Grant Agreement No.: 761488

CPN

D1.5: Content Licensing and Distribution framework

Projectcpn.eu

The role of this deliverable is twofold. It will first introduce the Personal Data Receipts, an end-user tool integrated into the CPN platform with the aim to increase GDPR compliance. Role of this tool is to improve the transparency on what personal data from readers the platform collects, how it uses them and what options are available to readers to manage them. In order to construct a Personal Data Receipt for the CPN platform, the work presented in this deliverable will initially review the user data flow within the CPN platform, with particular emphasis on the data collected and used by the Recommendation system.

The second part of the deliverable will focus on how more collaboration across media publishers and between media publishers and content creators can be generated and new forms of revenue streams produced. The deliverable introduces the proposal of the Distribution Framework, a blockchain-based infrastructure useful to notarize agreements for content distribution and revenue sharing between content creators and a network of publishers, without requiring trust in any central authority. In order to do that we will first define a minimum viable licensing system, defining the terms of use for media contents. This deliverable will review existing licensing models (including Creative Commons) and identify the most suitable ones or their required adaptation in order to fulfil the CPN content distribution requirements.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Work package	1
Task	1.5
Due date	31/08/2018
Submission date	31/08/2018
Deliverable lead	DIGICAT
Version	1.0
Authors	Michele Nati (DIGICAT) Angelo Capossele (DIGICAT) Anthony Garcia (DIGICAT) Robert Learney (DIGICAT) Maria Prokopi (DIGICAT) Tilman Wagner (DW) Olga Kisselmann (DW) Elena Perotti (WI) Giovanni De Gregorio (WI) Zoë De Ruyck (VRT)
Reviewers	Tilman Wagner (DW) Olga Kisselmann (DW)
Keywords	Blockchain, trust, licensing, digital contract, personal data, transparency, compliance, GDPR

Document Revision History

Version	Date	Description of change	List of contributor(s)
V0.1	15/08/2018	1st version of the document for internal comments	Michele Nati (DIGICAT) Angelo Capossele (DIGICAT) Anthony Garcia (DIGICAT) Robert Learney (DIGICAT) Maria Prokopi (DIGICAT) Tilman Wagner (DW) Olga Kisselmann (DW) Elena Perotti (WI) Giovanni De Gregorio (WI) Zoë De Ruyck (VRT)
V1.0	31/08/2018	Final version of the document	Angelo Capossele (DIGICAT) Anthony Garcia (DIGICAT) Robert Learney (DIGICAT)

Co-funded by the Horizon 2020
Framework Programme of the European Union

DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761488.

This document reflects only the authors' views and the Commission is not responsible for any use that may be made of the information it contains.

Project co-funded by the European Commission in the H2020 Programme		
Nature of the deliverable:	**Report (R)**	
Dissemination Level		
PU	Public, fully open, e.g. web	**X**
CL	Classified, information as referred to in Commission Decision 2001/844/EC	
CO	Confidential to CPN project and Commission Services	

EXECUTIVE SUMMARY

One of the biggest challenge media companies have to tackle nowadays is to maximize match accuracy between available content and media consumers, so that millions of content items produced by a vast variety of news producers and publishers can find their way to a culturally diversified media consumer base. CPN aims to tackle this challenge by developing a new approach to personalisation of digital content, allowing both large and small media companies to benefit from the value of being able to better target content to media consumers. The project will build an innovative virtual open platform with pluggable building blocks allowing both large and small media companies to effectively personalise their content distribution.

This deliverable presents two of the building blocks required to realize this vision. On one hand, content personalization requires distributors to capture useful consumer data to feed matching algorithms developed within CPN, but at the same time, be compliant with privacy and data protection regulations (GDPR). On the other hand, content creators need tools to transparently and fairly increase content re-use and encourage wider sharing of their work.

In the following, we provide a high-level summary of the key points addressed in the deliverable. For more details, please consult the respective sections of the deliverable.

Analysis of personal data flow and legal assessment

An important starting point for the realization of a transparent media consumer profiling is the understanding of the regulations involved. An analysis of the General Data Protection Regulation (GDPR) looks at the CPN project as subject to the scope of the Regulation since the processing regards data of subjects within the EU. In this context, the data processed within CPN should be considered personal data, except for the case in which personal data are anonymized. The analysis highlights the following points:

- A lawful basis for data processing
- Retention of personal data
- Individual rights on data
- Transparency requirements on information notice and consent

Currently, the following data types are considered as the set of data of the collection process:

Co-funded by the Horizon 2020 Framework Programme of the European Union

- Studies/ profession
- Marital status
- Age-range
- Questions related to news consumption habits

Based on the GDPR analysis and on the types of data collected, the legal assessment identifies that the processing of data in the context of the CPN project seems not to raise specific concerns in terms of GDPR compliance, assuming that actions required by the regulation are fulfilled.

