostienen la infraestructura del software, como Nomadic Labs, Nextel, LogicBlox, RedBorder y Zemsania. Además de eso, el Instituto ha colaborado en proyectos conjuntos con un número muy elevado de empresas que incluyen a la mayoría, si no todos, de los "sospechosos habituales" en el mundo TIC.

society as improvements in the products that we use, perhaps without noticing, every day.

Among the work that the Institute does, of which we will offer a sample in the next pages, we can cite research and advances in computer security, privacy, cryptography, secure computation, analysis of energy consumption, safety of mobile and aerospace software, distributed systems, blockchain, advanced programming languages, techniques to prove programs correctness, execution optimization, and the use of machine learning as a support for these areas. The way in which our work impacts society is not by directly creating or improving the programs, apps, and devices that we use every day, but by discovering the science and technology that makes it possible to develop better tools with which software engineers can generate in an easier and faster way applications that are safer and more efficient.

The relevance and timeliness of this research is witnessed by academic recognitions at several levels and the industrial interest it has generated. At the moment of writing, four researchers of the Institute are recipients of four Starting and Consolidator ERC grants, the most prestigious scientific grant awarded in Europe by a public body. In relative terms, the IMDEA Software Institute holds 27% of all the ERC Starting + Consolidator grants that are, at this moment, active in Spain in the area of Information Technologies. For an institution the size and age of IMDEA Software, this is a truly remarkable achievement.

The Institute has been awarded close to 150 grants, fellowships, and contracts since its inception, which brought to the Institute close to 30M€ in funds (roughly, an average of 30% of its budget during this period, being smaller at the beginning and reaching close to a 50% in the last years) from Regional, National, European, and International agencies, and from private companies. Among the latter, the Institute has had contracts with well-known companies such as Google, Amazon, Facebook, GMV, INDRA, BBVA, Microsoft Resarch, and Intel, and also with other lesser-known companies which work on products that sustain the software infrastructure, such as Nomadic Labs, Nextel, LogicBlox, RedBorder, and Zemsania. In addition to that, the Institute has collaborated in joint projects with a very large number of companies which include most, if not all, the "usual suspects" in the IT world.

Undoubtedly, the most precious asset of the Institute is its staff, from administration and support to the researchers. In this last category, the Institute has managed to attract a selected team coming from some of the most renowned universities and research centers in the world – among them some Spaniards who, after spending some years abroad, found an attracting opportunity to return to the their country. Our researchers routinely publish

Sin duda, el bien más preciado del Instituto es su personal, desde el de administración y apoyo hasta los investigadores. En esta última categoría, el Instituto ha logrado atraer a un selecto equipo proveniente de algunas de las universidades y centros de investigación más renombrados del mundo – entre ellos, españoles que, tras pasar unos años en el extranjero, encontraron una oportunidad atractiva para regresar a sus país. Nuestros investigadores publican artículos habitualmente en los mejores congresos y revistas y reciben galardones por su trabajo, tales como premios a los mejores artículos o mejores tesis. Además de eso, contribuyen a fortalecer el trabajo de formación de las universidades locales dirigiendo tesis de estudiantes a través de acuerdos con dichas universidades. En el momento de redactar este informe, los investigadores de IMDEA Software han dirigido 48 tesis doctorales, 50 tesis de máster y 23 tesis de grado.

Aparte de las actividades de investigación, el Instituto IMDEA Software también se ocupa de mantener REDIMadrid, la red regional de telecomunicaciones para investigación y educación superior, que da servicio 24/7 a más de 330.000 usuarios.

Actualmente trabajan en el Instituto más de 90 personas, 75 de las cuales están dedicadas a la investigación. Quienes tienen un doctorado lo obtuvieron de instituciones de primer nivel como Stanford, Carnegie Mellon, la Universidad de Cambridge, ETH Zurich, la Universidad de Purdue, la Universidad de California en San Diego, la Università de la Svizzera Italiana… El Instituto es un verdadero lugar internacional donde ahora mismo el 65% de su personal investigador es originario del extranjero.

Todos ellos, independientemente de su cargo o misión, comparten un mismo rasgo: la pasión y entrega que demuestran cada día en sus tareas y en el objetivo de mejorar la sociedad a través de la ciencia y la tecnología.

Se puede encontrar más información sobre el Instituto en su página web y, en particular, en los informes anuales disponibles allí. Estamos, como siempre, profundamente agradecidos al gobierno de la Comunidad de Madrid por su continuo apoyo.

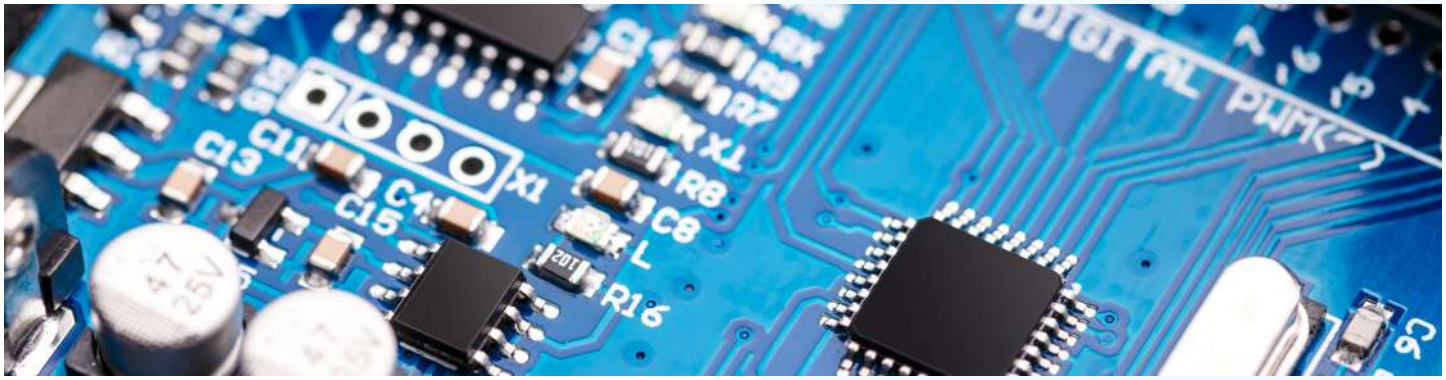
in the highest-ranked venues and receive prizes for their work, such as best paper and best thesis awards. In addition to that, they contribute to strengthening the education work of local universities by advising students through agreements with these universities. At the moment of writing, IMDEA Software researchers have been advisors of 48 PhD, 50 MSc, and 23 BSc theses.

Besides the research activities, the IMDEA Software Institute also takes care of managing REDIMadrid, the regional telecommunications network for research and higher-education, giving 24/7 service to more than 330.000 users.

More than 90 people work currently at the Institute, 75 of which are focused on research. Those with a PhD obtained it from top-level institutions such as Stanford, Carnegie Mellon, the University of Cambridge, ETH Zurich, Purdue University, the University of California at San Diego, the Università of the Svizzera Italiana… The Institute is a truly international place where right now a 65% of its research personnel is originally from abroad.

All of them, regardless of their position or mission, share a common trait: the passion and dedication that they show every day in their tasks and in the overarching objective of improving the society through science and technology.

More information on the Institute can be found at its web page and, in particular, in the annual reports available there. We are, as always, deeply grateful to the Regional Government of Madrid for its continuous support.



De la comunicación segura a la computación segura: nuevos retos en criptografía

From Secure Communication to Secure Computation: New Challenges in Cryptography

El objetivo de nuestra investigación es que podamos seguir utilizando toda la capacidad de las tecnologías de la información para nuestro beneficio sin renunciar a expectativas fundamentales de nuestra sociedad como son el derecho a la privacidad y la seguridad

Autores.- Ignacio Cascudo, Assistant Research Professor; Dario Fiore, Associate Research Professor; Pedro Moreno-Sánchez, Assistant Research Professor

La criptografía es la parte de la ciencia centrada en desarrollar técnicas para proteger la información y la comunicación, y juega un papel central en la seguridad de Internet. Gracias a la criptografía nos podemos comunicar a diario de forma segura con nuestras amistades o realizar compras online usando nuestra tarjeta de crédito sin miedo a que alguien pueda leer nuestros datos. La criptografía es un pilar fundamental de la *comunicación segura* que, en términos sencillos, permite el intercambio de información entre dos personas de forma que una tercera parte ajena a la conversación no pueda entender o modificar esta comunicación. La comunicación segura es crucial para la interacción digital.

