smashHit

Smart Dispatcher for Secure and Controlled sharing of Distributed Personal and Industrial Data

Public Innovation Concept

Public Report D1.3 Public Innovation Concept

This document comprises the smashHit innovation concept as a whole, leading from today's constraints in consent/contract processes to how smashHit will face those challenges with its innovative trusted and secure system concept.

MARCH 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No *871477*

Foreword

Welcome to our smashHit Public Innovation Concept. From the very first, we were absolutely convinced that the growing Data Economy has to become more attractive for its key stakeholders (data owners, data providers and data processors) to overcome existing barriers, as e.g. the complicated and time-consuming consent/contract processes, hindering to build-up innovative services using data from multiple sources.

With the growing ability of Cyber Physical Products (CPP) to generate, gather and share data with third parties among different data-sharing platforms, there will be a general need for flexible and easily manageable procedures to handle data owner consent and legal rules, to achieve effective and traceable contracting. The challenges and complexities of the General Data Protection Regulation (GDPR) directives make cumbersome mechanisms necessary to gain, record and manage consent. Also, data owners are wary about improper use of their data. The combination of understanding and relating to the value proposition, consumer trust, and cumbersome consent processes, results in a low opt-in rate for connected CPP data exchange (e.g. data from cars) and prevent the creation of innovative services (e.g. connected insurance programs, or smart city solutions).

Thus, we have conceptualised the smashHit System Concept as a trusted, secure and integrating privacy-by design reference Framework to simplify the consent/contract process, as well as to enable consent/contract tracing and sharing among multiple data-platforms. In addition, smashHit will offer solutions to identify data misuse as well as to support data processors in creation of legally binding contracts.

All along, we have followed the maxim to think about the needs of Data Owners and Data Customers, but also to win CPP manufacturers (e.g. car makers) to open up their products, by designing a convincing trustworthy Ecosystem.

Recently we have finalized the smashHit system concept and have started its detailed specification and development by our software and RTD development partners. Several public presentations of our smashHit system concept have been presented (see smashHit project website).

In this report you will find some more details about the smashHit system concept as a whole, leading from today’s constraints in consent/contract processes to how smashHit will face those challenges with its innovative trusted and secure system concept.

If you got curious about how all that is made possible, just continue on the following pages, enjoy the reading, and please contact us with your feedback or questions!

Table of Contents

1 smashHit Framework solution concept

The path towards the smashHit framework led through several thoughts as indicated in Figure 1.

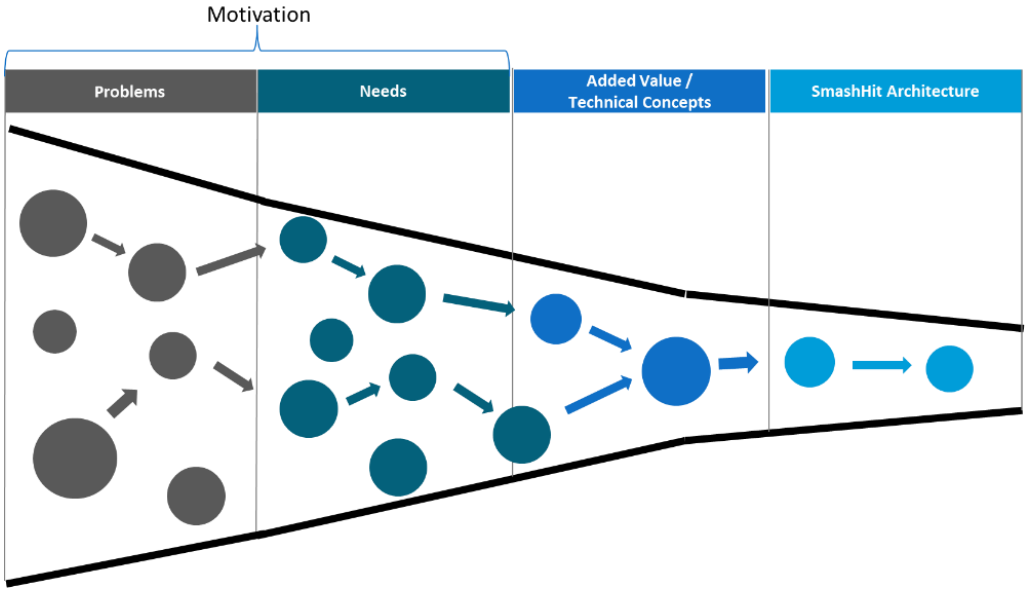

Figure 1: Path towards the smashHit Framework

A range of selected key problems and needs in the state-of-the-art solution for consent/contract handling in the data economy led to the starting point of the smashHit project. Driven by these needs, innovative approaches are compiled to build the smashHit architecture as key element of the project.

The following sub-chapters describes incrementally the outcomes along this way towards the overall smashHit framework.

1.1 Motivation (Problems & Needs)

In the last years many data sharing platform approaches for different purposes (like B2B data exchange, data marketplaces or data processing services) emerged, giving cross-sectorial industries access to the great spectrum of sensor data coming from high volume products from various industrial sectors (vehicles, smart home devices, smart cities data etc.). With the increasing number of connected sensors and actuators within such mass products (so called cyber physical products), this number will rise exponentially in short-term. This enormous amount of data offered by various data sharing platforms represent:

- a massive information resource to create new value, allowing the improvement of existing services or the establishment of diverse new innovative services, by combining data streams from various sources
- a major big data-driven business potential, not only for the manufacturers of Cyber Physical Products (CPP), but in particular also for cross-sectorial industries and various organisations with interdisciplinary applications

The majority of these data sharing platform approaches offer a scalable, secure and trusted sharing of data. However, they all lack in the provision of solutions for consent creation, management and observation between two or more involved actors that aim to share Personally Identifiable Information (PII) data. To fully exploit data sharing approaches in practice, they have to consider national and international laws like the European General Data Protection Regulation (GDPR). Among others, such regulations define that data owners can decide to whom and for which purposes their data can be used; they define how data providers have to handle data and they define

how data have to be processed. A data sharing platform that considers such regulations needs an additional layer of consent creation linked to a contract which would guarantee the compliance with all requirements of the contracting parties and national and international laws. This gap needs to be filled and is where smashHit aims to bring an added value.

1.1.1 Problems

As mentioned above, the rapid growth of the Data Economy also reflects on the number of B2B, B2C and B2G data-sharing platforms recently launched, which are contributing to fragment the market, preventing business opportunities due to lack of interoperability, inconsistent consent and legal rules among different data-sharing platforms actors and operators. For instance, Figure 2 identifies a set of different platform typologies (where diverse data, both closed and open, are streaming) that already operate on data markets, contributing to a fragmented data-sharing landscape, while data consumers look for multi data sources collaborations to exploit the added-value business based on combining data providers and integrating data from diverse platforms.