Personal Data Receipt

Fulfilling the identified legal requirements derived from the GDPR analysis requires a compliant solution providing transparency to an emerging knowledgeable consumer base. In this context, we present the concept of Personal Data Receipts (PDRs). PDRs are a human-readable digital record summarising in a simple and clear way what personal data an organisation is collecting about an individual, for what purpose, how it's stored, for how long, and if any third-party sharing is allowed. PDRs can be integrated with services using Consent Management Platforms and can be issued at the time new customers are on-boarded for given services. This section presents mock-ups and integration of the PDRs as well as its GDPR compliance in terms of:

- Right to be informed
- Right of access
- Right to rectification
- Right to erasure
- Right to restrict processing
- Right to data portability
- Right to object

Distribution framework

Fostering a more democratized digital content distribution requires tools to enable content producers to have direct control over how their content is distributed and monetized. Thus, the distribution framework is proposed to simplify the creation of licensing agreements for sharing content between content producers and the future network of publishers using the CPN platform and to create an audit trail of license and payment agreements between content producers and publishers. These agreements will be technically verifiable by auditing immutable records on a ledger shared by multiple parties. This deliverable presents an architectural overview of such a framework as well as a state-of-the-art examination of existing blockchain-based related projects.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Moreover, an analysis of the Creative Commons license framework is presented as a base for the development of a Minimum Viable Distribution Agreement for CPN. A list of use cases drives the identification of the following requirements:

- The set of parties (e.g., licensor and licensee) involved in the distribution agreement must be defined;
- Proof of ownership must be well defined, clear and easy to verify;
- Collaboration among different licensors must be defined in the distribution agreement by including the proportional ownership of each licensor;
- A distribution agreement must give the creator(s) the opportunity to specify the type of payment;
- A distribution agreement must clearly define what rights are granted to the licensee;
- A distribution agreement must clearly define what terms (obligations and/or restrictions) are requested of the licensee;
- A distribution agreement should include the legal text of the related license;
- A distribution agreement should include an easy-to-understand version of the license for non-technical people;
- A distribution agreement should include a machine-readable version of the related license;

As a result, the digital form of a distribution agreement is defined around the following sections:

- Parties involved;
- Ownership claim;
- License specification;
- Reward specification;

We finish with conclusions describing future developments which can be made to improve PDRs and the Distribution Framework for CPN.

Co-funded by the Horizon 2020 Framework Programme of the European Union

TABLE OF CONTENTS

Co-funded by the Horizon 2020
Framework Programme of the European Union

LIST OF FIGURES

Co-funded by the Horizon 2020
Framework Programme of the European Union

LIST OF TABLES

Co-funded by the Horizon 2020
Framework Programme of the European Union

ABBREVIATIONS

IP	Internet Protocol
TCP	Transmission Control Protocol
GDPR	General Data Protection Regulation
DLT	Distributed Ledger Technology
CPN	Content Personalisation Network
PDR	Personal Data Receipt
ICO	Initial Coin Offering
IPFS	InterPlanetary File System

Co-funded by the Horizon 2020
Framework Programme of the European Union

1 GDPR Summary

On 25 May 2018, the Regulation (EU) 679/2016 [1], also known as "General Data Protection Regulation" or "GDPR" has become applicable repealing the previous Directive 95/46/EC.

Scope of application

Articles 2-3 GDPR establish the scope of application. Having regard to the material scope, the Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system. While, the GDPR does not apply to the processing of personal data: (a) in the course of an activity which falls outside the scope of Union law; (b) by the Member States when carrying out activities which fall within the scope of Chapter 2 of Title V of the TEU; (c) by a natural person in the course of a purely personal or household activity; (d) by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security. Regarding the territorial scope, the GDPR applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not. Moreover, the GDPR applies to the processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to: (a)the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or (b)the monitoring of their behaviour as far as their behaviour takes place within the Union.

As a result, it is possible to consider that the CPN project is subject to the scope of the Regulation since the processing regards data of subjects within the EU.

Personal data

The GDPR defines personal data (art. 4 n. 1) as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person". In particular, Recitals 26 clarifies that, in order to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are

Co-funded by the Horizon 2020
Framework Programme of the European Union

reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.

The GDPR distinguishes between personal data and particular categories of data. In this last case, the processing is prohibited except in the case established by art. 9(2). Particular categories of data are personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual.

Regarding pseudonymization, personal data which have undergone this process, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. While the principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not, therefore, concern the processing of such anonymous information, including for a statistical or research purposes.

The data processed in the CPN project should be considered personal data, except in the case in which personal data are anonymized.

Data subjects, data controller and processor

The data subject is the person whose personal data are processed. The GDPR identifies specific role in the processing among which the most relevant ones are played by the data controller and the data processor. According to art. 4 n. 7 GDPR, the “controller” is the natural or legal person, public authority, agency or another body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law. While, according to art. 4 n. 8 GDPR, the “processor” is the natural or legal person, public authority, agency or another body which processes personal data on behalf of the controller.

According to article 24 GDPR, the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary. However, this obligation depends on the nature, scope, context and purposes of processing as well as the risks of

Co-funded by the Horizon 2020
Framework Programme of the European Union

varying likelihood and severity of the rights and freedoms of natural persons, the controller.

In the case of two or more controllers, art. 26 applies. Indeed, where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. In particular, joint controllers shall in a transparent manner determine their respective responsibilities for compliance with the GDPR obligations, in particular as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in articles 13 and 14 GDPR, by means of an arrangement between them unless, and in so far as, the respective responsibilities of the controllers are determined by Union or Member State law to which the controllers are subject. This arrangement shall duly reflect the respective roles and relationships of the joint controllers vis-à-vis the data subjects. The essence of the arrangement shall be made available to the data subject.