Al mismo tiempo, la creciente digitalización de nuestra sociedad está creando nuevas maneras de interactuar y compartir información. Esto conlleva nuevos problemas de seguridad que no se pueden solventar con las herramientas criptográficas más tradicionales.

Por ejemplo, hoy en día es habitual almacenar y procesar nuestros datos (emails, documentos, fotos) en proveedores de servicios en Internet que nos permiten acceder a esta información cuando y desde donde queramos. Sin embargo, los beneficios de utilizar estos servicios pueden conllevar tener que revelar nuestros datos a estos proveedores. Y es que, aunque podríamos guardar estos datos de forma cifrada, esto nos haría perder funcionalidades. Por ejemplo, ya

The overall goal of our research is that the power of information technologies can continue to be used for our benefit without having to give up fundamental expectations in our society such as the right to privacy and security

Authors.- Ignacio Cascudo, Assistant Research Professor; Dario Fiore, Associate Research Professor; Pedro Moreno-Sánchez, Assistant Research Professor

Cryptography is the science of developing methods for protecting information and communication, and plays a central role in the security of the Internet. Thanks to cryptography, every day we can communicate securely with our friends, and send our credit card information when we shop online without the risk that anyone can read it. Cryptography is today's pillar of *secure communication* which, in simple terms, is the problem of enabling two parties to exchange information in such a way that a third party does not learn or tamper with this information. Secure communication is fundamental for digital interactions.

At the same time, the increasing digitalization of our society is leading to the emergence of new and more sophisticated ways in which our computing devices interact and share information. These new interaction paradigms also raise new security concerns that cannot be addressed using traditional cryptographic tools for secure communication.

For example, today it is common to store and process our data (emails, documents, pictures) on third-party providers that let us access this information whenever we want and from anywhere we are. However, the benefit of this paradigm may come at the cost of revealing our data to these providers. Indeed, although we could store data in encrypted form, this would also make us lose

no podríamos efectuar búsquedas en nuestros emails almacenados remotamente.

Otro ejemplo lo tenemos en la aplicación de novedosas técnicas de inteligencia artificial que, gracias al acopio y procesado de ingentes cantidades de datos provenientes de múltiples fuentes (sensores, clientes, etc.), nos ofrecen productos que pueden mejorar nuestras vidas. Por ejemplo, pensemos en cuántas veces usamos asistentes virtuales para obtener sugerencias acerca de la mejor ruta a casa, basándose en el tráfico actual, nuestras preferencias o nuestros hábitos personales. Sin embargo, tanto entrenar esta inteligencia artificial usando grandes cantidades de datos, como utilizarla para nuestras búsquedas concretas, proporcionan una gran cantidad de información privada de los usuarios al proveedor de este servicio.

Por último, las blockchain han aparecido en esta última década como una forma de ofrecernos servicios financieros de forma digital. Comenzando con la irrupción de Bitcoin, la tecnología de blockchain ha evolucionado como soporte para el equivalente digital del dinero en efectivo, en forma de monedas digitales emitidas por bancos centrales (CBDC), así como para permitir transacciones complejas gobernadas por determinadas reglas lógicas (los contratos inteligentes o smart contracts). Por ejemplo, utilizando tecnologías de blockchain, podemos programar un pago de forma que la cantidad y el receptor final dependan de la decisión tomada por un juez en un caso judicial. Sin embargo, estos avances pueden tener como consecuencia la pérdida de nuestra privacidad, ya que por ejemplo es posible que sea necesario hacer públicas nuestras transacciones y balances.

En resumen, aunque los beneficios de la digitalización son numerosos y evidentes, la recolección y uso de nuestros datos digitales pueden exponernos a riesgos y afectar a nuestra seguridad y privacidad, ambos derechos fundamentales en muchos países. El problema principal es que, en estas nuevas aplicaciones, nuestros datos ya

some important functionalities. For example, we would no longer be able to search over our remotely stored emails.

Also, new artificial intelligence techniques, thanks to massive collection of data from multiple parties (sensors, customers, etc.) and its processing, can provide us with new products that improve our life. As an example, let us think how many times we use a computer assistant to get suggestions on the best way to reach our home based on current traffic and our personal preferences or habits. However, both training this artificial intelligence on a data bank, and then using it on our specific query gives, in principle, a massive amount of sensitive information to the provider of this service.

Blockchain has also emerged during the last decade as the means to bring financial services to everyone in the digital form. Starting from the inception of the Bitcoin cryptocurrency, the blockchain technology has evolved to support the digital counterpart of cash in the form of central bank digital currency (CBDC), or allow complex transfers governed by a certain spending logic (or smart contract). For instance, using blockchain, we could program a payment such that the receiver and the amount are established according to the final decision of a judge on the corresponding court case. However, the technical advance of blockchain technology may come at the cost of privacy guarantees, e.g., making the complete set of balances and executed transactions publicly available.

In summary, while the benefits of digitalization are numerous and clear, the pervasive collection and use of digital data may also expose us to more risks, undermining our security and privacy, which are fundamental rights enshrined in many national regulations. The main issue is that, under these new paradigms, our data gets out of our control and is given to third parties. But, what is the guarantee that these parties behave honestly?

no están bajo nuestro control, sino que pasan a ser controlados por terceras partes. Pero ¿qué garantías tenemos de que estas terceras partes se comporten honestamente?

Desafortunadamente, las herramientas de la criptografía tradicional no bastan para ofrecer seguridad en estas situaciones. El problema es que ya no solo estamos enviando y recibiendo información, sino que esta información necesita ser procesada. Además, en el problema tradicional de la comunicación segura nos teníamos que proteger de un adversario externo, ajeno a la comunicación. Sin embargo, en este nuevo paradigma de aplicaciones con acceso a nuestros datos y comunicaciones, el adversario puede ser una de las entidades que toman parte en la computación, por ejemplo el proveedor de servidores para cierta computación o el proveedor de almacenamiento en la nube. Esto no solo provoca que tengamos que resolver nuevos problemas de privacidad (esto es, permitir que terceras partes puedan procesar nuestros datos sin darles acceso completo a ellos) sino que a menudo también necesitamos asegurarnos de que estas terceras partes estén procesando nuestros datos de forma honesta, acorde a la funcionalidad que nosotros esperamos de la aplicación.

Dicho de otro modo, ahora necesitamos formas de asegurar la *computación* segura y no sólo la *comunicación* segura.

En el instituto IMDEA Software trabajamos en el diseño de nuevos métodos de computación segura. Es muy probable que estos métodos pasen a ser un pilar de la interacción digital en el futuro, permitiendo el intercambio y procesado de información de forma que se pueda hacer uso esta información, pero a la vez protegiendo su seguridad y privacidad. Más concretamente, trabajamos en dos clases principales de cuestiones: la computación privada y la computación verificable.

Computación privada

La computación privada ofrece herramientas para procesar datos sin necesidad de 'verlos'. Como ejemplo de uso, pensemos en el caso de varios hospitales (pertenecientes a distintas administraciones) que estén llevando a cabo un estudio científico conjunto, utilizando sus historiales médicos. Estos historiales son confidenciales y no pueden cruzar barreras administrativas; pero la combinación de los historiales en diversos hospitales daría información mucho más valiosa para el estudio que la que cada hospital puede obtener de forma independiente con los historiales de los que dispone. Las herramientas propias de la computación privada permiten a estos hospitales, o incluso a una tercera parte (por ejemplo otro organismo de investigación) realizar el estudio y obtener su resultado sin necesidad de llegar a conocer los historiales.

Unfortunately, traditional cryptography falls short of providing meaningful security in this paradigm. The problem is that information is not only sent and received, but also processed. Also, while in the traditional problem of secure communication the attacker is an external party not involved in the communication, in this new paradigm the adversary may be one of the parties taking part in the computation – for example the cloud storage or computation provider. This not only poses new challenges regarding privacy (how to allow third parties to process our confidential data without giving them full access to it), but also with respect to verifiability: often we also need to make sure that these parties process our data exactly in the way they are supposed to.

In other words, we need ways to ensure *secure computation*, rather than only communication.