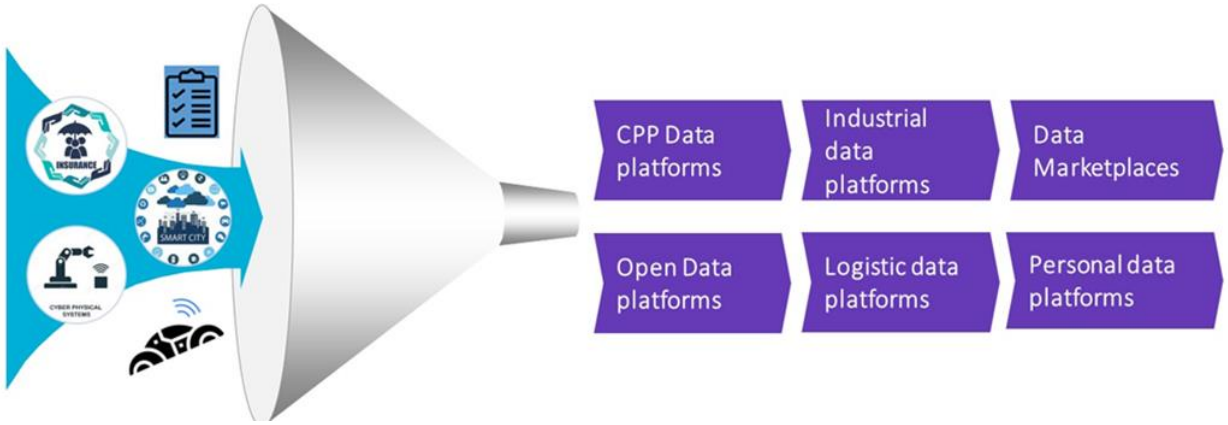

Figure 2: Fragmented data sharing landscape

There is a short-term challenge for the data economy to look for an interoperable data-sharing framework to enforce and manage multi-platform agreements for exchanging data whilst protecting and assuring compliance with GDPR, Privacy and Security Policy enforcement rules of individual data providers, national (country) data privacy and protection rules and EU-legal directives and legislation surrounding personal and industrial data generation, storage and sharing. Due to this high level of market fragmentation, the situation today is characterized by far too complex and individual value chains resulting in economic inefficiency. Therefore, the use and integration of data from various platforms is limited due to missing solutions for agreement on consent, legal rules, effective contracting, data use traceability, security & privacy issues etc.

The requirement analysis at beginning of project has shown, that this situation is mainly characterised by the following major challenges and barriers for all stakeholders in the value chain, the service providers (data processors) and for the CPP manufactures (data providers) and the CPP owners (data owners):

- The number of steps required to gain consent create customer “friction”.
- Time taken to obtain the required data is too long.
- Verifying a person giving consent is often insufficient as well as duplicated at each consent step.
- Problem of broken consent chain is unsolved.
- Lack of tools to prevent data misuse prevent willingness of data providers to provide data.
- Insufficient control for data owners over their data.
- Missing support in creating legally binding contracts.

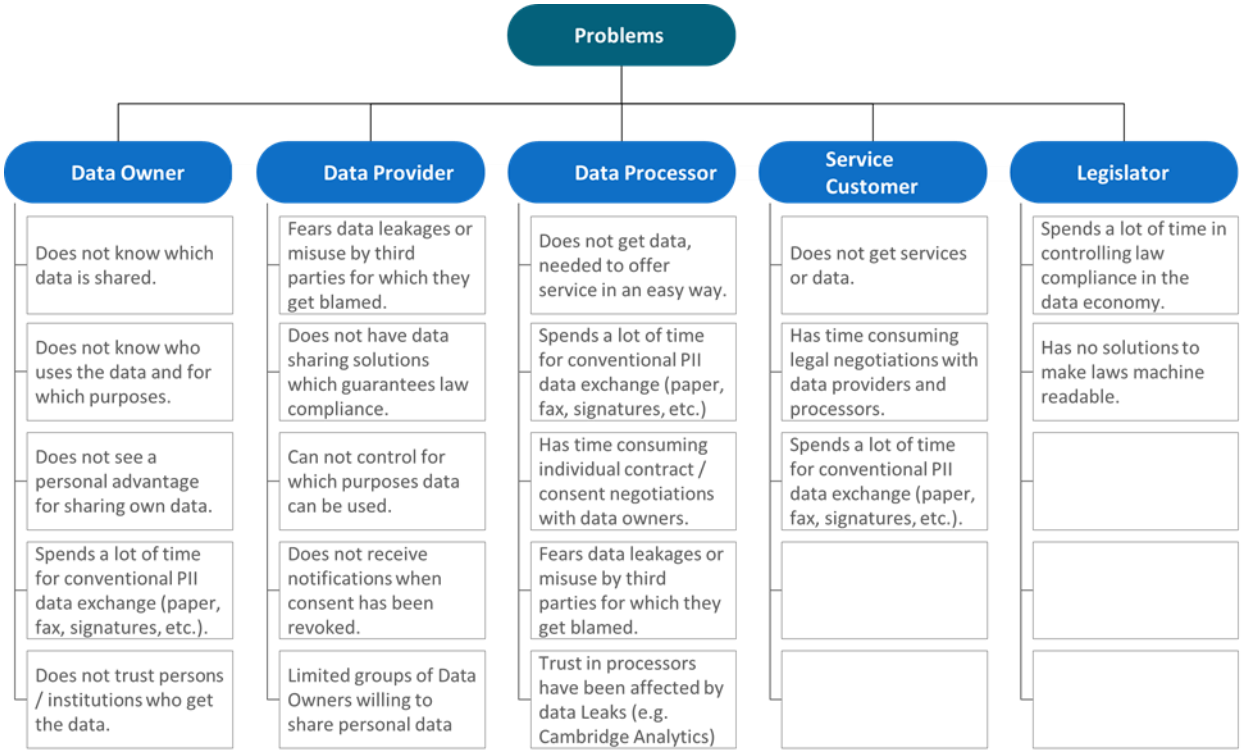

Figure 3: Selected overall problems of different data economy stakeholders

Figure 3 introduces a range of overall problems that lead to a bottleneck in the current data sharing economy. Although the majority of people already sees the potential behind a data economy, there are still several doubts based on the perspective/role of different participants. The problems for the different data economy stakeholders are described in more detailed below:

Data Owners (like a car owner/driver who has ownership of produced car data while driving) claim that they often do not know and cannot reconstruct who gets which data, for which purposes they are used and how long is the contract runtime. They also often just simply do not trust the organisations in the data economy because the verification of a person giving consent is often insufficient and the contracts are too cryptic and out of their control. They are also often annoyed about steps required to gain consent to get just one service, because they have duplicated consent steps.

Data Providers (e.g. OEMs) who collect data and/or produce and sell smart (cyber physical) products that produce sharable data. The missing possibility for reconstruction of data leakage is one of the most challenging obstacles for data providers that has to be overcome. In current data market solutions Data Providers (like VW) are unaware of the data processing chain. In case of a criminal data misuse, the absence of features to continuously monitor the data traceability represents a key barrier, which does not allow the Data Providers to identify the origin of a data leakage and to prevent GDPR punishments. Without clear indications regarding the data leak origins, the Data Provider is liable for the damages and penalties under GDPR. Even more critical for large CPP manufactures is the danger of reputation damage threat after breach. It's in the nature of things, that if vehicle data is misused, the OEM as a producer of the vehicle and key instance responsible for gathering vehicle data, will be in the spotlight of the suspicion.

A further important aspect connected with consent traceability is a problem which could occur if a product (e.g. vehicle) is sold and the owner did not revoke the given consent for gathering data from this product. In this case the Data Provider (OEM) does not get an alert and will carry on with gathering data from a product without having the consent to do that. Although the OEM has neither the responsibility nor the possibility to monitor if given consent is broken, this case has also a high risk of reputation damage for the Data Provider.

Data Processors who needs data from various data platforms to build services upon them. A key challenge for them is to simplify the process to obtain consent and easy contracting. For a lot of services (e.g. usage based vehicle insurance: UBI) one of the key problem is due to the number of partners and steps required to gain consent, which create customer “friction” and is a deterrent to providing consent. For such type of services, the consent also requires additional personal information in order to verify the person giving consent is indeed the product/data owner. The request for this data is duplicated at each consent step which is irritating for the consumer. Also, the time taken to obtain the required data, and then process the data into a quotation is too long if consent is not provided prior to the request for insurance. More and more consumers shop for insurance on-line. The time for the whole end-to-end process from requesting a quotation to getting it, needs to be sub-second at worst. If the process of buying insurance is interrupted for the consumer to change web sites and provide consent to the OEM, then the consumer is likely to buy elsewhere or not buy a “connected” policy. Moreover, Data Processors have similar problems w.r.t. data traceability as described for Data Providers, e.g. regarding identification of broken consent chain or backtracking of data misuse. A special problem exists for SME or start-ups, who needs support in creating legally binding contracts, assuring compliance with GDPR, Privacy and Security Policy enforcement rules.