Regarding the processor, the GDPR provides more detailed rules. Preliminary, the GDPR clarifies that the controller shall use only processors providing sufficient guarantees to implement appropriate technical and organisational measures in such a manner that processing will meet the requirements of the GDPR. Moreover, the relationship between the controller and the processor shall be governed by a contract or other legal act under Union or Member State law, that is binding on the processor with regard to the controller and that sets out the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller. The specific content of such agreement is provided for by article 28(2) and the contract or the other legal act shall be in writing, including in electronic form.

Regarding sub-processing, the GPRD establishes that the processor shall not engage another processor without prior specific or general written authorisation of the controller. In case of written authorisation, where a processor engages another processor for carrying out specific processing activities on behalf of the controller, the same data protection obligations as set out in the contract or other legal act between the controller and the processor shall be imposed on that other processor by way of a contract or other legal act under Union or Member State law. Where that other processor fails to fulfil its data protection obligations, the initial processor shall remain fully liable to the controller for the performance of that other processor's obligations.

In the framework of the CPN project, the joint data controllers are DW, VRT, DigiCat, Livetech and the data processor/s is/are EN.

Lawful basis for data processing

Co-funded by the Horizon 2020
Framework Programme of the European Union

The lawful basis for processing are set out in Article 6 of the GDPR. At least one of these must apply whenever personal data are processed: the individual has given clear consent to the data controller in order to process their personal data for a specific purpose; the processing is necessary for a contract that the data controller has with the individual, or because they have asked the data controller to take specific steps before entering into a contract; the processing is necessary for you to comply with the law (not including contractual obligations); the processing is necessary to protect someone's life; the processing is necessary to perform a task in the public interest or for your official functions, and the task or function has a clear basis in law; the processing is necessary for the legitimate interests of the data controller or the legitimate interests of a third party unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.

Many of the lawful bases for processing depend on the processing being "necessary". This does not mean that processing always has to be essential. However, it must be a targeted and proportionate way of achieving the purpose. The lawful basis will not apply if the data controller can reasonably achieve the purpose by some other less intrusive means.

The data controller must determine its lawful basis before starting to process personal data in order to comply with the principle of accountability and transparency requirements. Indeed, the principle of accountability requires the data controller to be able to demonstrate that it is complying with the GDPR, so that the data controller should be able to explain its choice regarding the lawful basis which applies to each processing purpose.

If the data controllers' purposes change over time, the data controller may not need a new lawful basis as long as your new purpose is compatible with the original purpose. However, the GDPR clarifies that this mechanism does not apply to processing based on consent. Consent must always be specific and informed. The data controller needs to either get fresh consent which specifically covers the new purpose, or find a different basis for the new purpose.

It is necessary to include information about the lawful basis (or bases) in the privacy policy. Under the transparency provisions of the GDPR, the information to provide the data subject includes also: the intended purposes for processing the personal data and the lawful basis for the processing.

In the case of the CPN project, the following legal basis would apply:

(a) Consent: the individual has given clear consent to the data controller in order to process their personal data for the purposes of the CPN project.

(b) Contract: the processing is necessary for the contract which the user or the supplier agree in order to use the services offered by the CPN platform.

Co-funded by the Horizon 2020
Framework Programme of the European Union

(c) Legal obligation: the processing is necessary for data controllers to comply with the law e.g. communication to public authorities or judicial orders.

(d) Public task: the processing is necessary for the data controllers to perform a task in the public interest or for their official functions, and the task or function has a clear basis in law.

(e) Legitimate interests: the processing is necessary for the legitimate interests of the data controller or the legitimate interests of a third party unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.

Retention of personal data

The GDPR does not provide specific terms of retention of personal data. However, it is necessary to focus on the storage limitation principle as provided for by article 5(1)(e) GDPR. Indeed, personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with article 89(1) GDPR subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject.

As a result, the time of retention depends on different factors such as the usefulness of data for the processing. Erasing or anonymising personal data no longer needed reduces the risk that it becomes irrelevant, excessive, inaccurate or out of date. Personal data held for too long will, by definition, be unnecessary. It is necessary to have a lawful basis for retention.

Having regard to the CPN project, it will be necessary to address each retention period by issuing a retention policy and automated system of deletion or anonymisation when the purpose of the processing expires.

Summary of individual rights on data

The GDPR provides the following data subjects' rights:

a) The right to be informed: articles 13 and 14 GDPR specify what individuals have the right to be informed about. In particular, this obligation consists mainly in providing the privacy notice to the data subjects. The privacy notice should be concise, transparent, intelligible, easily accessible and uses clear and plain language. Where personal data are transferred to a third country or to an

Co-funded by the Horizon 2020
Framework Programme of the European Union

international organisation, the data subject shall have the right to be informed of the appropriate safeguards pursuant to Article 46 relating to the transfer. Moreover, the same information regards data which are transferred to third parties.

b) The right of access: article 15 GDPR provides that the data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and specific information listed by art. 15 GDPR (e.g. the purposes of the processing). The controller shall provide a copy of the personal data undergoing processing (potential reasonable fee for any further copies requested by the data subject). Where the data subject makes the request by electronic means, and unless otherwise requested by the data subject, the information shall be provided in a commonly used electronic form.

c) The right to rectification: article 16 GDPR states that the data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Considering the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement.

d) The right to erasure: article 17 GDPR introduces the right of erasure establishing that the data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay according to the cases described by article 17(1) GDPR. In those cases, listed by article 17(3) GDPR, the data subject cannot rely on its right of erasure.