At the IMDEA Software Institute, we are working on the design of novel cryptographic methods that enable secure computation. These methods are likely to become the pillar of the future's digital interactions by allowing individuals to exchange and process information in ways that simultaneously enable the usage of this information and the protection of their security and privacy.

Specifically, we are working on two main classes of cryptographic problems: privacy-preserving computation and verifiable computation.

Privacy preserving computation

Privacy-preserving computation provides tools for processing data without needing to ``see it''. As an example, let us consider the case of several hospitals (belonging to different administrations) jointly running a scientific study on their medical records. This data is confidential and cannot cross administrative boundaries; yet the combination of the different records would give much more valuable information for the study than what each hospital can obtain by itself. Privacy-preserving computation tools allow the hospitals, or even a third party (e.g. another research organization), to obtain the result of the study without the need of revealing the data.

With the advent of machine learning, which can extract very useful information from huge amounts of data, applications like the one above are becoming ubiquitous in many fields. Moreover, training of machine learning models is expensive, and the resulting models are often a valuable asset. Privacy-preserving computation not only allows one to keep the data confidential, but can also hide the trained machine learning model that is being used in the analysis of that data.

Con el desarrollo del aprendizaje automático y la inteligencia artificial, ambas técnicas caracterizadas por el hecho de que pueden extraer información muy útil a partir de ingentes cantidades de datos, están surgiendo múltiples casos de uso como el descrito anteriormente, en diversos campos. Además, el entrenamiento de modelos de aprendizaje automático es normalmente un proceso caro y, por tanto, los modelos resultantes son un bien preciado. La computación privada puede no solo tratar los datos de forma confidencial, sino que también puede mantener oculto el modelo de aprendizaje automático que se está utilizando para ese análisis de datos.

Entre las tecnologías de computación segura más importantes se encuentran el cifrado homomórfico, el cifrado funcional y la computación multi-parte segura. Los cifrados homomórfico y funcional son tipos especiales de cifrado que permiten operar sobre información privada sin necesidad de conocer el mensaje subyacente. Por ejemplo, dados dos números cifrados, podemos crear el cifrado de su suma sin necesidad de descifrarlos. El poseedor de la clave privada puede entonces descifrar el resultado final después de haber sido procesado por otras entidades.

La computación multiparte segura estudia cómo llevar a cabo computaciones privadas de forma colaborativa y descentralizada. En este caso, en lugar de cifrar la información confidencial, ésta se distribuye entre distintos participantes de la computación, de forma que la información recibida por cada uno de ellos, por sí misma, es independiente de los datos iniciales y no revela ninguna información sobre ellos. Solo cuando un conjunto suficientemente grande de los participantes se pone de acuerdo se pueden reconstruir estos datos. Los participantes de la computación pueden llevar a cabo operaciones sobre los datos distribuidos y reconstruir sólo el resultado final de la computación. Dada la naturaleza distribuida de la computación, ninguno de los participantes llega a conocer el conjunto inicial de datos.

Privacy-preserving computation tools include homomorphic and functional encryption, and secure multiparty-computation. Homomorphic and functional encryption are special kinds of encryption that can operate on encrypted data without knowing the underlying message. For example, a party can take two encrypted numbers and create an encryption of the sum of these numbers without decrypting them. The owner of the private key can decrypt and recover the result after the processing.

Secure multiparty computation studies how to perform privacy-preserving calculations in a collaborative, decentralized manner. In this case, the confidential data is not encrypted, but rather distributed among several parties, in such a way that the information received by each of them looks, by itself, completely unrelated to the original data. Only when a large enough number of parties agree can the data be reconstructed. Parties can compute on the distributed data and the computation result can be reconstructed at the end of the day. Because of the distributed nature of the computation, none of the parties get to know the whole original set of data.

Nevertheless, widespread adoption of homomorphic encryption and secure multiparty computation technologies presents some challenges. Perhaps the most important of them are that solutions incur in large computation times (especially in the case of homomorphic encryption) and a large amount of communication between parties (relevant in the case of secure multiparty computation), in comparison to the cost of computation if the privacy requirements were not present. Another challenging aspect in certain applications is to address the possibility that some parties may actively cheat and try to lead others to wrong outcomes.

No obstante, la adopción generalizada de las tecnologías de cifrado homomórfico y computación multiparte segura presenta algunos desafíos. Quizás el más importante es el hecho de que las soluciones actuales incurren en grandes tiempos de cómputo (especialmente en el caso del cifrado homomórfico) y en una gran cantidad de comunicación entre las partes (en el caso de la computación multiparte segura), en comparación con su coste si los requisitos de privacidad no estuvieran presentes. Otro reto relevante en ciertas aplicaciones es cómo evitar que algunos participantes de estos protocolos actúen de forma maliciosa con el objetivo de llevar a otros a obtener resultados equivocados.

En nuestro Instituto dedicamos parte de nuestros esfuerzos a minimizar los costes de cómputo y comunicación mencionados anteriormente, utilizando para ello una serie de técnicas matemáticas. Estas estrategias incluyen encontrar técnicas más eficientes para calcular de forma segura varias evaluaciones del mismo tipo en diferentes conjuntos de datos (lo que en informática suele llamarse SIMD - instrucción única, datos múltiples) o encontrar protocolos especializados para ciertos tipos de cálculos, como los utilizados en la evaluación de redes neuronales.

Computación verificable y pruebas de conocimiento cero

La computación verificable aborda la pregunta: "¿cómo podemos delegar una cierta computación a un proveedor de servicios en el que no confiamos y estar seguros de que los resultados que recibimos son correctos?"

Este problema no es trivial cuando consideramos una de las siguientes propiedades: eficiencia y conocimiento cero. En el caso de la eficiencia, queremos que la persona que delega el cálculo quede convencida de la exactitud del resultado, utilizando muchos menos recursos de los que dicha persona requeriría para llevar a cabo el cálculo por sí sola. De lo contrario, uno podría ejecutar el cálculo, y la ventaja de delegarlo se desvanece.

Las pruebas de conocimiento cero consideran el caso en el que la computación involucra información que no se debe revelar a la parte que realiza el cálculo. Por ejemplo, la ejecución de una transacción financiera puede involucrar información confidencial (como el saldo de la cuenta, o el precio de compra/venta). Sin embargo, el objetivo es que todo el mundo se convenza de que los resultados de las transacciones se calculen correctamente, sin hacer trampa. Una computación verificable de conocimiento cero permite convencer a cualquiera de que el cálculo es correcto sin revelar más información que la estrictamente necesaria sobre los datos privados involucrados en él.

La computación verificable de conocimiento cero genera una prueba de la validez del resultado que se puede verificar de manera eficiente.

At our Institute we are devoting research efforts to minimizing these overheads using a number of mathematical techniques. These strategies include finding more efficient techniques to compute securely several evaluations of the same type on different sets of data (which in computer science is usually called SIMD - single instruction, multiple data) or finding specialized protocols for certain types of computations, such as the ones used in evaluation of neural networks.

Verifiable computation and zero-knowledge proofs

Verifiable computation tackles the question: ``how can we delegate a computation to a third party that we do not trust and be sure that the results we receive are correct?''

This problem is nontrivial when we consider either one of the following properties: efficiency and zero-knowledge. For efficiency, the party that delegates the calculation should be convinced of the result's correctness using much less resources than what doing the computation on its own requires. Otherwise, one could execute the calculation, and the advantage of delegating it vanishes.

Zero-knowledge tackles the case when the party performing the computation introduces its own secret information. For instance, executing a financial transaction may involve secret information (e.g., account balance, buy/sell price). Yet, everyone wants to make sure that the results of the transactions are correctly computed, without cheating. A verifiable computation that is zero-knowledge allows convincing anyone that the computation is correct without revealing any more information on these private inputs.

Zero-knowledge verifiable computation generates a proof of the validity of the result that can be verified efficiently. Zero-knowledge proofs are an active and important research area that advanced significantly in the last ten years. In spite of this, the widespread adoption of this technology still presents significant challenges related to their efficiency – notably the size of proofs and the cost to generate and verify them.

