Resumen ejecutivo

Título

Contribuciones desde la ingeniería de software y sistemas basada en modelos a la privacidad y la protección de datos mediante un enfoque multidisciplinar.

Autor

Yod Samuel Martín García

Objetivo

Este trabajo pretende ofrecer **métodos y herramientas dirigidos a los ingenieros que desarrollan software y sistemas**, para operativizar los principios y la regulación de la protección de datos, integrándose con aquellos métodos y herramientas que los ingenieros ya usan de manera cotidiana en las diferentes actividades que tienen lugar a lo largo de las etapas del ciclo de vida del desarrollo del sistema, para que, en última instancia, y siguiendo los **principios de responsabilidad proactiva y protección de datos por diseño**, puedan crear sistemas que cumplan con el RGPD, se adhieran a los principios de protección de datos y cuiden los derechos de los interesados. En particular, nos hemos centrado en *integrar la privacidad y la protección de datos en varias disciplinas de ingeniería: gestión de riesgos, ingeniería de requisitos, diseño, aseguramiento de sistemas, e ingeniería de métodos*.

Motivación

Los ingenieros son en última instancia los responsables de elaborar, construir y mantener los sistemas, servicios y productos de software y hardware que dan soporte a los tratamientos de datos personales sujetos al RGPD, y de implantar las medidas técnicas para protegerlos. Sin embargo, *los ingenieros carecen de métodos y herramientas cercanos a su mentalidad*, que les permitan introducir sistemáticamente las soluciones de privacidad y protección de datos en del flujo de trabajo al que están acostumbrados. Por ello, nuestra visión es que los ingenieros puedan crear servicios que cumplan con la privacidad… sin tener que ser expertos en ella. Así, asumimos la misión de trasladar los conocimientos establecidos de privacidad y protección de datos a la práctica general de la ingeniería de software y sistemas.

Diseño y metodología

El trabajo responde a dicha misión desde el cuerpo de conocimiento y la experiencia de la ingeniería de software y sistemas mediante métodos, herramientas y técnicas alineadas con las que los ingenieros trabajan en su práctica cotidiana. En particular, aprovechamos el enfoque de la *ingeniería de software y sistemas basada en modelos (MBSSE)*, que ha demostrado su adecuación para organizar sistemáticamente el conocimiento del sistema y de su contexto, integrar puntos de vista propios del sistema con otros orientados a aspectos transversales, y facilitar el procesamiento humano o automatizado de los modelos.

No obstante, hemos tratado de abordar el trabajo desde un *enfoque metodológico multidisciplinar*, combinando métodos típicos de las ciencias sociales y humanidades (análisis de contenido cualitativo), desarrollo de productos (modelos mentales), e ingeniería (ingeniería de software y sistemas basada en modelos MBSSE o, teoría de la información, etcétera). Hemos llevado a cabo análisis legales detallados tanto desde la perspectiva interna del RGPD (y las guías auxiliares del EDPS), como en comparación con la legislación de la UE para otros requisitos, y en relación con estándares industriales (ISO 29100, ISO 27550, ISO 29134, etcétera).

Resultados principales

El trabajo presenta las siguientes *contribuciones*, que muestran la introducción la privacidad y la protección de datos a lo largo del ciclo de desarrollo:

- Una *definición operativa de privacidad* y protección de datos, y un *modelo de adversario* apropiado para la ingeniería de privacidad y de protección de datos.
- Un *análisis de las brechas entre las necesidades de los ingenieros* para implementar los principios de privacidad y protección de datos (y lograr el cumplimiento del RGPD de la UE), y el respaldo ofrecido por las herramientas de gestión de la privacidad, más una propuesta para apoyar estas necesidades desde las disciplinas de ingeniería mencionadas.
- Un *lenguaje específico de modelado aspectos de dominio (DSAL)* independiente de la implementación, para anotar una variedad de modelos de sistema con información de características de privacidad y protección de datos.
- Una *definición operativa de la minimización de datos* en múltiples dimensiones cuantificables a partir de los riesgos sobre los derechos de los interesados, y su aplicación a una serie de *métricas de riesgos de privacidad* de ámbito organizacional.