e) The right to restrict processing: article 18 states that the data subject shall have the right to obtain from the controller restriction of processing in the cases listed by article 18(1) GDPR. 2. Where processing has been restricted, such personal data shall, with the exception of storage, only be processed with the data subject's consent or for the establishment, exercise or defence of legal claims or for the protection of the rights of another natural or legal person or for reasons of important public interest of the Union or of a Member State.

f) The right to data portability: article 20 clarifies that the data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller. Such data should be transmitted in a structured, commonly used and machine-readable format. Moreover, data subjects have the right to transmit his or her data (or have its data transmitted by the data controller) to another controller without hindrance from the controller to which the personal data have been provided, where the processing is based on consent or on a contract and the processing is carried out by automated means.

g) The right to object: article 21 GDPR established the right to object. In particular, the data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data

Co-funded by the Horizon 2020
Framework Programme of the European Union

concerning him or her which is based on point (e) or (f) of Article 6(1) GDPR, including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims. Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to the processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing. In this case, the personal data shall no longer be processed for such purposes. In the context of the use of information society services, and notwithstanding Directive 2002/58/EC, the data subject may exercise his or her right to object by automated means using technical specifications. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

h) Rights in relation to automated decision making and profiling: article 22 GDPR provides the right of the data subject not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. However, this right does not apply if the decision is necessary for entering into, or performance of, a contract between the data subject and a data controller; is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or is based on the data subject's explicit consent. It is particularly important that the controller ensure such right providing human intervention to express the controller point of view. In any case, automated decisions shall not be based on special categories of personal data referred article 9(1) GDPR, unless point (a) or (g) of article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

In the case of CPN, the above-mentioned rights should be ensured and explained in the privacy policy. In this case, the right to data portability and the right to object deserve specific attention due to the automated processing and profiling of data subjects.

Transparency requirements on information notice and consent

The privacy policy is at the basis of transparency. Indeed, the GDPR provides a list of information which the data controller should make available to data subjects in order to fulfil such obligation.

Co-funded by the Horizon 2020
Framework Programme of the European Union

In particular, where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information: a) the identity and the contact details of the controller and, where applicable, of the controller's representative; b) the contact details of the data protection officer, where applicable; c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing; d) where the processing is based on point (f) of Article 6(1), the legitimate interests pursued by the controller or by a third party; e) the recipients or categories of recipients of the personal data, if any; f) where applicable, the fact that the controller intends to transfer personal data to a third country or international organisation and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second subparagraph of Article 49(1), reference to the appropriate or suitable safeguards and the means by which to obtain a copy of them or where they have been made available.

Moreover, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing: a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period; b) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject or to object to processing as well as the right to data portability; c) where the processing is based on point (a) of Article 6(1) or point (a) of Article 9(2), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal; d) the right to lodge a complaint with a supervisory authority; e) whether the provision of personal data is a statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data; f) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

While, where personal data have not been obtained from the data subject, the controller shall provide the data subject with the following information: a) the identity and the contact details of the controller and, where applicable, of the controller's representative; b) the contact details of the data protection officer, where applicable; c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing; d) the categories of personal data concerned; e) the recipients or categories of recipients of the personal data, if any; f) where applicable, that the controller intends to transfer personal data to a recipient in a third country or international organisation and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second subparagraph of Article

Co-funded by the Horizon 2020
Framework Programme of the European Union

49(1), reference to the appropriate or suitable safeguards and the means to obtain a copy of them or where they have been made available.

Even in this case, the controller shall provide the data subject with the following information necessary to ensure fair and transparent processing in respect of the data subject: a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period; b) where the processing is based on point (f) of Article 6(1), the legitimate interests pursued by the controller or by a third party; c) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject and to object to processing as well as the right to data portability; d) where processing is based on point (a) of Article 6(1) or point (a) of Article 9(2), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal; e) the right to lodge a complaint with a supervisory authority; f) from which source the personal data originate, and if applicable, whether it came from publicly accessible sources; (g) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

In the case in which data have not been obtained from the data subject, the controller shall provide the information referred to in paragraphs 1 and 2: a) within a reasonable period after obtaining the personal data, but at the latest within one month, having regard to the specific circumstances in which the personal data are processed; b) if the personal data are to be used for communication with the data subject, at the latest at the time of the first communication to that data subject; or c) if a disclosure to another recipient is envisaged, at the latest when the personal data are first disclosed. Moreover, the transparency obligations do not apply in this case if the data subject already has the information; the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation referred to in paragraph 1 of this Article is likely to render impossible or seriously impair the achievement of the objectives of that processing. In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available; obtaining or disclosure is expressly laid down by Union or Member State law to which the controller is subject and which provides appropriate measures to protect the data subject's legitimate interests; or where the personal data must remain confidential subject to an obligation of professional secrecy regulated by Union or Member State law, including a statutory obligation of secrecy.

Consent is one of the lawful bases for processing and explicit consent can also legitimise use of special category data. Article 7 GDPR establishes the conditions for consent. First of all, according to the principle of accountability, where

Co-funded by the Horizon 2020
Framework Programme of the European Union

processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data. Moreover, in cases in which the data subject's consent is given in the context of a written declaration which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language. Consent should be obvious and require a positive action to opt in. Consent requests must be prominent, unbundled from other terms and conditions, concise and easy to understand, and user-friendly.

The data subject shall have the right to withdraw his or her consent at any time. The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. Prior to giving consent, the data subject shall be informed thereof. It shall be as easy to withdraw as to give consent.