At our Institute, we are actively working on improving this efficiency from different angles. We explore new design approaches to construct proof systems from simpler building blocks and we investigate the development of mathematical techniques to speed up the cost of proof generation for specific classes of computations of wide interest. As a more concrete example, our research includes the development of efficient proof systems for standard digital signatures in blockchain. The goal is to ensure that a transaction is not only authorized by the right person (e.g., the sender) but only when the right conditions are met. For instance, our technique allows that the authorization of a financial

Las pruebas de conocimiento cero son un área de investigación activa e importante que ha avanzado significativamente en los últimos diez años. A pesar de esto, la adopción generalizada de esta tecnología aún presenta desafíos importantes relacionados con su eficiencia, en particular el tamaño de las pruebas y el coste de generarlas y verificarlas.

En nuestro Instituto, trabajamos para mejorar esta eficiencia desde diferentes ángulos. Exploramos nuevos enfoques para diseñar sistemas de prueba a partir de técnicas y herramientas más simples e investigamos el desarrollo de técnicas matemáticas para acelerar la generación de pruebas para casos específicos de computaciones que sean de amplio interés. Como ejemplo más concreto, nuestra investigación incluye el desarrollo de sistemas de prueba eficientes para firmas digitales en blockchain. El objetivo es garantizar que una transacción no sólo sea autorizada por la persona adecuada (por ejemplo, el remitente), sino que además solo sea aceptada cuando se cumplan las condiciones adecuadas. Por ejemplo, nuestra técnica permite que la autorización de una transacción financiera demuestre de forma adicional la veracidad de otra información (por ejemplo, que el remitente tiene fondos suficientes para realizar la transacción). Lo interesante es que dicha prueba está diseñada de tal manera que no revela información confidencial (el saldo en sí) y requiere una cantidad de computación y comunicación mínima.

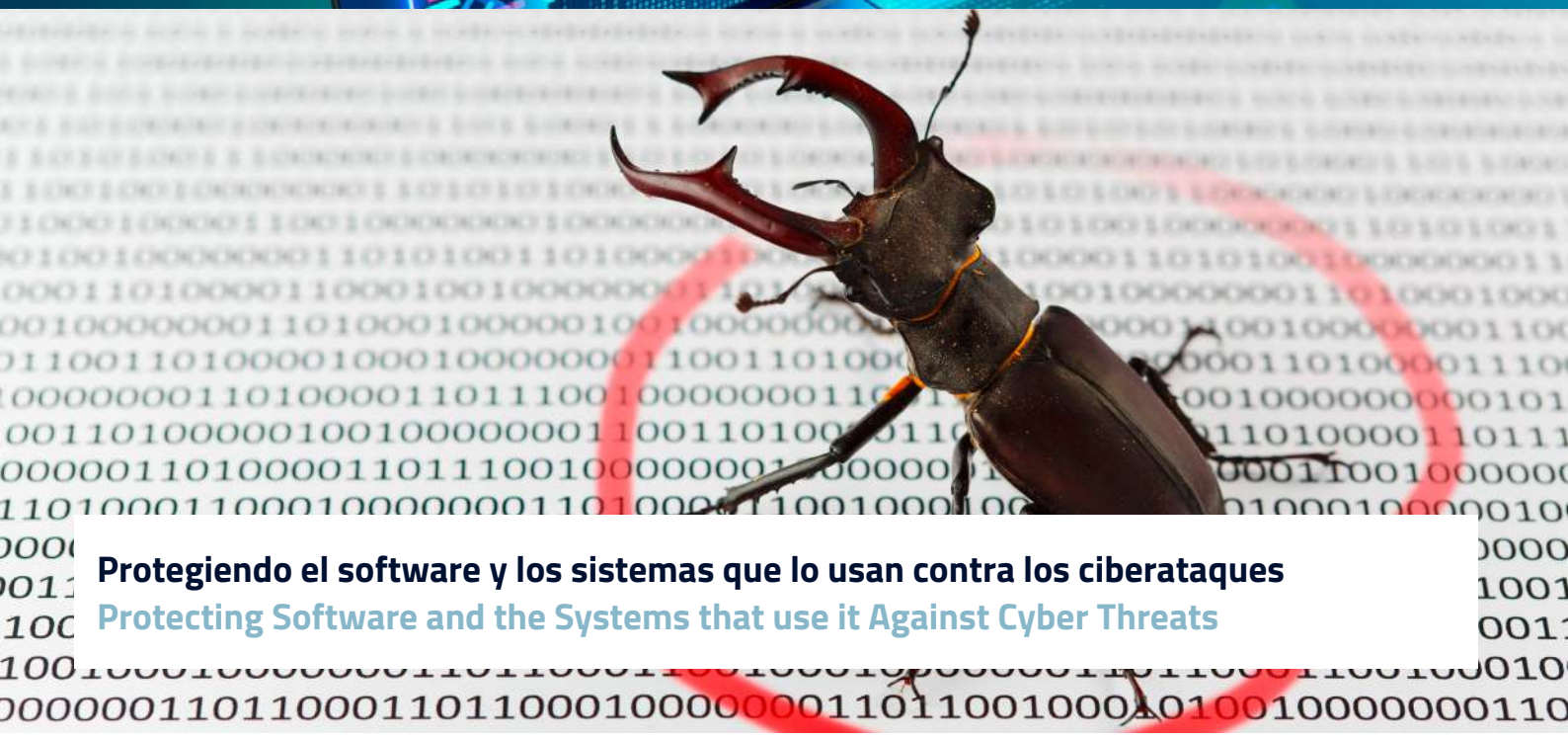
En resumen, el objetivo general de nuestra investigación es que podamos seguir utilizando toda la capacidad de las tecnologías de la información para nuestro beneficio sin tener que renunciar a expectativas fundamentales de nuestra sociedad como son el derecho a la privacidad y la seguridad.

La investigación que hemos descrito ha recibido recientemente financiación del Consejo Europeo de Investigación en el marco del proyecto PICOCRYPT del ERC, además de fondos procedentes de otras fuentes públicas y privadas, como los gobiernos nacional y regional, GMV, BBVA, la Fundación Tezos, Chaincode Labs, Protocol Labs, NEC Labs y otros.

transaction simultaneously proves the veracity of other information of interest (e.g., the sender has enough funds to perform the transaction in the first place). Interestingly, such functionality is designed in such a manner that it does not reveal sensitive information (the balance itself), and minimizes the computation and communication overhead.

The overall goal of our research is that the power of information technologies can continue to be used for our benefit without having to give up fundamental expectations in our society such as the right to privacy and security.

These research efforts described here have recently received funding from the European Research Council under the ERC project PICOCRYPT and funds coming from other private and public sources, such as the National and Regional governments, GMV, BBVA, the Tezos Foundation, Chaincode Labs, Protocol Labs, NEC Labs, and others.



Protegiendo el software y los sistemas que lo usan contra los ciberataques

Protecting Software and the Systems that use it Against Cyber Threats

El grupo de Ciberseguridad del Instituto IMDEA Software desarrolla múltiples líneas de investigación que tienen como objetivo prevenir las vulnerabilidades del software y defenderse frente a ellas

Autores.- Juan Caballero (Associate Research Professor); Alessandra Gorla (Assistant Research Professor); Marco Guarnieri (Assistant Research Professor)

Casi todos los días aparecen nuevas amenazas contra el desarrollo de las sociedades digitales: ataques que inhabilitan en plena pandemia los sistemas de la administración pública que hacen posible el pago de prestaciones por desempleo, empresas que se quedan sin acceso a sus sistemas digitales durante semanas o meses, campañas de desinformación para influir en los resultados de elecciones democráticas, ataques contra infraestructuras críticas como las instalaciones de gas y nucleares, filtraciones masivas de credenciales de usuarios o el espionaje de los dispositivos móviles de activistas, políticos e incluso jefes de estado.