Potential **Service Consumers** claim similar problems as the Data Owners, which are related to complex consent/contract processes or missing traceability of data usage, especially for services that use PII data because they need time-consuming legal negotiations with all actors in the chain. At the end, all service consumer’s problems are problems/challenges for the data processors, since these obstacles will hinder the service consumers to give consent to an offered service.

Finally shown are **Legislators** like the European Union who claim that the controlling of law compliance in the data economy is hard to do because there are no adequate technical solutions available. Currently, there is no bridge between the analogue and the digital machine world which leads to the situation that laws are not interpretable by machines and always need human support acting as an interpreter.

1.1.2 Needs

In contrast to today's fragmented data sharing landscape, which is characterised by non-heterogeneous technical designs and proprietary implementations and by this means is hindering the use and integration of data from various platforms due to missing solutions for effective consent/contracting and data use traceability etc., the smashHit Project focuses on user oriented solutions that make Data Marketplaces more attractive and trustful for its key stakeholders. Therefore, smashHit has to overcome several obstacles by establishing an inter-data-platform framework solution enabling simplified multi-platform Consent/Contract processes for exchanging data also ensuring a trustful data use traceability. To achieve these key challenges, smashHit needs to develop the following main characteristics:

- Improve Citizens trust
  - by means of trace the use of data over diverse platforms
- Improve OEM & Data Customer trust by means of:
  - fingerprinted data to ensure traceability/unchangeability along value chain and data quality
  - Consent tracing -> notification of all involved contractors in case of broken consent chain
- Simplify consent process through:
  - authentication of the individual (certification of consent)
  - single point of consent (management and distribution of consent)
- Support in Consent/Contract generation:
  - legally binding Contracts, taking into account relevant legislations/legal rules

Such an innovative inter-data-platform framework solution will have several positive effects on the European Data Economy, as e.g.:

- To overcome existing barriers, as e.g. the complicated and time-consuming consent/contract processes
- Increase trust for all stakeholders in the data market, due to strong data traceability mechanism
- Increase willingness of Data Providers (OEM) to provide CPP data for 3rd party use
- Open new opportunities for value-creation by creation of innovative services using data from multiple platforms

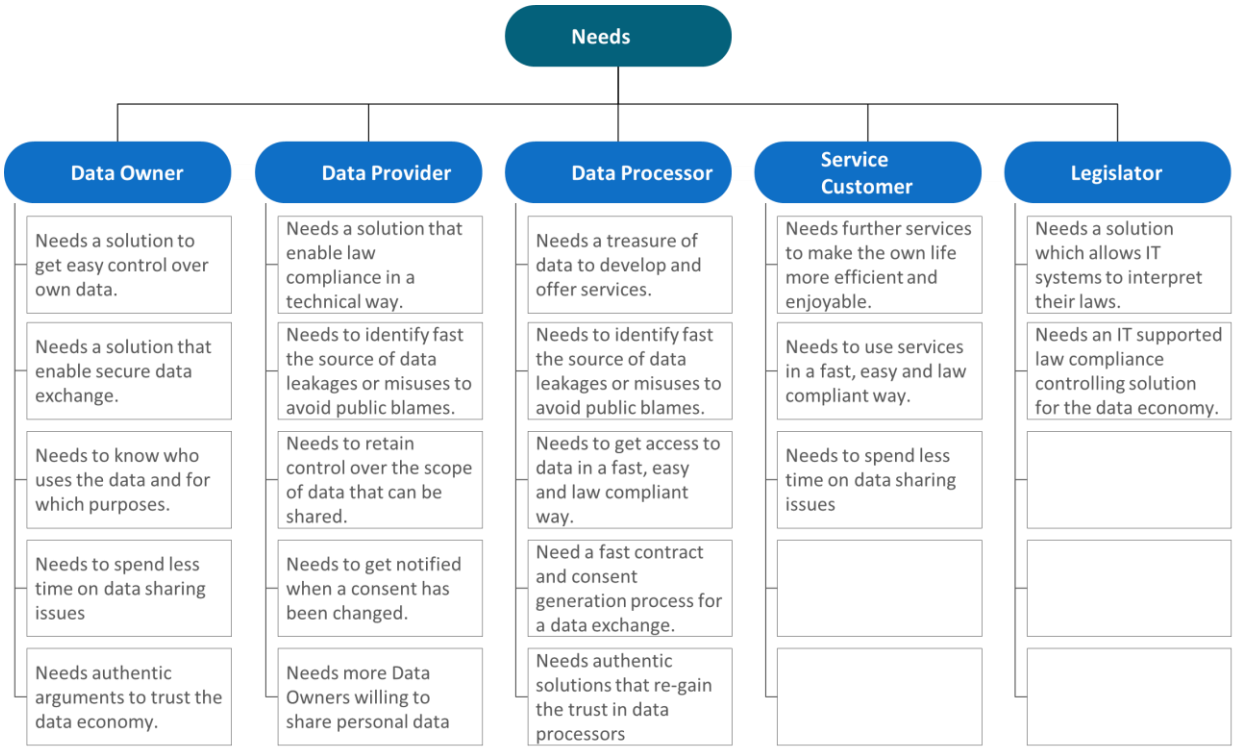

Figure 4: Selected overall needs of different data economy participants

Figure 4 introduces a range of derived needs that (when satisfied) would open the addressed bottlenecks in the data sharing economy.

Data owners need a solution that would enable them to easily trace the use of their data over diverse platforms. The solution must provide the data owner the power to manage their data, enabling them to identify for each of their products, who gets which product data, for which purposes they are used and how long is the contract run-time. In addition the solution has to simplify consent processes, to create the trust and motivate the users to share their data by offering solutions users find useful and necessary.

Data Providers (e.g. OEMs) need a solution that allows to identify the origin of a data leakage in case of data misuse in an efficient way. Further they need to have the possibility to monitor if given consent is broken. Both cases have a high risk of reputation damage for the Data Provider.

Generally, we can state, that a reliable solution to eliminate the above data traceability challenges related with 'reconstruction of data leakage' and minimisation of 'Reputation Damage Risks' will considerably increase the willingness of large CPP manufactures to provide more data from their products.

Data Processors are offering services based on data from various data platforms. A key request for them is to simplify the process to obtain consent and easy contracting. Moreover, Data Processors have similar needs w.r.t. data traceability as described for Data Providers, e.g. regarding identification of broken consent chain or backtracking of data misuse.

Service Customers have similar need as the Data Owners, as e.g. simplified consent/contract processes and a clear traceability of data usage for all their products.

Finally, **Legislators** need a solution which enables to apply their laws in a technical way. A unified solution would also simplify the controlling of law compliance in the data economy.

1.2 smashHit added value and technical concepts

smashHit is aiming to design and implement a trusted, secure and privacy-by design reference Framework to simplify the consent/contract process as well as to enable consent tracing and sharing among multiple data-platforms. In addition, smashHit will offer solutions to identify data misuse as well as to support data customers in creation consent. The solution shall address the identified needs of different actors in the data economy. Therefore, smashHit aims to provide answers for these needs compiled in form of a smashHit Framework which can be divided into five different main building blocks shown in Figure 5. These building blocks are partly using basic concepts of state-of-the-art technologies that can be used for further exploitation to fill the identified gaps.