Particular conditions apply to minors' consent in relation to information society services. The processing of the personal data of a child shall be lawful where the child is at least 16 years old. Where the child is below the age of 16 years, such processing shall be lawful only if and to the extent that consent is given or authorised by the holder of parental responsibility over the child. However, Member States may provide by law for a lower age for those purposes provided that such lower age is not below 13 years.

As a result, it is necessary that the CPN website publishes a privacy policy including all the information required by the GDPR in order to comply with transparency obligations. Moreover, regarding consent, in the case in which the consent is required since it is the only lawful condition of the processing, it will be necessary to ask to data subjects to express their consent according to the conditions established by the GDPR.

1.1 Personal Data and profiling

This section explores the relationship between personal data and their use in automated profiling in more detail, as this is relevant to CPN project.

It is clear now that most of the challenges outlined in the previous section regarding the lawful basis for data processing will characterise the emerging sector of news personalization. The main idea behind gathering data for news personalisation is to better understand the behaviour and preferences of one's readers, which could perhaps be achieved through a publisher's mobile application. The goal would be *to deliver to the user the news he/she likes to read, at the time, and in the form he/she prefers to consume it.* It is clear how achieving this will require a high level of reader profiling, which due to the population sizes involved and the multi-dimensionality of the underlying data, will be performed by automated algorithms. A number of unforeseen privacy and legal compliance issues might arise from this processing.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Although automated profiling might fulfil the requirements for such a matching and personalisation process to achieve efficiency, there is also a risk that such processes might result in significant effects on the data subjects, as they may unwittingly lock them even further within their ‘filter bubble’. Additionally, such preference-based profiling might also extract, correlate, and reveal (more or less accurately) sensitive data about the readers (e.g., sexual orientation, marital status, address, political inclinations), thus requiring organizations to apply additional effort to protect themselves against the risks of collecting this metadata, while being further exposed to the unwanted potential of discriminating between individuals. This is particularly true when such information is used and shared for targeted marketing, which is currently the predominant revenue stream for the industry. Finally, it is anticipated that these kinds of personalization services might also indiscriminately impact upon children, thus requiring us to carefully think about how to address all the related legal compliance challenges.

As result, to be on the safe side, all profiling data (such as activities and preference data) collected by CPN platform should require informed consent. However, this need could be waived during the pilot as long as it is made explicit that users joining the pilots will be exposed to a technology still under development and therefore their participation will require agreeing to full data sharing.

2 Personal data flow within CPN

This section analyses the personal data currently collected and used within the CPN platform, to establish whether the treatment is compliant with GDPR. The legal analysis will regard the following aspects:

- what data are collected;
- for what purpose;
- what are the legal basis on which data are collected;
- for how long will the personal data be held;
- what control is offered to end-users over their data (e.g., data portability right etc)
- if third party sharing is involved

2.1 CPN ecosystem

This section describes the roles of the three different actors involved into the CPN platform:

- Readers (the **data subjects**), are those to whom the personal data belongs
- CPN platform (the **data controller**) is the entity setting the purpose for which the data are collected and used (e.g. to provide a personalised news distribution service). This is represented by the union of VRT, DIAS and DW for the purpose of the pilots
- The service provisioning partners (the **data processors**) are the entities providing the service on behalf of the CPN platform (the data controller). This is represented by LiveTech for the purpose of the pilots

2.2 Use cases

Here we list some use cases describing the interactions of readers with the CPN platform in order to receive the personalised news service. For more information about use cases, please refer to the CPN Deliverable D1.2. We start from the registration phase.

Data collection process 1 - onboarding

We can illustrate this process with the following example: Michael currently lives in Belgium but was originally born and grew up in Germany. After years of hard work

Co-funded by the Horizon 2020 Framework Programme of the European Union

Michael is now planning to buy a vacation house in Cyprus. He's an eager reader of news from the three countries that form part of his life. He recently discovered that he could access news from VRT, DW and DIAS using the CPN platform. He decides to register.

During registration the platform asks him to provide data on his *home address*, *date of birth*, and *language preferences* in order to source relevant local news sources. These data are stored in the databases of CPN's data controllers for a maximum duration. If at any time Michael wants to amend his data, delete them, or move them, the CPN platform allows him to do so by providing a Personal Data Receipt outlining the current state of data which it holds on him, and a route to control this data.

Data collection process 2 - usages

After registration, Michael starts to use the CPN app to browse news. When he travels to Cyprus he starts to see news about children's TV from Germany. Because he doesn't read every news item related to Germany, and doesn't have any interest in children's TV, the system learns this and Michael starts to receive less news about these events. This is possible because the platform collects and uses *current location data*, and *age data* from Michael, based on the legal basis of contract under GDPR - the processing of Michael's personal data is necessary for the contract which the user or the supplier agree in order to use the services offered by the CPN platform.