La ciberseguridad se ha convertido en algo crítico que debe abordarse no solo desde el punto de vista económico, sino también para garantizar la libertad y la privacidad a las que estamos acostumbrados en las sociedades democráticas. Con este objetivo, el grupo de Ciberseguridad del Instituto IMDEA Software se estableció hace más de una década para desarrollar la ciencia que permite proteger el software y los sistemas informáticos que se basan en él. La investigación del grupo abarca, entre otros, la protección del software en las diversas etapas de su desarrollo, el análisis de programas involucrados en violaciones de privacidad como las aplicaciones móviles, y la detección y atribución de ciberataques. El grupo tiene una establecida reputación internacional. Todos sus investigadores de plantilla tienen doctorados de las principales universidades europeas y norteamericanas,

The Cybersecurity group at the IMDEA Software Institute has multiple research lines that aim to prevent software vulnerabilities and provide protections against them

Authors.- Juan Caballero (Associate Research Professor); Alessandra Gorla (Assistant Research Professor); Marco Guarnieri (Assistant Research Professor)

Almost every day we are reminded of the ever-growing threats against the development of digital societies. The range of such threats is wide: attacks disabling the public administration systems that make possible paying unemployment benefits in the middle of a pandemic, companies left without access to their digital systems for weeks or months, disinformation campaigns to influence the results of democratic elections, hacking of critical infrastructures such as gas and nuclear facilities, massive privacy leaks of user credentials, and spying of the mobile devices of activists, politicians, and even heads of state.

Cybersecurity has thus become a critical issue that needs to be addressed not only from the economic point of view, but also to guarantee the freedom and privacy that we are used to in democratic societies. Towards this goal, the Cybersecurity group at the IMDEA Software Institute was established over a decade ago for developing the science that enables securing software and the digital systems that are built upon it. The group's research encompasses, among others, the protection of the software at the heart of digital systems, the analysis of software involved in privacy violations such as mobile applications, and the detection and attribution of cyberattacks. The group has a well-established international reputation. All its faculty have PhDs from leading European and North American universities such as Carnegie Mellon

como la Universidad Carnegie Mellon (EE.UU.), ETH Zurich (Suiza) y la Universidad de Lugano (Suiza). El grupo colabora activamente con los mejores investigadores internacionales y nacionales, así como con otros grupos del Instituto IMDEA Software.

El proceso de desarrollo a menudo introduce defectos en el software resultante. A veces, estos defectos sólo irritan a los usuarios porque les impiden usar el software según lo previsto. Sin embargo, algunos defectos pueden ser explotados por un atacante para comprometer el sistema que ejecuta el software vulnerable. La explotación de vulnerabilidades es un importante vector de infección que conduce a la instalación de software malicioso (es decir, malware). Las vulnerabilidades más peligrosas son las de día cero, que una vez encontradas no se notifican a los desarrolladores del software vulnerable para que de esta manera puedan utilizarse en ciberataques. Existen mercados donde se comercia con vulnerabilidades de día cero y donde éstas se pueden vender por decenas de miles, cientos de miles o incluso un millón de dólares, según la popularidad del software vulnerable. Entre las diferentes clases de vulnerabilidades, algunas denominadas cero-clic son particularmente peligrosas porque no requieren de la participación del usuario durante la explotación. Por lo tanto, pueden usarse para infectar discretamente un dispositivo. Por ejemplo, vulnerabilidades de día cero y cero-clic han sido utilizadas en ataques recientes que infectaron los dispositivos móviles de activistas y políticos. El grupo de Ciberseguridad del Instituto IMDEA Software desarrolla múltiples líneas de investigación que tienen como objetivo prevenir las vulnerabilidades del software y defenderse frente a ellas.

Una línea de investigación aborda la protección contra los canales laterales de microarquitectura. Los canales laterales suceden cuando la ejecución de alguna acción por parte de un programa depende de un

University (USA), ETH Zurich (Switzerland), and University of Lugano (Switzerland). It actively collaborates with top international and national researchers, as well as other groups at the IMDEA Software Institute.

Software defects are often introduced during software development. Sometimes those defects only irritate users, preventing them from using the software as intended. Some defects, however, can be exploited by a determined attacker to compromise the system running the vulnerable software. Exploiting vulnerabilities is a major infection vector leading to the installation of malicious software (i.e., malware). Most dangerous are zero-day vulnerabilities, which once found are not reported to the developers of the vulnerable software so that they can be used in cyberattacks. There exist markets for zero-day vulnerabilities, where they can be sold for tens of thousands, hundreds of thousands, or even a million dollars, depending on the popularity of the vulnerable software. Among the different classes of vulnerabilities some called zero-click are particularly dangerous because they do not require any user involvement during exploitation. Thus, they can be used to discreetly compromise digital systems. For example, zero-day, zero-click vulnerabilities have been involved in recent attacks that infected the mobile devices of activists and politicians. The Cybersecurity group at the IMDEA Software Institute has multiple research lines that aim to prevent software vulnerabilities and provide protections against them.

One research line addresses the protection against micro-architectural side-channels. When the execution of some action by a program depends on a secret (e.g., a password), attackers can often recover this secret---using a so-called side-channel---by

secreto (por ejemplo, una contraseña). En estos casos los atacantes a menudo pueden recuperar este secreto midiendo las características físicas de la ejecución del programa, como el tiempo de ejecución o el consumo de energía, lo que se denominan canales laterales. Los canales laterales de microarquitectura son un tipo específico de canales laterales que surgen de la interacción entre el hardware y el software. Han dado lugar a ataques notables como Spectre y Meltdown que han afectado a casi todos los procesadores modernos, es decir, miles de millones de sistemas informáticos. Como parte de un proyecto de investigación financiado por Intel, un fabricante líder de microchips, nuestros investigadores han desarrollado técnicas novedosas para identificar canales laterales de microarquitectura y han propuesto defensas contra ellos, como las basadas en la compilación segura [1].

Otra línea de investigación, en colaboración con investigadores de otros grupos de IMDEA Software, trata de erradicar las vulnerabilidades durante el desarrollo de software mediante el uso de lenguajes de programación seguros y técnicas de verificación de programas. Además, también desarrollamos líneas de investigación cuyo objetivo es identificar y eliminar vulnerabilidades después de que se haya desarrollado el software, pero antes de que se distribuya a los usuarios finales. Una de estas líneas de investigación se ocupa del testeo automático del software. El proceso de testeo valida la funcionalidad de un programa, incluida su seguridad y escalabilidad. El proceso de testeo es a menudo el coste más grande durante el desarrollo de software. Como tal, la automatización del proceso de testeo puede conducir no solo a una mayor seguridad, sino también a ahorros significativos en el desarrollo del software. El testeo automático es fundamental para múltiples plataformas. Por ejemplo, en el proyecto Horizon 2020 ElasTest, financiado por la Unión Europea, nuestros investigadores han desarrollado nuevas técnicas de testeo automático para servicios web de gran escala. En otro proyecto llamado Madrid Flight on Chip, financiado por la Comunidad de Madrid, nuestros investigadores colaboran con empresas aeronáuticas nacionales para desarrollar técnicas de última generación para testear sistemas de vuelo y satélites basados en tecnologías multiprocesador system-on-a-chip.

Otro aspecto importante de la ciberseguridad es el análisis de amenazas como malware y software potencialmente no deseado, que se utilizan para establecer una presencia en sistemas infectados como estaciones de trabajo, ordenadores portátiles, teléfonos móviles y dispositivos IoT. El malware es software con un objetivo claramente malicioso, como el ransomware, que puede cifrar los datos en un dispositivo y exigir un pago para recuperarlo, o los gusanos que se propagan automáticamente de una máquina a otra mediante la explotación de vulnerabilidades.

Por el contrario, el software potencialmente no deseado es software que, si bien no es claramente malicioso, puede afectar a la privacidad y que, en general, la mayoría de los usuarios querrían ser informados

measuring physical characteristics of the program's execution like execution time or energy consumption. Micro-architectural side-channels are a specific type of side-channels that arise from the interaction between the hardware and the software. They have led to noteworthy attacks such as Spectre and Meltdown that affect almost all modern processors, that is, billions of IT systems. As part of a research project funded by Intel, a leading chip manufacturer, our researchers have developed novel techniques to identify micro-architectural side-channels and have proposed defenses against them such as those based on secure compilation [1].

Another research line, in collaboration with researchers from other groups at IMDEA Software, deals with eradicating vulnerabilities by design during software development through the use of secure programming languages and program verification techniques. In addition, our research also aims to identify and remove vulnerabilities after the software has been developed, but before it is deployed to end users. One such research line deals with automatic testing of software. Testing validates the functionality of a program including its security and scalability. Testing is often the largest cost in the whole process of software development. As such, automating the testing process can lead not only to improved security, but also to significant savings throughout software development. Automated testing is fundamental for multiple platforms. For example, in the Horizon 2020 *ElasTest* project, funded by the European Union, our researchers have developed novel automated testing techniques for large scale Web services. In another project called *Madrid Flight on Chip*, funded by the Regional Government of Madrid, our researchers collaborate with national aeronautics companies to develop next generation techniques for testing flight and satellite systems based on multiprocessor system-on-a-chip technologies.