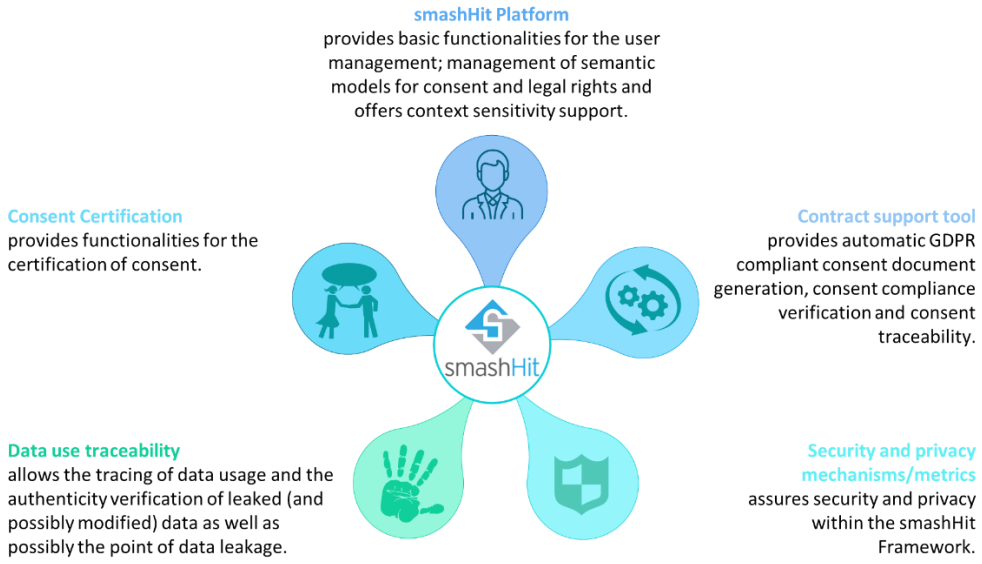

Figure 5: Main modules of the smashHit framework

smashHit Platform

The smashHit platform provides basic functionalities in the smashHit framework. Different actors aim to get consent creation, management and observation functionalities. For this requirement it is necessary to create basic user management functionalities that provides access to the smashHit framework. Further it is needed to describe an approach to CRUD information related to users, laws or consent. The smashHit Platform will exploit a range of state-of-the-art technologies to achieve these requirements.

Each actor which is using the smashHit platform has to be a trusted participant who must prove the identity of his person.

To describe the handled data in the smashHit framework will be used an ontology approach. smashHit will integrate state of the art ontologies as basis to model user accounts, consent, laws and other concepts that will appear within the project. The ontologies will be used as a schema to build the knowledge graph/semantic data model.

The world is complex and context dependent so that many situations cannot be handled always in a static way. An integrated context sensitive system design will enable context extraction to enable smashHit to act not like a static system but provide dynamically (based on the current contexts and rules) adequate functionalities and options.

The smashHit platform shall provide a user-friendly interface to the smashHit framework.

Consent Certification

The consent certification module includes functionalities along the life cycle of consent certificates and is a core component of smashHit. The module includes functionalities for consent management and support.

Semantic models will be used to describe consent (-chains) between two or more participants. The described consent will be digitally certified and saved in an database.

This approach will enable to manage and control consent about data usage with two or more involved participants in the “data use traceability module”.

Data use traceability

The data use traceability component module has two purposes: (1st) to trace data flows by fingerprinting/watermarking, and (2nd) the identification of data leakages. The module addresses the need of OEMs, Data Providers and Data Processors to get a solution which enables the identification of data leakages or misuses. Basis for this component are data fingerprinting and watermark technologies that shall enable the identification of the last data source within the smashHit ecosystem. The fingerprint or watermark will provide answers about where the leaked or misused data was located or forwarded the last time which will avoid the blaming of guiltless participants in a consent chain.

Contract support tool (automatic contracting)

The smashHit solution shall accelerate the consent creation process to increase the efficiency in the data economy and to reduce the efforts needed for a consent creation and consent management. This component aims to allow a semi-automatic consent form creation process based on a law compliant data usage and processing rules. This approach will be based on ontologies and reasoning technologies (having the smashHit semantic model as basis).

Data owners need a solution which provides monitoring functionalities over their data. This will be enabled by the consent certification in the previously described module. Certified contracts and consent(-chains), saved in a knowledge graph enables to always have an overview about signed contracts and consents, and enables the tracing of consent, allowing the identification of a broken consent chain.

Security and privacy mechanisms/metrics

The smashHit Framework shall be developed and run in accordance with the General Data Protection Regulation (GDPR). Therefore, the principles of privacy by design and by default are cornerstones for the development of the smashHit infrastructure. This will be achieved by security & privacy policy languages and mechanisms, event-response language and mechanism within the smashHit infrastructure. This component focussed on how to combine privacy/security policies and their associated mechanisms in a complementary, if not integrated, way.

1.3 smashHit System architecture

The integrated modules and sub-components are building the smashHit system architecture shown in Figure 6.

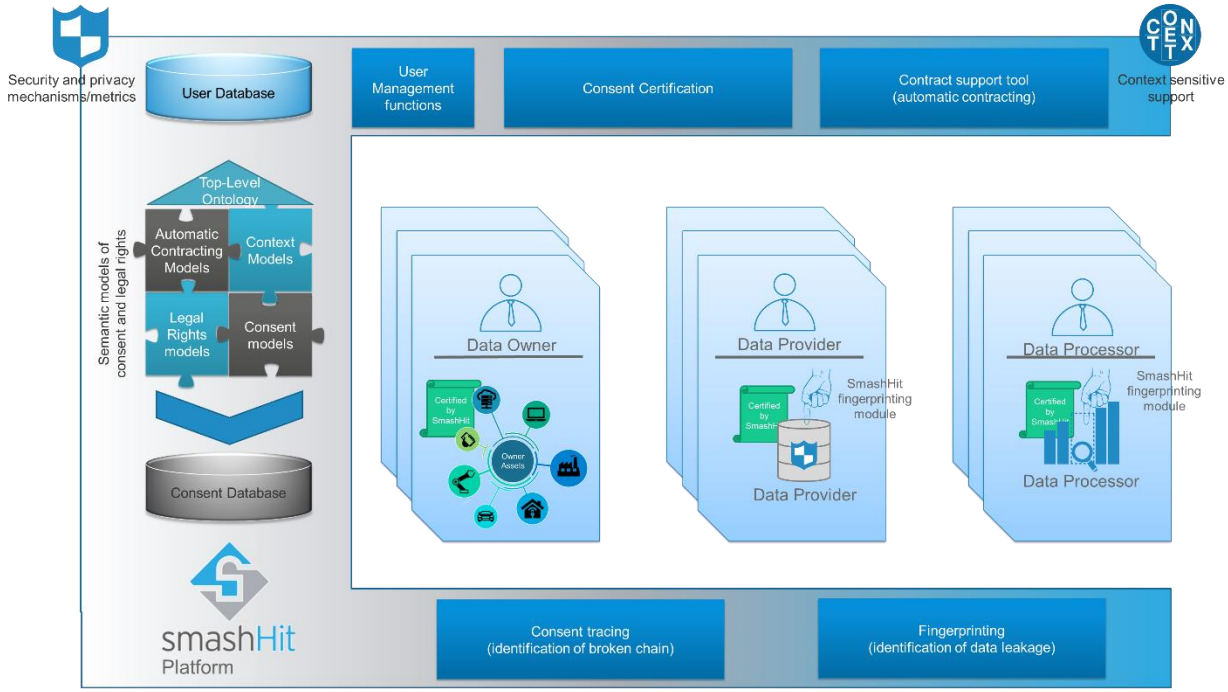

Figure 6: smashHit System Architecture[1]

In summary, the different components are providing the following functionalities:

smashHit Platform

- **User management functions –** This component includes the functionality regarding the life cycle of the actors involved in the smashHit framework. It interacts with the user database which stores information related to user accounts.
- **Semantic models of consent and legal rights** – This component represents the used ontologies and resulting semantic data models used for the description of legal rights, consents, context and automatic contracting rules, terms and conditions. The different ontologies are linked by a top-level ontology that describes very basic concepts to be extended by additional linked sub-ontologies.
- **Context sensitivity support tool** – This component adds context sensitive features to the overall smashHit framework and enables a context dependent behaviour of the system.