- Un marco metodológico de *gestión de requisitos* para abordar los requisitos de privacidad y protección de datos desde una doble perspectiva (basada en riesgos y orientada a objetivos) y la aplicación de esta última para operativizar los contenidos de la norma ISO/IEC 29100.
- Un sistema de *patrones de diseño* de privacidad, incluidos elementos morfológicos (estructura), elementos sintácticos (relaciones) y elementos léxicos (instancias).
- Un método para el *aseguramiento y la garantía de privacidad y protección de datos* (especialmente, evaluaciones de impacto de privacidad y protección de datos EIPD), con un marco de referencia orientado a procesos, patrones de argumentación, y correspondencia entre la ley (GDPR), estándares técnicos (ISO/IEC 29134) y pautas industriales (plantilla EIPD para red eléctrica inteligente).
- Un *metamodelo metodológico* para los métodos de ingeniería de privacidad y protección de datos, un conjunto de procesos de privacidad y protección de datos que se introducirán a lo largo del ciclo de vida de desarrollo del sistema (SDLC), sus interacciones, y la arquitectura de un conjunto de herramientas de software para respaldar ese proceso.

Novedad y aplicabilidad

Los *paradigmas dominantes para abordar la privacidad y la protección de datos tienden a ignorar el papel de los desarolladores* de productos, sistemas y servicios; por ejemplo, (1) la privacidad y la protección de datos desde la regulación (p.ej. RGPD) emplean conceptos ajenos al ámbito de la ingeniería; (2) los principios de privacidad y protección de datos por diseño (PbD / DPbD) resultan abstractos, aspiracionales, y carentes de un correlato técnico operativo; (3) las tecnologías de mejora de la privacidad (PETs, en sus distintas variantes como 'PETs duras' u orientadas a la minimización y 'PETs blandas' u orientadas a la transparencia) en realidad ofrecen un enfoque 'artesanal' para abordar la privacidad y protección de datos, ya que no proporcionan los métodos sistemáticos y económicos que serían esperables de un proceso ingenieril; y (4) las herramientas software de gestión de privacidad y protección de datos abarcan otras áreas corporativas como la Tecnología de la Información o la Dirección Operativa, pero no abordan el proceso de desarrollo.Todo ello hace que los ingenieros de desarrollo de sistemas rara vez consideren la privacidad entre sus preocupaciones más importantes, a pesar de la relevancia que debería tener.

Por ello, abogamos por *complementar estos paradigmas mediante la introducción de la ingeniería de privacidad a lo largo del ciclo de desarrollo y la práctica de la ingeniería del software y sistemas*. Nuestro trabajo aborda cómo se puede integrar la implementación del marco legal con las principales herramientas y métodos que los ingenieros usan y aplican de manera cotidiana, adaptándose a sus necesidades para ayudarles en el cumplimiento del RGPD a lo largo de todo el ciclo de vida del desarrollo de software y les ofrece métodos para que implementen de manera sistemática los controles de privacidad más adecuados sin que tengan que ser expertos en privacidad ni investigar a fondo caso por caso.

Nuestras contribuciones han quedado validadas por la aplicación de sus resultados a distintos proyectos de ámbito europeo y con una fuerte participación industrial, como PRIPARE, TRUESSEC.EU y PDP4E. En ese último (de 3.4 M€ de presupuesto), el autor ha sido coordinador científico técnico, y *distintas organizaciones han trasladado las propuestas de nuestro trabajo y las han aplicado a herramientas software destinadas a ingenieros*, introduciendo aspectos de ingeniería de privacidad en herramientas software de desarrollo ya existentes (DST, Papyrus, OpenCert), de código fuente abierto, integradas en el ecosistema Eclipse, y con las que los ingenieros ya están familiarizados. Más aún, los socios industriales han aplicado los resultados a sendos pilotos en la industria energética y la automovilística. Y el enfoque basado en modelos que introducimos en la arquitectura es clave para lograr la adaptabilidad, flexibilidad, reusabilidad e interoperabilidad de las herramientas, fomentando así su adopción y proyección futura.