Another important aspect of cybersecurity is the analysis of threats such as malware and potentially unwanted software, which are used to establish a presence in infected systems such as workstations, laptops, mobile phones, and IoT devices. Malware is clearly malicious software such as ransomware that may encrypt the data in a device and demand a payment to recover it, or worms that automatically spread from machine to machine by exploiting vulnerabilities.

In contrast, potentially unwanted software is software that, while not clearly malicious, may still be privacy-invasive, and that in general most users may want to be aware if they inadvertently install them on their devices. If malware is considered black, and benign software white, then potentially unwanted software is the gray in between. One example of potentially unwanted software are invasive advertisement libraries and mobile apps that include them. A user may install an app and give it access to sensitive

si lo instalaran sin darse cuenta en sus dispositivos. Si el malware se considera negro y el software benigno blanco, entonces el software potencialmente no deseado es el gris en el medio. Un ejemplo de software potencialmente no deseado son las librerías de anuncios invasivas y las aplicaciones móviles que las incluyen. Un usuario puede instalar una aplicación y darle acceso a datos privados (por ejemplo, geolocalización o contactos) porque la aplicación proporciona alguna funcionalidad que el usuario desea, pero es posible que el usuario no sepa que la aplicación incluye una librería de anuncios de un tercero que también obtiene acceso a los datos privados. Otro caso son las aplicaciones de doble uso, que se desarrollan con una intención posiblemente benigna, pero pueden ser utilizadas en escenarios maliciosos. Por ejemplo, las herramientas de acceso remoto se pueden usar para ayudar a las personas mayores a configurar sus dispositivos a distancia. Pero, si se instalan sin el consentimiento del propietario, como por ejemplo en casos de violencia de pareja, claramente se convierten en malware.

El grupo de Ciberseguridad del Instituto IMDEA Software desarrolla técnicas novedosas para analizar malware y otro software no deseado, y las aplica para comprender el ecosistema de ciberamenazas. Por ejemplo, en colaboración con investigadores de Norton LifeLock, una de las mayores empresas de seguridad (anteriormente conocida como Symantec), nuestros investigadores han analizado cómo se distribuye el malware y las aplicaciones no deseadas a los usuarios de Android y han descubierto que, a pesar de la gran cantidad de recursos que Google invierte para defender su mercado de aplicaciones Play, sigue siendo la mayor fuente de instalaciones de aplicaciones no deseadas, ya que los desarrolladores de software malicioso tienen un incentivo para vencer esas defensas para que sus aplicaciones maliciosas puedan llegar a una audiencia amplia [2]. En otro

data (e.g., geolocation, contacts) because the app provides some functionality that the user desires, but the user may not be aware that the app includes an advertisement library from a third party that also gets access to the sensitive data. Another case are dual-use apps, which are developed with a possibly benign intent, but can be abused in more nefarious scenarios. For example, remote access tools can be used to assist seniors in configuring their devices. But, if installed without the owner's consent such as in intimate partner violence cases they clearly become malware.

The Cybersecurity group at the IMDEA Software Institute develops novel techniques to analyze malware and other unwanted software, and applies them to understand the cyber-threat ecosystem. For example, in collaboration with researchers from Norton LifeLock, one of the largest security companies (previously known as Symantec), our researchers have analyzed how malware and unwanted apps are distributed to Android users finding that despite the large amount of resources that Google invests to defend its Play store, it is still the largest source of unwanted app installations as developers of malicious software have an incentive to defeat those defenses so that their malicious apps can reach a wide audience [2]. In another example, our researchers have collaborated with researchers from the IMDEA Networks Institute to analyze the privacy-invasive behaviors of parental control apps [3], used by parents to monitor and limit their children's online behaviors, finding that over one third of those apps send personal information without appropriate user consent and 72% of the apps share data with third parties such as online advertising and analytics services, without mentioning their presence in privacy policies. Current research projects develop analysis techniques for other platforms such as the analysis of advertisement libraries in

ejemplo, nuestros investigadores han colaborado con investigadores del Instituto IMDEA Networks para analizar los comportamientos que afectan a la privacidad en las aplicaciones de control parental [3], utilizadas por los padres para monitorear y limitar el acceso a Internet de sus hijos. Los resultados de esta investigación muestran que más de un tercio de esas aplicaciones envían información personal sin el debido consentimiento del usuario y el 72% de las aplicaciones comparten datos con terceros, como agencias de publicidad, sin mencionar esto en sus políticas de privacidad. También tenemos en marcha proyectos de investigación que desarrollan técnicas de análisis para otras plataformas, como el análisis de librerías de publicidad en dispositivos iOS, técnicas de aprendizaje automático para la detección y clasificación de malware de Windows y la identificación de relaciones de cibercrimen en la cadena de bloques de Bitcoin.

[1] Marco Patrignani, Marco Guarnieri. Exorcising spectres with secure compilers. Proceedings of the 28th ACM Conference on Computer and Communications Security, November 2021. DOI: 10.1145/3460120.3484534

[2] Platon Kotzias, Juan Caballero, Leyla Bilge. How Did That Get In My Phone? Unwanted App Distribution on Android Devices. Proceedings of the 42nd IEEE Symposium on Security and Privacy. May 2021. DOI: 10.1109/SP40001.2021.00041

[3] Alvaro Feal, Paolo Calciati, Narseo Vallina-Rodriguez, Carmela Troncoso, Alessandra Gorla. Angel or Devil? A Privacy Study of Mobile Parental Control Apps. Proceedings of Privacy Enhancing Technologies (PoPETS) 2020. DOI: 10.2478/popets-2020-0029

iOS devices, machine learning techniques for the detection and classification of Windows malware, and the identification of cybercrime relationships in the Bitcoin blockchain.

[1] Marco Patrignani, Marco Guarnieri. Exorcising spectres with secure compilers. Proceedings of the 28th ACM Conference on Computer and Communications Security, November 2021. DOI: 10.1145/3460120.3484534

[2] Platon Kotzias, Juan Caballero, Leyla Bilge. How Did That Get In My Phone? Unwanted App Distribution on Android Devices. Proceedings of the 42nd IEEE Symposium on Security and Privacy. May 2021. DOI: 10.1109/SP40001.2021.00041

[3] Alvaro Feal, Paolo Calciati, Narseo Vallina-Rodriguez, Carmela Troncoso, Alessandra Gorla. Angel or Devil? A Privacy Study of Mobile Parental Control Apps. Proceedings of Privacy Enhancing Technologies (PoPETS) 2020. DOI: 10.2478/popets-2020-0029

Software "Ecológico": Verificando y Controlando el Consumo de Recursos

"Greener" Software: Verifying and Controlling Resource Consumption

Aunque ha habido mejoras en la tecnología de baterías y recolección de energía, se necesita una reducción significativa en las demandas de energía de dichos dispositivos para hacer realidad la visión completa de Internet de las Cosas y explotar todo su potencial

Autores.- Dr. Pedro López-García, Investigador Científico en el CSIC y Colaborador en IMDEA Software; Dr. Manuel Hermenegildo, Catedrático en la UPM y Profesor Distinguido en IMDEA Software; Dr. José F. Morales, Profesor Ayudante Doctor en la UPM e Investigador en IMDEA Software; Víctor Pérez, Ingeniero de Software en Codeplay Software

El consumo de energía y el impacto medioambiental de las tecnologías de la computación se han convertido en una gran preocupación a nivel mundial. Es ya un problema importante en sistemas que van desde *granjas de servidores* con un enorme consumo de energía, hasta miles de millones de dispositivos complejos alimentados con baterias que son frecuentemente cargadas, como *teléfonos y relojes inteligentes, tabletas, sensores y dispositivos médicos implantables y portables.* Como resultado del enorme crecimiento de la *computación en la nube*, el *tráfico en Internet*, la *computación de altas prestaciones* y las *aplicaciones distribuídas*, los centros de cálculo actuales consumen grandes cantidades de energía, no sólo para procesar y transportar datos, sino también para refrigeración.