Co-funded by the Horizon 2020
Framework Programme of the European Union

2.3 List of Data

Here is a table summarizing the data collected for the CPN Service:

Table 1: Collected data for the CPN Service

Type	Purpose	How long	Where	Legal basis (Service provisioning, legitimate interest, Consent, legal obligation)	What can the readers do with this data (remove, portability, update)	Are we sharing this data outside of CPN ecosystem? If so, with who and why?
User information: E-Mail IP-address **Locality:** Country Region Latitude and longitude **User platform info:** Browser, Operating system, Mobile or not **Web page info:** Domain, Host, Path, internal_referrer, Title IP address or URL (of website) **Consumption behaviour** Idle or not, Page read or not Engagement time, Scroll depth, Time spent on page, Interaction on page **Date and time:** Utc time step	Perform the CPN contract. User profiling and recommendations	Project duration after the opt-in and, in any case, until the opt-out	The CPN platform is hosted on an ENG server farm at Pont Saint Martin (Italy). The web app is hosted in AWS cloud at Frankfurt (eu-central-1 zone)	Contract and specific consent for profiling (service provisioning)	Opt-in/ opt-out, Access, erasure, portability, rectification, object	We share parts of this data anonymised, without any personalised data, with researchers, as stated in the DoW (through an open repository) The thus shared datasets will comply to the GDPR and to the FAIR principles.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Here is a table summarizing the collected data for the CPN evaluation:

Table 2: Collected data for the CPN evaluation

Type	Purpose	How long	Where	Legal basis (Service provisioning, legitimate interest, Consent, legal obligation)	What can the readers do with this data (remove, portability, update)	Are we sharing this data outside of CPN ecosystem? If so, with who and why?
User information interview: Studies/ profession Marital status Age-range Questions related to news consumption habits	User profiling	Before qualitative interviews	Belgium, Germany, Cypress	Contract and legitimate interest	Access, erasure, portability, rectification	only anonymised in project deliverables

2.4 Legal assessment

The processing of data in the context of the CPN project does not seem to raise specific concerns in terms of GDPR compliance. Personal data will be processed in order to enact the contract between the user and the platform, and where the data will be used for user profiling purposes, the processing will occur on a sound legal basis i.e. the consent of the user. Moreover, special categories of personal data, as per article 9 GDPR, will not be processed in the framework of the CPN platform. It is important to note that the processing of data will be based solely in the European Union. Regarding the retention of personal data, the GDPR does not establish a specific term after which data must be erased. As a result, in the case of CPN platform, personal data can be processed until the contractual relationship between the user and the platform is in force or until the data subject expresse**s** his/her intention to opt-out (i.e. withdrawal of the consent). In these cases, personal data should be removed immediately. In other cases (e.g. legal proceedings), the retention of personal data should be assessed according to the specific circumstances of the case at stake. As a general rule, the retention of personal data should respect the purposes according to which personal data have been processed. It is crucial that such processing flow will be well-explained both in the privacy policy and the cookie policy. Regarding dissemination of the result, it would be preferable to anonymise the gathered personal data. Pseudo-anonymisation techniques can be used, but the

Co-funded by the Horizon 2020 Framework Programme of the European Union

data controllers and processors would in this case be responsible for maintaining and ensuring that a specific individual would not be identifiable through those data.

Co-funded by the Horizon 2020
Framework Programme of the European Union

3 CPN Personal Data Receipt

This section describes the Personal Data Receipts adapted to the CPN project. It shows a customised version of Digital Catapult's PDRs [4] and how they are integrated within the CPN platform.

3.1 What is a PDR

Trends demonstrate that organisations are currently embracing digital transformation and creating more data-driven businesses through the use of customers' personal data. As this practice grows beyond the current predominant social media platforms and targeted advertising industry, more economic value is expected to be generated in new sectors, including digital manufacturing and digital health. According to the Internet Advertising Bureau (Europe), access to and use of personal data is worth approximately €100bn per year [2].

The UK government's Department for Digital, Culture, Media and Sport (DCMS) predicts a £241 billion growth in UK revenue between 2015-2020 derived from the use of personal data, with an 11% increase in customer numbers and a 10% growth in new opportunities [3].

With the ability to deliver advanced and personalised digital services for their customers, more companies are now also increasing their potential to generate additional revenue streams via the advertising industry.

As a result of the growing volume of sales and marketing communications that consumers are exposed to on a daily basis, more than half of consumers increasingly perceive a lack of control over how their personal data is being used, with 60% feeling nervous about sharing personal data when using digital services. This is due to the lack of transparency regarding how their data is collected and used, as well as the challenges consumers face in tracking and controlling how the data they share is actually used.

As a result, customers are demanding far simpler and clearer privacy statements that focus more on the user experience, in particular when services are accessed from mobile devices. Failing to provide such transparency reduces trust and increases the likelihood of customer churn.

Terms and Conditions (T&Cs) provide a cumbersome way to onboard savvy consumers to digital services. They now demand for more information. T&Cs are often agreed because they don't offer alternatives, or are seldom read in full or understood. Moreover, they often lack information on how choices can be altered or a service terminated, offerings now required under GDPR.

Co-funded by the Horizon 2020
Framework Programme of the European Union

We present what Digital Catapult believes to be a compliant solution that provides the transparency required by GDPR and an emerging knowledgeable consumer base: Personal Data Receipts (PDRs).

PDRs are a human-readable digital record summarising in a simple and clear way what personal data an organisation is collecting about an individual, for what purpose, how it's stored, for how long, and if any third-party sharing is allowed.

PDRs can be integrated with services using Consent Management Platforms. PDRs can be issued at the time new customers are on-boarded for given services and can also be tailored to include a link to the providers Consent Management Platform in order to amend permissions, the existence of which could otherwise be unknown or difficult for less technical customers to discover.