Consent Certification – This component includes functionalities regarding the life cycle of the consent certifications, i.e. the consent certification creation, consent management and consent distribution among the parties.

Data use traceability – This component will allow to find data leakages or misuses by using data fingerprinting or watermark techniques and provides data owners the power to manage their data contracts.

Contract support tool – This component provides automatic contraction functionalities that enables automatic consent document generation and execution, as well as semantic consent representation and visualisation. Furthermore, a consent tracing functionality will be provided, enabling the identification of broken consent chain and guaranteeing that the data exchanged are in agreement with the data consented by data owner.

[1] Context Symbol: https://commons.wikimedia.org/wiki/File:ConTeXt_Unofficial_Logo.svg

Security and privacy mechanisms/metrics – This component assured security and privacy within the smashHit Framework.

Users

Three main smashHit users are identified who represent different roles.

- **Data Owners** – Are the owners of data. They decide for which purposes their data can be used to whom they give their data.
- **Data Providers** – Are the data sources who collect data from data owner assets for a later provision.
- **Data Processors** – Are data actors who request and get data from data providers, to build and offer services based on these data.

1.4 smashHit Information Flows

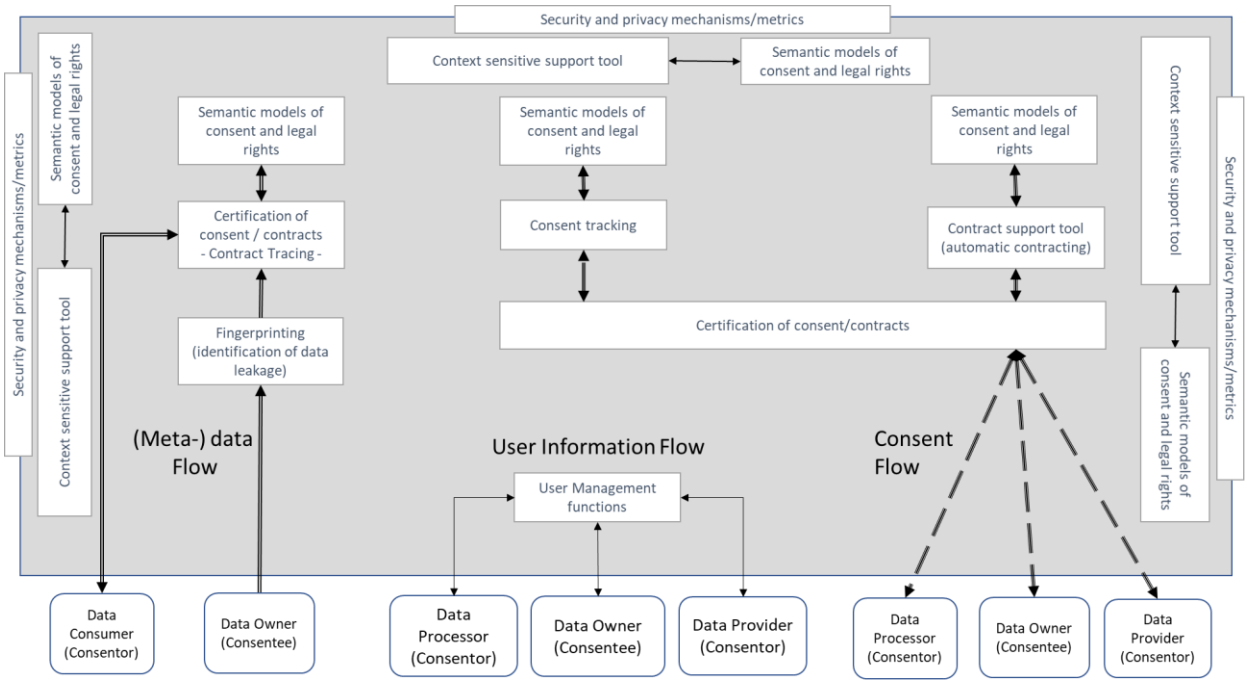

Figure 7: High-level smashHit architecture diagram depicting three major types of information flow

Figure 7 shows a high-level basic system architecture diagram which describes individual components and their interactions.

smashHit will handle three basic types of data flows:

1) user information,
2) consent (contract)
3) (Meta-) Data (CPP data streams combined with personal and industrial data streams).

As part of the **user information flow**, smashHit accounts that data producers, owners and consumers have created are managed by the “User Management functions” component. Here new users need to be vetted prior to them getting access to the full smashHit functionality.

Consent Flow in Figure 7 covers the

- management and storing of (compositions of) contracts between consentees and consentors
- generation of certificates to enable private and secure data sharing
- automated creation of contracts using semantic models
- tracking of consent to identify broken consent chains

(Meta-) Data Flow in Figure 7 allows to trace the (mis-) use of data by fingerprinting the data as it is accessed and processed by data consumers.

The three groups of users that access smashHit are shown at the bottom of the figure. Multiple copies of the same user groups are shown to avoid overcrowding the figure with too many intersecting arrows.

1.5 smashHit Scenarios

Figure 8: Main smashHit scenarios

As shown in Figure 8, smashHit offers five exemplary main high-level usage scenarios:

- **Initialisation and Setup** - deals with the initialisation and setup of the smashHit framework. It includes the registration of users and individual pre-configurations as e.g. data usage and provision rules set by a data owner.
- **Consent (Chain) Creation** – describes the main feature of smashHit. It is the creation of consent (chains) between two or more participants.
- **Consent and Contract Management** – describes the management of consents and contracts by the data owner.
- **Broken Consent Chain** – describes the handling of a brake in a consent chain caused by expiration, broken rules or external events.
- **Data Leakage and Data Misuse** – describes the handling in case of a data leakage or data misuse.

Each scenario will be described in the following sub-sections.