Asimismo, *hemos lanzado varias iniciativas para la adopción de nuestras contribuciones*: (1) la *estandarización en ISO/IEC* a través del *PWI 27564 Privacy models*, (2) la creación de un *grupo de interés en la Eclipse Foundation* para el desarrollo de una serie de modelos de ingeniería de privacidad abiertos y reutilizables, y (3) un documento de *orientación de políticas* (policy brief) enviado a la Comisión Europea para informar la manera más efectiva de articular las decisiones políticas para facilitar su implementación por los ingenieros.

Conclusiones

Esperamos que el trabajo redunde en una mayor adopción de los principios de privacidad en el desarrollo de sistemas y servicios. Es necesario responder al desafío (reconocido por el EDPS y ENISA) de llevar a la práctica el paradigma de la "privacidad desde el diseño" para asegurar su viabilidad e impacto, dotando a los ingenieros de herramientas para que apliquen la privacidad y protección de datos. Si estas herramientas se integran con su práctica habitual, evitaremos sus resistencias a la adopción de la privacidad en sus desarrollos, reduciremos sus costes de cumplimiento normativo, y redundará en una mejorará la confianza en la tecnología por parte de los ciudadanos.

Executive Summary

Title

Contributions from model-based software and systems engineering to privacy and data protection through a multidisciplinary approach.

Author

Yod Samuel Martin Garcia

Objective

This work aims to offer **methods and tools aimed at engineers who develop software and systems**, to operationalize the privacy and data protection principles and regulation, integrating them with those other methods and tools that engineers already use on a daily basis in the different activities that take place throughout the stages of the system development life cycle (SDLC), so that, ultimately, and sticking to the **principles of accountability and data protection by design**, they are able to create systems that comply with the GDPR, adhere to the principles of data protection and take care of the rights of the data subjects. In particular, we have focused on *integrating privacy and data protection into several engineering disciplines: risk management, requirements engineering, design, systems assurance, and method engineering.*

Motivation

Engineers are ultimately responsible for developing, building and maintaining the systems, services and software and hardware products that support the personal data processing activities that are subject to the scope of the GDPR, and for implementing technical measures to protect such data. However, *engineers lack methods and tools close to their mindset* that allow them to systematically introduce privacy and data protection solutions into the workflow they are accustomed to. That's why our vision is that engineers be able to create services that comply with privacy... without needing to be experts in that, for which we assume the mission of transferring the established knowledge and wisdom on privacy and data protection to the general practice of software and systems engineering.

Design and methodology

The work responds to such mission from the body of knowledge and experience of software and systems engineering, through methods, tools and techniques aligned with those that engineers employ in their daily practice. In particular, we take advantage of the *Model-Based Software and Systems Engineering (MBSSE*) approach, whose suitability has been proven to systematically organize knowledge about the system and its context, integrate system-specific points of view with others that address cross-cutting aspects, and facilitate human or automated processing of models.

Nonetheless, we have tried to approach our work from a *multidisciplinary methodological approach*, combining methods usually employed in Social Sciences and Humanities (Qualitative Content Analysis), product development (Mental Models), and Software and Systems Engineering (MBSSE or model-based software and systems engineering, information theory, etc.). We have carried out detailed legal analyses both from the internal perspective of the GDPR (and the EDPS ancillary guides), and in relation to EU legislation for other requirements, as well as compared to industry standards (ISO 29100, 27550, 29134…)

Main results

Our work has yielded the following *contributions*, which show the introduction of privacy and data protection throughout the development cycle:

- An *operational definition of* privacy and data protection, and a *suitable adversary model* for privacy and data protection engineering.
- An *gap analysis between engineers' needs* to implement privacy and data protection principles (and comply with EU GDPR), and the support offered by privacy management tools, plus a proposal to support these needs from the engineering disciplines mentioned.
- A specific *domain-aspect modeling (DSAL) language* independent of implementation, to annotate a variety of system models with information on privacy and data protection features.
- An *operational definition of data minimization* into multiple, quantifiable dimensions, derived from the risks to the data subjects' rights, and their application to a series of organization-wide, *privacy risk metrics*.
- A methodological framework for *requirements management* to address privacy and data protection requirements from a dual perspective (risk-based and goal-oriented) and the application of the latter to operationalize the contents of ISO/IEC 29100.