3.2 PDR MOCKUPS (DIGICAT)

This section provides a mock-up of the Personal Data Receipts developed for the CPN project. Beyond the scope of the pilot, it is expected that customer interest or activity data will constitute personal data collected for the scope of automated profiling. We can initially group all such data into two larger categories, governed by the same principles in terms of purpose and legal basis for collection, storage and retention period, data sharing and user rights. Future embodiments could implement a finer-grained level of detail to match an extended privacy policy and a user data management dashboard.

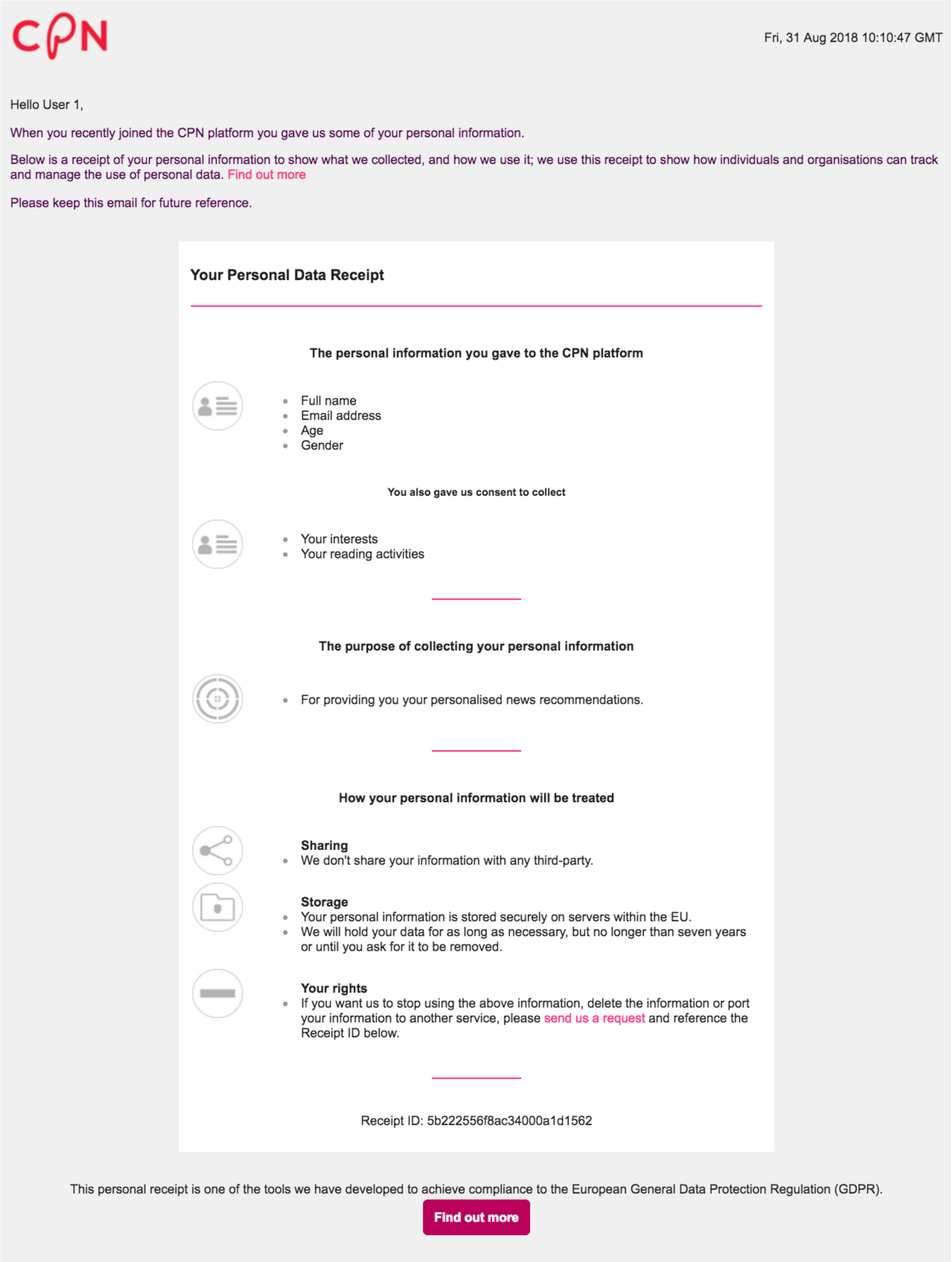
Personal Data Receipts were initially developed in English, but this can be tailored for the requirements of different regional publisher to adapt a preferred language.

We demonstrate the creation of 4 different versions of Personal Data Receipts, respectively for the following 4 categories of platform users.

Example 1:

- Giving consent to use both activity and interest data

Co-funded by the Horizon 2020
Framework Programme of the European Union



Fri, 31 Aug 2018 10:10:47 GMT

Hello User 1,

When you recently joined the CPN platform you gave us some of your personal information.

Below is a receipt of your personal information to show what we collected, and how we use it; we use this receipt to show how individuals and organisations can track and manage the use of personal data. Find out more

Please keep this email for future reference.

Your Personal Data Receipt

The personal information you gave to the CPN platform

- Full name
- Email address
- Age
- Gender

You also gave us consent to collect

- Your interests
- Your reading activities

The purpose of collecting your personal information

- For providing you your personalised news recommendations.

How your personal information will be treated

Sharing
- We don't share your information with any third-party.

Storage
- Your personal information is stored securely on servers within the EU.
- We will hold your data for as long as necessary, but no longer than seven years or until you ask for it to be removed.