1.5.1 Initialisation & Setup

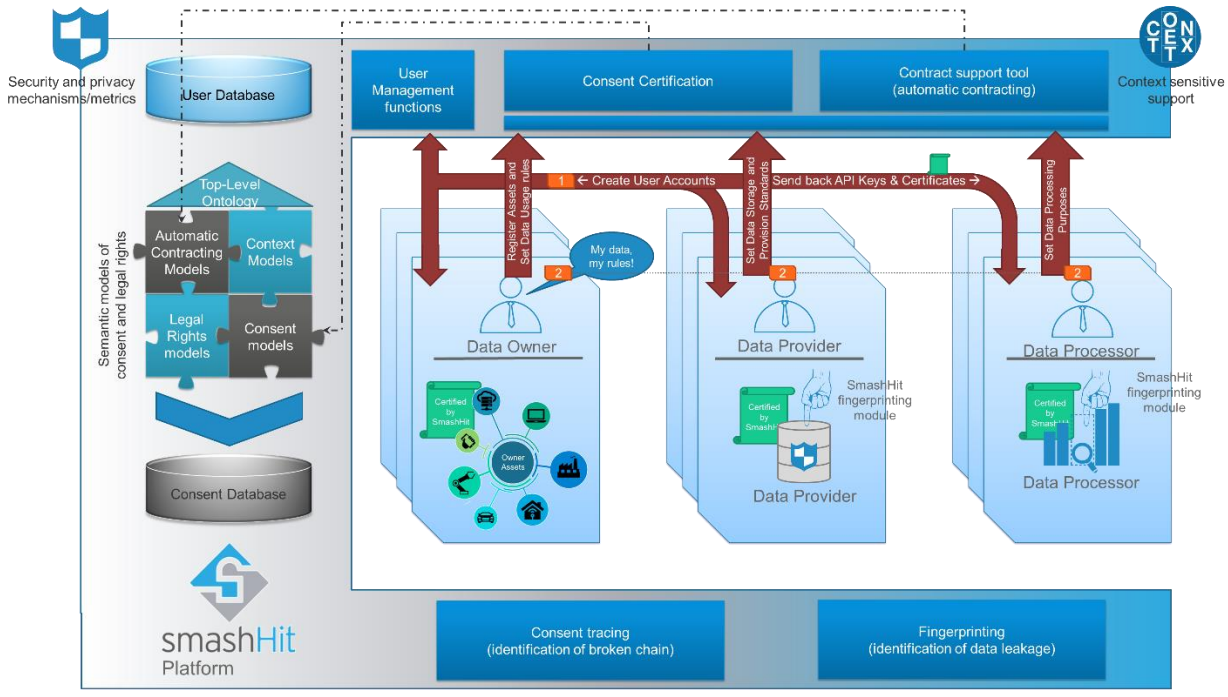

Figure 9: Scenario - Initialisation & Setup

The initialisation and setup scenario describes basic tasks that have to be done by users (data owners, data providers and data processors) to enable the regular use of the smashHit solution.

A smashHit account needs to be 'low-friction' and easily created, accessed and managed by users/consumers, who may be both Data Owner (driver), Data Provider (OEMs) and Data Processor (LexisNexis Risk Solutions, insurers). As part of the set-up, there needs to be a method of authentication, credentialing and verification of the entity that creates and owns the account, and against which consent contracts are created. Therefore, it is essential that the identity of the account holder is verified. The smashHit account then becomes the unique point of reference for the entity of the individual. This is a pre-requisite for the wider smashHit project, and not specific to this use case.

Step 1 in Figure 9 describes the creation of user accounts. Each user of the smashHit solution needs to register themselves. The users interact with the "User Management functions" component which provides related functionalities. After the registration, this component sends back an API Key and Certificates to the registered users which they later use for authentication.

Step 2 describes that different user roles can define pre-configurations using the "consent certification", "Contract support Tool" and "context sensitive support tool" component that help smashHit later to semi-automatise internal processes:

- Data Owners register their assets (like a car, a mobile phone, a weather station, etc.; but also virtual data producing services or only data buckets) that are used as sharable data sources. Further, the data owner defines rules for the use of their data. This enables smashHit to provide consent in a semi-automatic way (like automatically acknowledge the use of weather data from a registered weather station asset if data is used for non-military purposes).
- Data Providers can set their data storage and provision standards which allows data owners to know how data is shared and to set requirements (like provide data if data traceability features like watermarks or fingerprints are used).
- Data Processors can set data processing purposes e.g. to enable automatic contract creations.

1.5.2 Consent (Chain) Creation

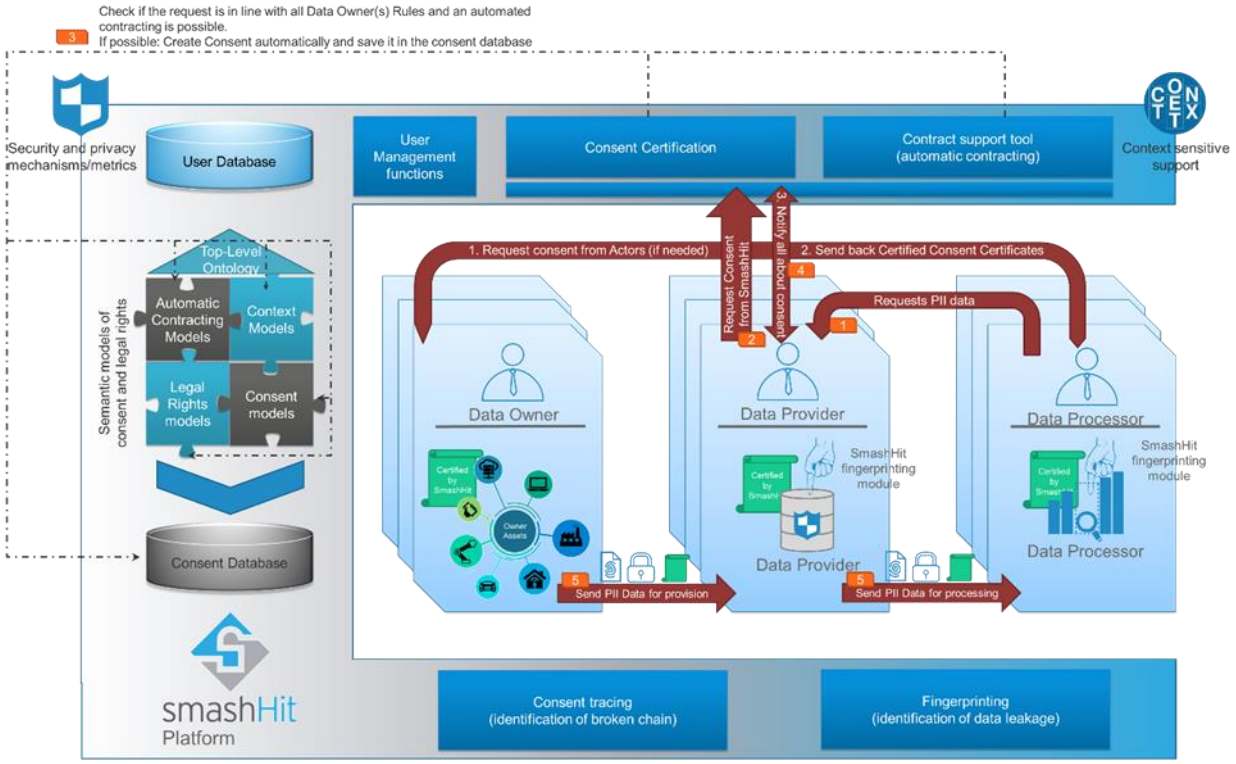

Figure 10: Scenario - Consent Creation

The consent (chain) creation is the main scenario of the smashHit scenario and addresses the core functionality.

In the scenario of Figure 10 a data processor needs PII from a data owner which requires a consent.

Step 1 shows the request of a data processor to get PII data from a data provider.

Step 2 shows that the data provider sends a request to the “Consent Certification” component to initiate the consent creation process.

Step 3 shows that smashHit is now checking if a consent was already granted or could be generated automatically based on set rules by all involved actors in the initialisation and setup phase. If possible, smashHit will create the consent automatically without requesting a manual agreement by the involved users.

Step 4 is divided into three sub-steps. Sub-step 1 addresses a request for agreement by the involved users in case that the consent could not be generated automatically in step 3. Sub-step 2 distributes the consent certificates to all involved users in case of a successful consent creation. In sub-step 3, all involved users will be informed about the result of the consent creation process.

Step 5 shows the final step. The PII data gets forwarded from the data owner to the data provider (if not done before), and the data provider forwards the PII data to the data processor. Optionally the data will be fingerprinted to enable data traceability if this was agreed in the consent creation process.