El consumo de energía es también muy relevante en el contexto del paradigma de *Internet de las Cosas,* en donde miles de millones de pequeños dispositivos autónomos, empotrados en todo tipo de objetos, en nuestra ropa, o pegados a nuestros cuerpos, tendrán que operar e intercomunicarse contínuamente durante largos periodos de tiempo, incluso años.

Although there have been improvements in battery and energy harvesting technology, a significant reduction in the energy demands of such devices is needed to make the full Internet of Things vision come true and all its potential be exploite

Authors.- Dr. Pedro López-García, Tenured Researcher at CSIC and Affiliate Faculty at IMDEA Software; Dr. Manuel Hermenegildo, Full Professor at UPM and Distinguished Professor at IMDEA Software; Dr. José F. Morales, Assistant Professor at UPM and Researcher at IMDEA Software; Víctor Pérez, Software Engineer at Codeplay Software

Energy consumption and the environmental impact of computing technologies is a major worldwide concern. It is a significant issue in systems ranging from energy-hungry server farms to billions of frequently charged smartphones, tablets, smart watches, sensors, and portable/implantable medical devices. As a result of the huge growth in cloud computing, Internet traffic, high-performance computing, and distributed applications, current data centers consume very large amounts of energy, not only to process and transport data, but also for cooling.

Energy consumption is also highly relevant in the context of the Internet of Things paradigm, where very large numbers of small autonomous devices, embedded in all kind of objects, in our clothes, or stuck to our bodies, will operate and intercommunicate continuously for long periods of time, such as years.

Although there have been improvements in battery and energy harvesting technology, a significant reduction in the energy

Aunque ha habido mejoras en la tecnología de baterías y recolección de energía, se necesita una reducción significativa en las demandas de energía de dichos dispositivos para hacer realidad la visión completa de *Internet de las Cosas* y explotar todo su potencial.

A pesar de los rápidos avances recientes en *hardware* de bajo consumo energético, la mayor parte del ahorro potencial de energía se desperdicia por el software que no aprovecha estas características de ahorro de energía ofrecidas por el *hardware* subyacente y realiza una muy deficiente gestión dinámica de tareas y recursos. Para abordar este reto, los investigadores *del Instituto IMDEA Software* han promovido la eficiencia energética como un objetivo de primera clase en el diseño de software, y han desarrollado técnicas y herramientas que facilitan la producción de dispositivos "más ecológicos", es decir, dispositivos con un consumo certificable y más eficiente de la energía disponible y, en general, de los recursos (por ejemplo, tiempo de ejecución o memoria, así como otros recursos definidos por el usuario como accesos a la red o transacciones). En particular, mencionaremos nuestras novedosas técnicas estáticas de obtención de *perfiles*, que son más útiles para el desarrollo de software consciente de los recursos que el análisis del consumo de recursos estándar.

Estas técnicas y herramientas de última generación se implementan e integran en el sistema pionero *CiaoPP*, que proporciona un marco general, formal y práctico (basado en la interpretación abstracta) para predecir con alta precisión los recursos consumidos por un

demands of such devices is needed to make the full Internet of Things vision come true and all its potential be exploited.

In spite of the recent rapid advances in power-efficient hardware, most of the potential energy savings are wasted by software that does not exploit these hardware energy-saving features and performs poor dynamic management of tasks and resources. To face this challenge, researchers at the IMDEA Software Institute have promoted energy efficiency to a first-class goal in software design, and developed techniques and tools that facilitate the production of ``greener'' devices, i.e., devices that make a certifiably more efficient use of their available energy and, in general, of resources (e.g., execution time or memory, as well as other user-defined resources like network accesses or transactions). In particular, it is worth mentioning our novel static profiling techniques, which are more useful for resource-aware software development than standard resource usage analysis.

These state-of-the-art techniques and tools are implemented and integrated into the pioneering CiaoPP system, which provides a general, sound, and practical framework (based on abstract interpretation) for predicting with high accuracy the resources consumed by a given piece of software, for debugging/certifying such consumption with respect to specifications, and for generating dynamic optimization strategies. The system is adaptable to different languages, hardware, and resources because it is built

determinado sistema software, para la *depuración/certificación* de dichos consumos con respecto a las especificaciones, y para la generación de estrategias de optimización dinámica. El sistema es adaptable a diferentes lenguajes, hardware y recursos porque está construido sobre un analizador estático especializable con un *lenguaje de aserciones* versátil, en combinación con técnicas dinámicas para el modelado. Puede ayudar a los programadores a reducir significativamente el uso de recursos de los programas, incluido su consumo de energía y/o el tiempo total de ejecución, lo que resulta en mejoras significativas en la duración de la batería (por ejemplo, en teléfonos inteligentes y otros dispositivos pequeños) o reducciones en el consumo de electricidad (por ejemplo, en *centros de cálculo*).

En estrecha colaboración con la industria, las herramientas conscientes del consumo de energía del Instituto IMDEA Software se integran en sus productos y se prueban en aplicaciones industriales concretas. Esto se ha realizado en parte dentro del contexto de varios proyectos de investigación europeos, nacionales y regionales.

Aunque los ahorros de energía logrados mediante la aplicación de las técnicas y herramientas mencionadas anteriormente pueden ser arbitrariamente grandes, dependiendo de la medida en que un programa esté diseñado teniendo en cuenta la eficiencia energética y utilice características de ahorro de energía, nuestros resultados experimentales con casos de estudio muestran una mayor eficiencia energética en un factor de 6% a 50%. Además, según una estimación de Intel, las técnicas para la eficiencia energética a nivel de *software* pueden obtener ahorros de un factor de tres a cinco veces más allá de lo que se puede lograr a través de las técnicas centradas en el *hardware*.

Un área de aplicación importante en donde nuestras herramientas para la estimación del uso de recursos son bastante relevantes la constituyen las plataformas de contratos inteligentes. Para limitar el tiempo de ejecución, dichas plataformas hacen uso de un concepto denominado *"gas"*, por lo que cada instrucción del lenguaje del contrato inteligente suele tener asociado un coste en términos de este recurso. Si una transacción excede su consumo de gas permitido, se detiene su ejecución y se revierten sus efectos. Sin embargo, incluso si una transacción no tiene éxito debido al agotamiento del *gas*, se incluye en el blockch*ain* y se cobran las tarifas. Del mismo modo, existen limitaciones y costes relacionados con el tamaño del *almacenamiento*. El coste de ejecutar un contrato se puede expresar en términos de estos dos recursos, el *gas* consumido y el *almacenamiento*. En este contexto, conocer de antemano el coste de ejecutar un contrato puede ser útil, ya que permite a los usuarios saber cuánto se les cobrará por la transacción y si se excederán o no los límites de *gas*. Hemos desarrollado un enfoque general que permite el desarrollo rápido, flexible y efectivo de herramientas para la estimación de *gas* y *almacenamiento* para contratos inteligentes, que pueden ser especialmente útiles en el entorno rápidamente cambiante de las tecnologías *blockcha*in.

around a customizable static analyzer with a versatile assertion language, in combination with dynamic techniques for modelling. It can help programmers significantly reduce resource usage of programs, including their energy use and/or total execution time, resulting in significant improvements in battery life (e.g., in smart phones and other small devices), or reductions in electricity consumption (e.g., at data centers).

In close collaboration with industry, IMDEA Software Institute's energy-aware tools are integrated into their products, and tested on concrete industrial applications. This has been done in part within the context of several European, national and regional research projects.

Although the energy savings achieved by applying the techniques and tools mentioned above can be arbitrarily large, depending on the extent to which a program is designed taking energy efficiency into account and uses energy saving features, our experimental results with case studies show increased energy efficient by a factor of 6% to 50%. In addition, an estimate from Intel is that energy-efficient software can realize savings of a factor of three to five beyond what can be achieved through energy efficient hardware.