- A system *of privacy design patterns*, including morphological elements (structure), syntactic elements (relationships), and lexical elements (instances).
- A method for *ensuring and guaranteeing privacy and data protection* (especially privacy impact assessments and DPIA data protection), with a framework oriented to processes, argumentation patterns, and correspondence between the law (GDPR), technical standards (ISO/IEC 29134) and industry guidelines (EIPD template for smart grid).
- A *methodological metamodel* for privacy and data protection engineering methods, a set of privacy and data protection processes that will be introduced throughout the system development lifecycle (SDLC), its interactions, and the architecture of a set of software tools to support that process.

Novelty and applicability

Dominant paradigms for addressing privacy and data protection *tend to ignore the role of product, system and service developers*. For instance, (1) privacy and data protection from regulation (e.g. GDPR) employ non-engineering concepts; (2) privacy and data protection by design (PbD /DPbD) principles are abstract, aspirational, and lack an operational technical correlate; (3) Privacy enhancement technologies (PETs, in their variants such as 'hard PETs' or minimisation-oriented and 'soft PETs' or transparency-oriented) actually offer a 'craftsman's' approach to addressing privacy and data protection, as they do not provide the systematic and economical methods that would be expected from an engineering process; and last, (4) privacy and data protection management software tools cover other corporate areas such as Information Technology or Operational Management, but do not address the development process. All of this means that systems development engineers rarely consider privacy among their most important concerns, despite the relevance it should have.

Therefore, we vouch for complementing these paradigms by *introducing privacy engineering throughout the development cycle and the practice of software and systems engineering*. Our work addresses how the implementation of the legal framework can be integrated with the mainstream tools and methods that engineers use and apply on a daily basis, adapting to their needs to help them comply with the GDPR throughout the entire software development lifecycle (SDLC). Plus, we offer them methods to have the most appropriate privacy controls implement consistently, without either having to be privacy experts or thoroughly investigate on a case-by-case basis.

Our contributions have been validated by the application of their results to different projects at European level, with a strong industrial participation, such as PRIPARE, TRUESSEC.EU and PDP4E. In the latter (of 3.4 M€ of budget), where the author has been the scientific and technical lead, *different organizations have transferred the proposals from our work and have applied them to software tools aimed at engineers*, introducing aspects of privacy engineering in already existing open-source software development tools (DST, Papyrus, OpenCert), which are integrated into the Eclipse ecosystem, and with which engineers are already familiar. Moreover, these industrial partners have applied the results to pilots in the energy and automotive industries. And the model-based approach we introduce into our solutions' architecture is key to achieving the adaptability, flexibility, reusability and interoperability of tools, thus fostering their adoption and future projection.

We *have also launched several initiatives for the adoption of our contributions*: (1) *standardization in ISO/IEC* through *PWI 27564 Privacy* models, (2) the constitution of an *interest group at the Eclipse Foundation* for the development of a series of open and reusable privacy engineering models, and (3) a *policy brief sent to the European Commission* to inform the most effective way to articulate policy decisions to facilitate their implementation by engineers.

Conclusions

We hope that the work will result into greater adoption of privacy principles in the development of systems and services. It is necessary to respond to the challenge (as recognized by the EDPS and ENISA) of **bringing the paradigm of "privacy by design" into practice**, to ensure its viability and impact, by providing engineers with tools to apply privacy and data protection. If these tools are integrated with the daily practice they are already used to, we will elude their resistance to adopting privacy in their developments, reduce their compliance costs, and ultimately improve citizens' trust in the technology.