1.5.3 Consent and Contract Management

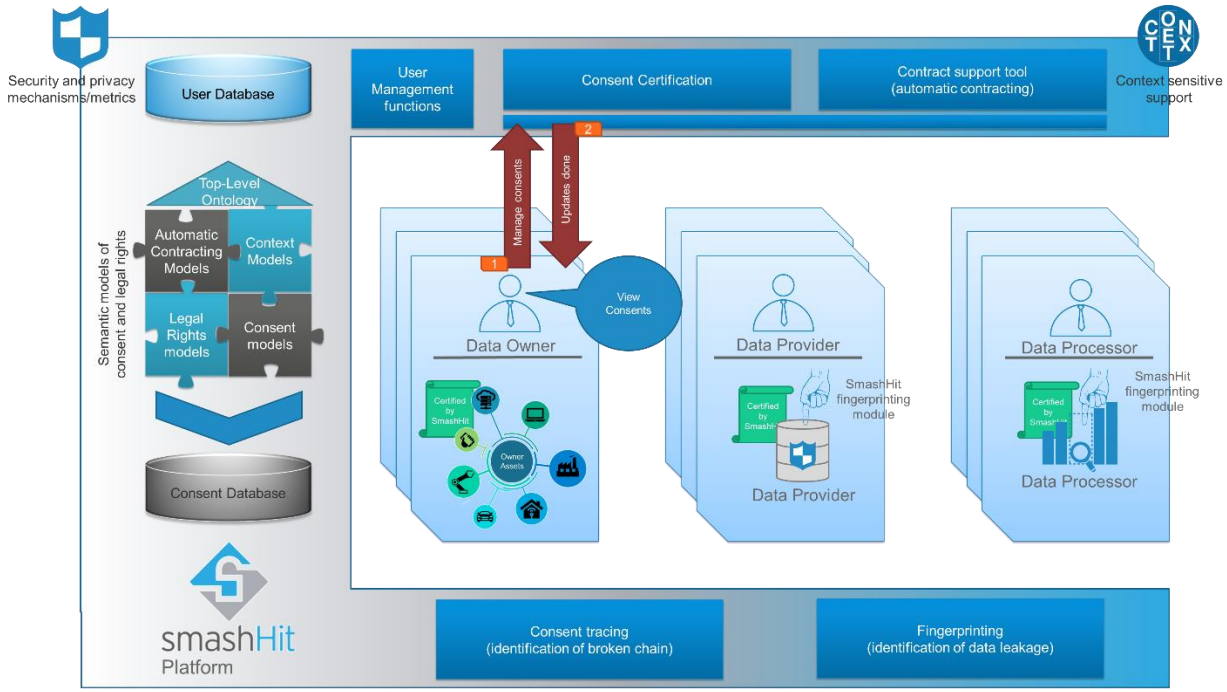

Figure 11: Scenario - Consent and Contract Management

smashHit needs to have a user-friendly interface for the data owner to access his consent listings. The listings should show the consents given, the metadata that forms the basis of the consent, to whom the Data Owner has granted the consent, any time frame against which the consent is valid, and for any explicit purpose. The Data Owner should be able to VIEW and ACCESS a link to the Data Consumer's or Data Producer's means of withdrawing consent. If consent is withdrawn (using the Data Producer's or Data Consumer's own process), the Data Consumer or Data Owner has an obligation to inform smashHit that the contract is no longer valid. In future iterations of the smashHit service the Data Owner may have the ability to withdraw or terminate a consent contract from within the smashHit portal. In future iterations of the smashHit service the Data Owner may also have the ability to restore any consent contracts that has been previously cancelled.

Figure 11 shows basically the interaction with the "Consent Certification" component (which collaborates with the "context sensitive support" component), which enables this functionality.

Step 1 shows that the data owner wants to manage own consents. This could mean to view own consents or, in future iterations, to withdraw, terminate or restore own consents.

Step 2 shows an optional notification about performed actions to the data owner.

1.5.4 Broken Consent Chain

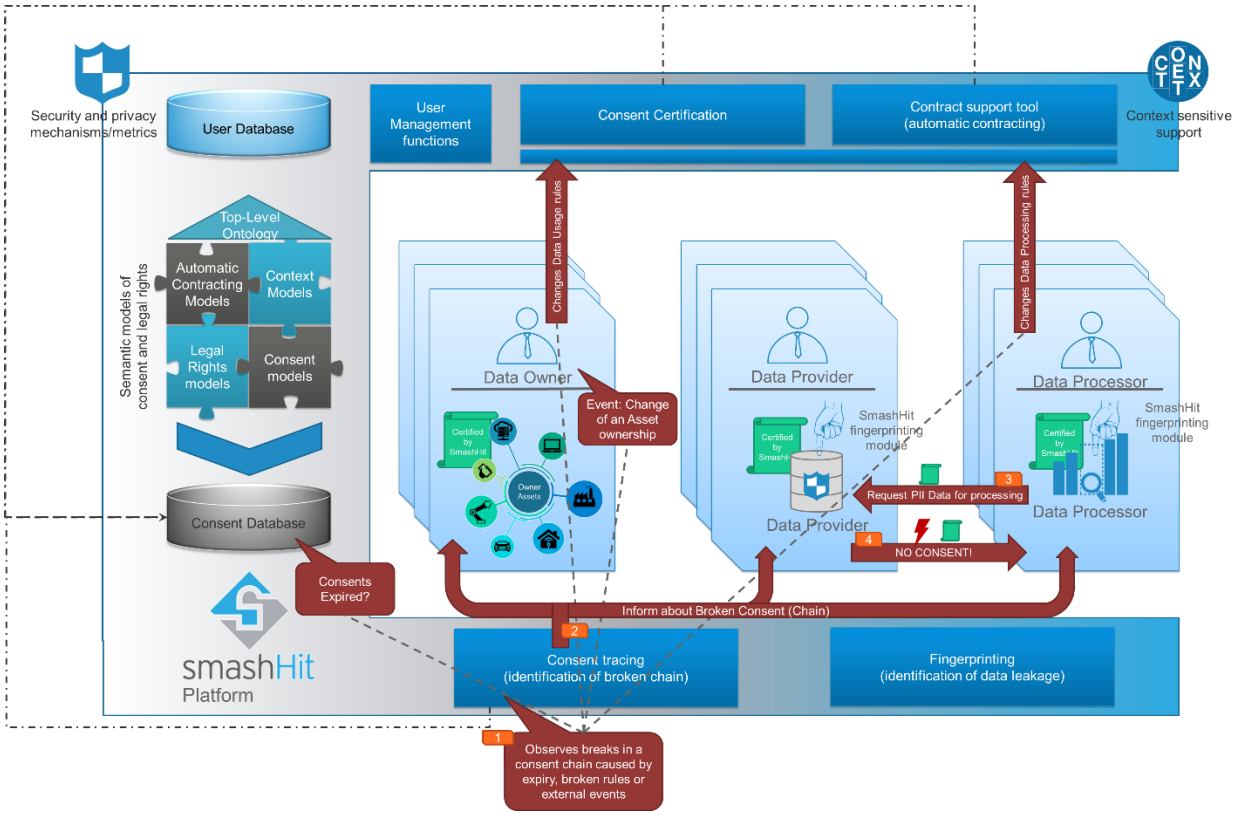

Figure 12: Scenario - Broken Consent Chain

A broken consent cain may occur for many reasons. If consent is withdrawn, the Data Consumer or Data Owner has an obligation to inform smashHit that the contract is no longer valid. The Data Owner may mark a contract as being invalid or under consideration but may not action this through the smashHit portal. Another reason could be that the data owner is changing the data usage rules or that a data processor changes the data processing purposes.

Such a situation is represented as a scenario in Figure 12.

Step 1 shows some reasons for a broken consent chain. Like an expired consent, a change of an asset ownership, the change of data usage rules by a data owner or the change of data processing rules by a data processor. The “consent tracing” component observed such changes in a consent (chain) and acts in case a break has been identified.

Step 2 shows that all actors involved in the consent chain get notified in case of a broken consent chain. So that all actors can behave according to the changed situation.