An important application area where our tools for resource usage estimation are quite relevant is smart contract platforms. In order to limit execution time, such platforms make use of a concept called "gas," so that each instruction of the smart contract language usually has an associated cost in terms of this resource. If a transaction exceeds its allowed gas consumption, its execution is stopped and its effects reverted. However, even if a transaction does not succeed because of gas exhaustion, it is included in the blockchain and the fees are taken. Similarly, there are limitations and costs related to storage size. The cost of running a contract can then be expressed in terms of these two resources, gas consumed and storage. In this context, knowing the cost of running a contract beforehand can be useful, since it allows users to know how much they will be charged for the transaction, and whether gas limits will be exceeded or not. We have developed a general approach that allows rapid, flexible, and effective development of tools for gas and storage estimation for smart contracts, which can be specially useful in the rapidly changing environment in blockchain technologies.

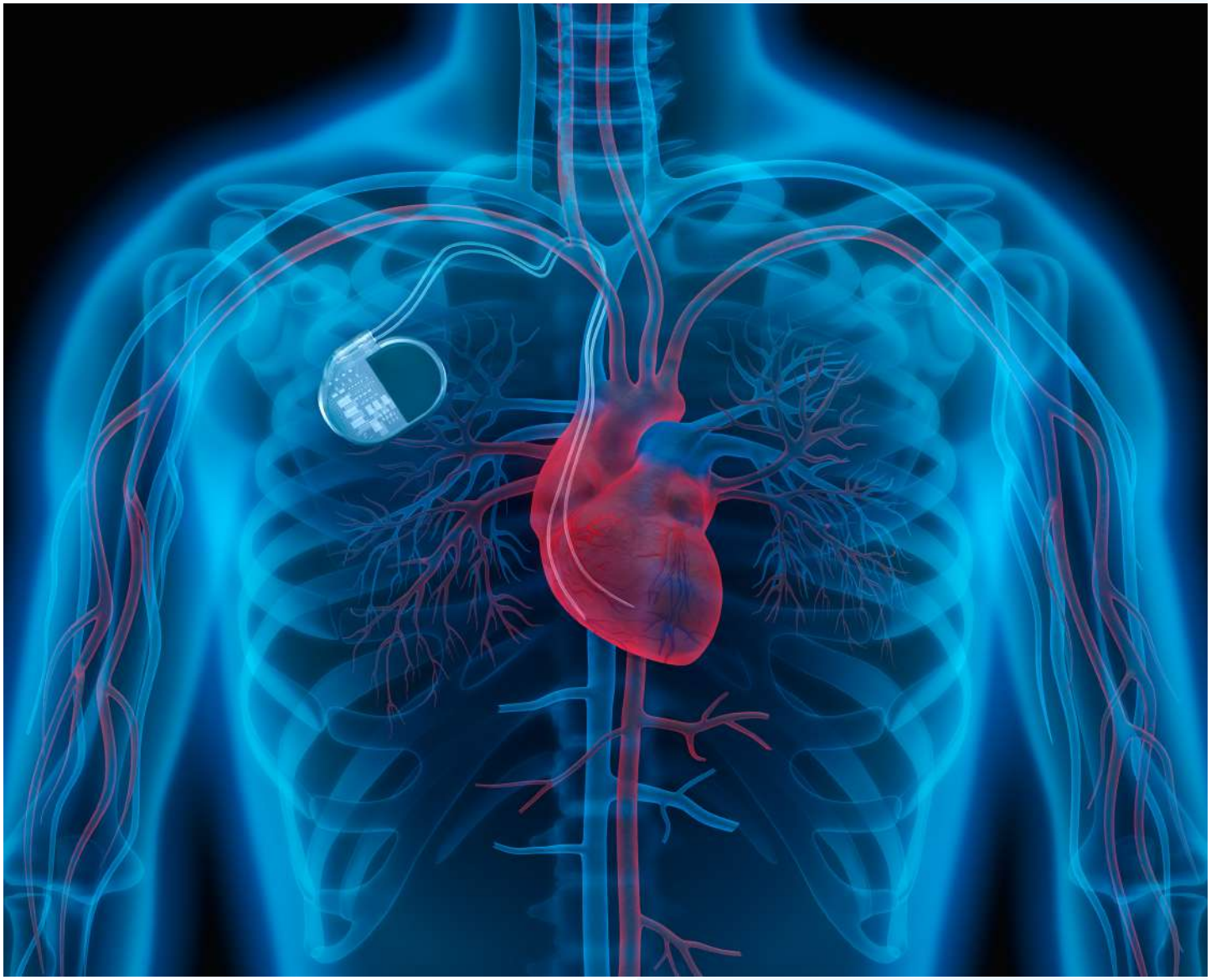
Current work focuses on maturing and extending the tools for analysis, verification, and optimization, in order to enable engineers to understand and quantify the impact of design decisions on energy, and developing a set of recommendations for the integration of such energy-aware tools in the software life-cycle.

Our research is aligned with some of the Sustainable Development Goals from the UN, such as: take urgent action to combat

Nuestro trabajo actual se centra en robustecer y ampliar las herramientas de análisis, verificación y optimización, a fin de permitir a los *ingenieros de software* comprender y cuantificar el impacto de las decisiones de diseño en la energía, y desarrollar un conjunto de recomendaciones para la integración de tales herramientas conscientes del consumo de energía en el ciclo de vida del software.

Nuestra investigación está alineada con algunos de los Objetivos de Desarrollo Sostenible de la ONU, tales como: tomar medidas urgentes para combatir el cambio climático y sus impactos, garantizar patrones de producción y consumo sostenibles, garantizar una vida saludable y promover el bienestar para todos en todas las edades, y hacer que las ciudades y los asentamientos humanos sean inclusivos, seguros, resilientes y sostenibles.

climate change and its impacts, ensure sustainable consumption and production patterns, ensure healthy lives and promote well-being for all at all ages, and make cities and human settlements inclusive, safe, resilient and sustainable.



La red madrileña ultrarrápida de comunicaciones avanzadas para la investigación y la innovación (REDIMadrid)

The Ultra-Fast Madrid Network for Research and Innovation (REDIMadrid)

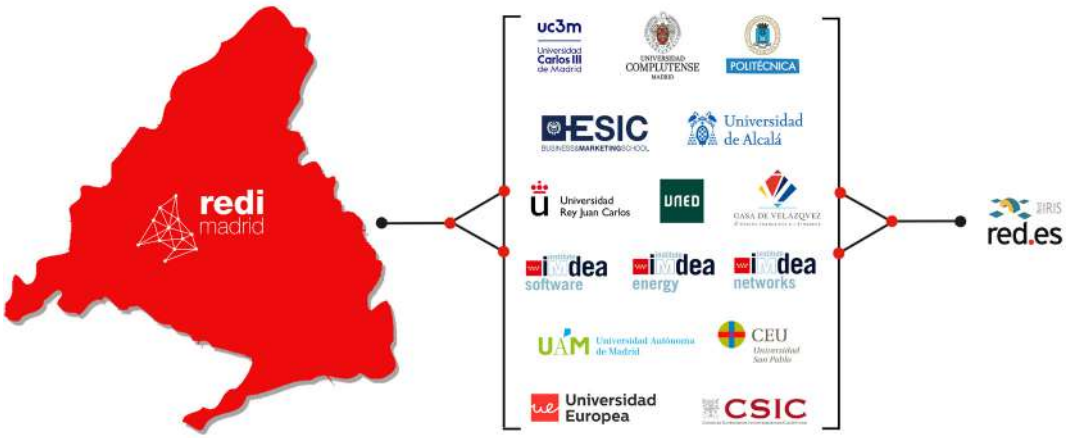
El Instituto IMDEA Software diseña, despliega y gestiona REDIMadrid, la red de datos de alta velocidad que conecta a las instituciones de investigación y educación superior de la Comunidad de Madrid. Esta red es una base fundamental para la investigación avanzada en todas las áreas del conocimiento y posibilita la investigación específica en redes de ordenadores, incluyendo el despliegue pionero en comunicaciones cuánticas *MadQuantum-CM.*

© César Sánchez, 2022

Para más información pulsa aquí

The IMDEA Software Institute designs, deploys, and maintains REDIMadrid, the high-speed network that connects research institutions, universities, and high education centers within the region of Madrid. This network is a cornerstone for advanced research in all areas of knowledge and enables as well performing specific research in computer networks, including flagship projects like *MadQuantum-CM*, the future Madrid quantum communication infrastructure.

For more information click here



madrimasd.org

suplemento
notiweb

instituto IMdea 15 ANIVERSARIO

institute IMdea software

fundación para el conocimiento madri+d