Step 3 and 4 show that the data processor is not anymore able to receive requested PII data because there is no consent given anymore.

1.5.5 Data Leakage and Data Misuse

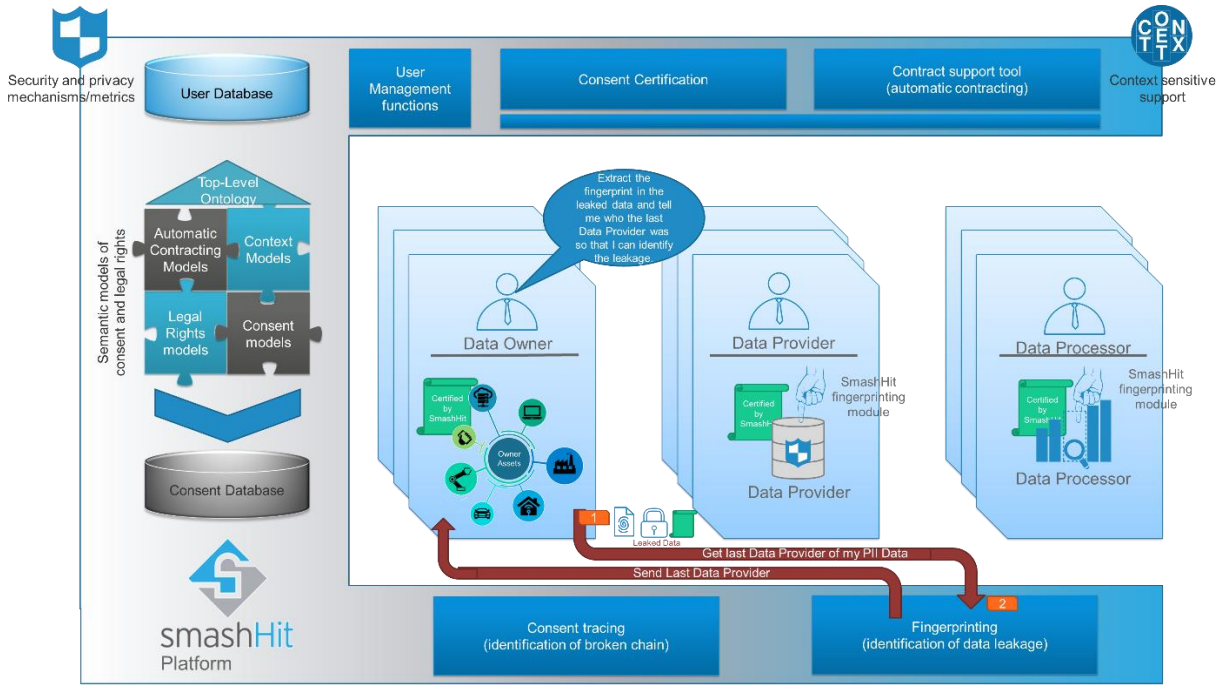

Figure 13: Scenario - Data Leakage and Misuse

The risks connected with missing data traceability related with ‘reconstruction of data leakage’ and minimisation of ‘Reputation Damage Risks’ are planned to be solved by smashHit fingerprinting solution. To lower the risks and increase the OEM acceptance, it is crucial to be able to pinpoint data breaches and data misuse to their origin in the data processing chain. To this end, a fingerprinting technology is desirable that connects data packages requested from the OEM APIs to their data customer’s user ID. This way, OEMs can quickly pinpoint the sources of breaches in outside third-party systems based solely of the raw data packages themselves.

Figure 13 shows a scenario in which data were leaked and are now publicly available in the internet. The data owner wants to identify the source of the leak.

Step 1 shows that the data owner sends the leaked PII data to the “Fingerprinting” component. The component extracts the fingerprints included in the PII data and traced back the last source where the data was leaked.

Step 2 shows that the “Fingerprinting” component informs the data owner about the last data provider who has fingerprinted the data and was probably the source for the data leakage.

2 smashHit Methodology Concept

Beside the specification and implementation of the smashHit platform solution (smashHit platform, Consent certification, Data use traceability etc.) the project has foreseen to develop and provide a set of support material, which serve the different stakeholders of the smashHit value chain (data owners, data providers, data consumers, data processors) as basic guidelines for participating in the overall smashHit workflow.

A key challenge of the Methodology is to address a privacy aware consent and contract management for secure sharing of data from and between diverse platforms including agreed consent and legal rules among stakeholders (data providers, data owners, data users). Therefore, the Methodology will provide guidelines how to use the smashHit platform and the privacy and security preserving services to integrate and assure data owner and service users consent and compliance with legal rules of all involved stakeholders over diverse platforms. The Methodology will provide the method to translate the legal rules of the stakeholders in the semantic model which will be used by smashHit solution to achieve a platform overlapping consent certification.

In this context, one of the key objectives of the support material is to empower smashHit platform stakeholders to

- easily implement specified and developed procedures and tools and
- to understand offered functions and tools.

The developed support materials will also address organisational, administrative and contractual measures concerning the interaction of the various stakeholders with the smashHit platform.

To achieve such envisaged Methodology Concept and to identify required tailored support material to be developed, the specific view and needs for each of the smashHit platform solution stakeholders was taken into account, which results into the following four types of support materials:

Developer Guidelines: This type represents specific developer guidelines for industrial smashHit platform users (data providers, data consumers & data processors), aiming to give these user groups all necessary knowledge on how to connect their systems with the heart of the smashHit platform (smashHit platform)

User Guidelines: This type represents specific user guidelines for all type of smashHit solution end users (data owners, data providers, data consumers, data processors, smashHit platform administrator), aiming to help users on how to use the platform (e.g. the specific functionalities provided by the platform, such as user authentication, consent/contract certification, data use traceability, automatic contracting etc.) and gives knowledge about the different functionalities available.

Concept Papers: These documents provide easy understandable general descriptions of the key innovative project outcome, like the overall smashHit platform solution.

smashHit Platform Policy Guidelines: As a complementary support material, this document provide covers legal, privacy and consent regulations for key processes/activities in respect to the actions/roles of the various stakeholders and their interaction required for the operation of the smashHit platform solution.

The final set of smashHit support material, will be presented in the scope of Deliverable 2.2 (due at end of June 2022) and will be made available for the public.

About smashHit

The objective of smashHit is to assure trusted and secure sharing of data streams from both personal and industrial platforms, needed to build sectorial and cross-sectorial services, by establishing a Framework for processing of data owner consent and legal rules and effective contracting, as well as joint security and privacy preserving mechanisms. The vision of smashHit is to overcome obstacles in the rapidly growing Data Economy which is characterized by heterogeneous technical designs and proprietary implementations, locking business opportunities due to the inconsistent consent and legal rules among different data-sharing platforms actors and operators. The Framework will provide methods and tools, such as the smashHit platform, to assure common consent over data shared using semantic models of consent and legal rules. The new tools include traceability of use of data, data fingerprinting and automatic contracting among the data owners, data providers, service providers and users. These tools are specifically critical for enormous volumes on data streaming from the usage of mass products with cyber physical features (e.g. vehicles). These data streams offer new opportunities to build innovative services, but their combination with other personal and industrial data is subject to complex ownership and consent aspects, as the data streaming from these products belong to persons or organizations who are owners or users of the products. The project will be based on the solutions developed or under development in previous and current projects (AutoMat, Cross-CPP, CAMPANEO, DALICC etc.). smashHit is driven by 2 industrial Business Cases involving several existing industrial and personal data platforms owned by the leading data providers in three diverse sectors (automotive industry, insurance, smart city), and will provide 3 demonstrators of various applications of the developed solutions. More information is available at www.smashHit.eu

Funded by the Horizon 2020 Framework Programme of the European Union