Your rights
- If you want us to stop using the above information, delete the information or port your information to another service, please send us a request and reference the Receipt ID below.

Receipt ID: 5b222556f8ac34000a1d1562

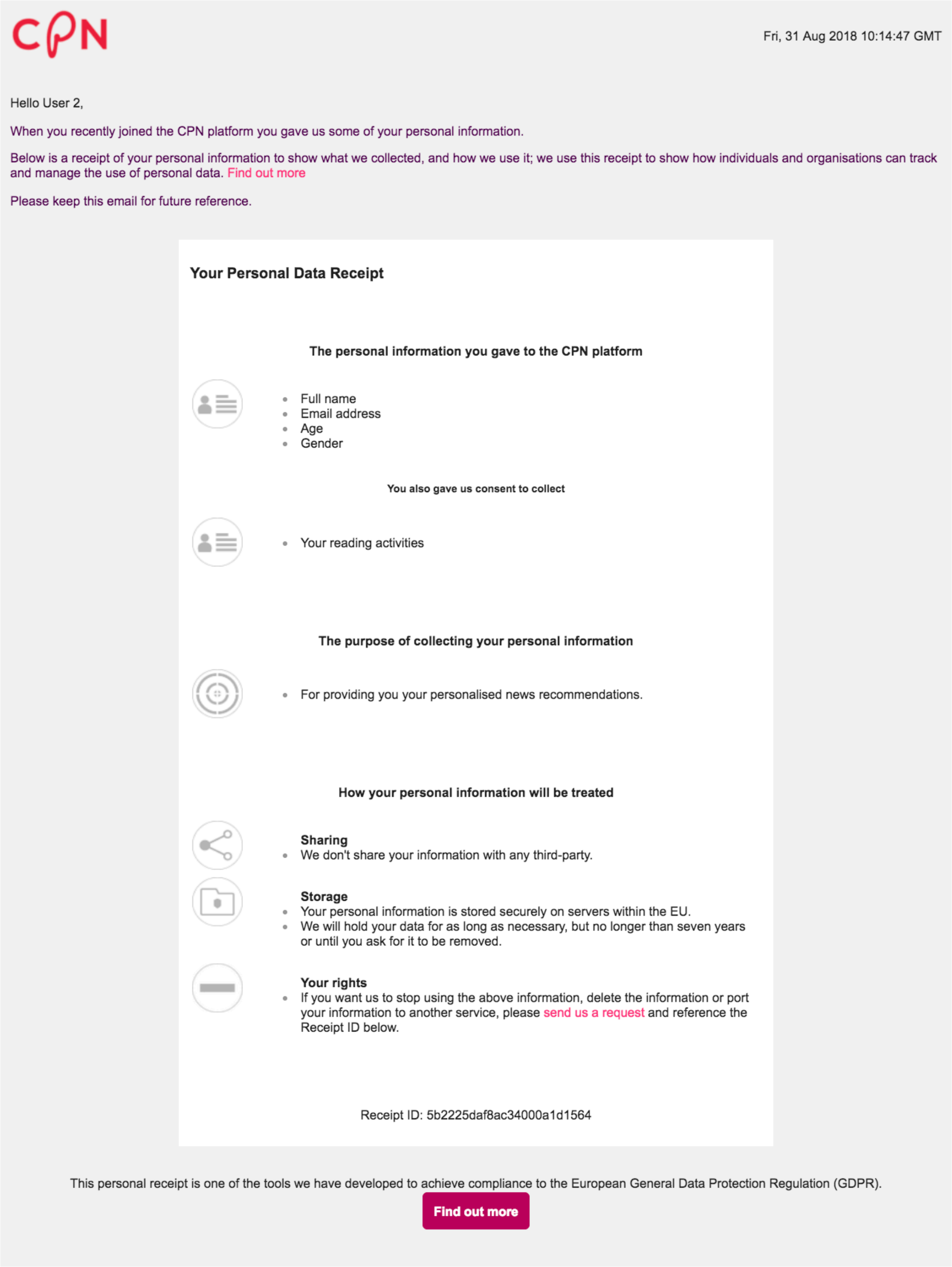
This personal receipt is one of the tools we have developed to achieve compliance to the European General Data Protection Regulation (GDPR).

Figure 1: PDR example 1

Example 2:

- Giving consent to use only activity data

Co-funded by the Horizon 2020
Framework Programme of the European Union



Fri, 31 Aug 2018 10:14:47 GMT

Hello User 2,

When you recently joined the CPN platform you gave us some of your personal information.

Below is a receipt of your personal information to show what we collected, and how we use it; we use this receipt to show how individuals and organisations can track and manage the use of personal data. Find out more

Please keep this email for future reference.

Your Personal Data Receipt

The personal information you gave to the CPN platform

- Full name
- Email address
- Age
- Gender

You also gave us consent to collect

- Your reading activities

The purpose of collecting your personal information

- For providing you your personalised news recommendations.

How your personal information will be treated

Sharing
- We don't share your information with any third-party.

Storage
- Your personal information is stored securely on servers within the EU.
- We will hold your data for as long as necessary, but no longer than seven years or until you ask for it to be removed.

Your rights
- If you want us to stop using the above information, delete the information or port your information to another service, please send us a request and reference the Receipt ID below.

Receipt ID: 5b2225daf8ac34000a1d1564

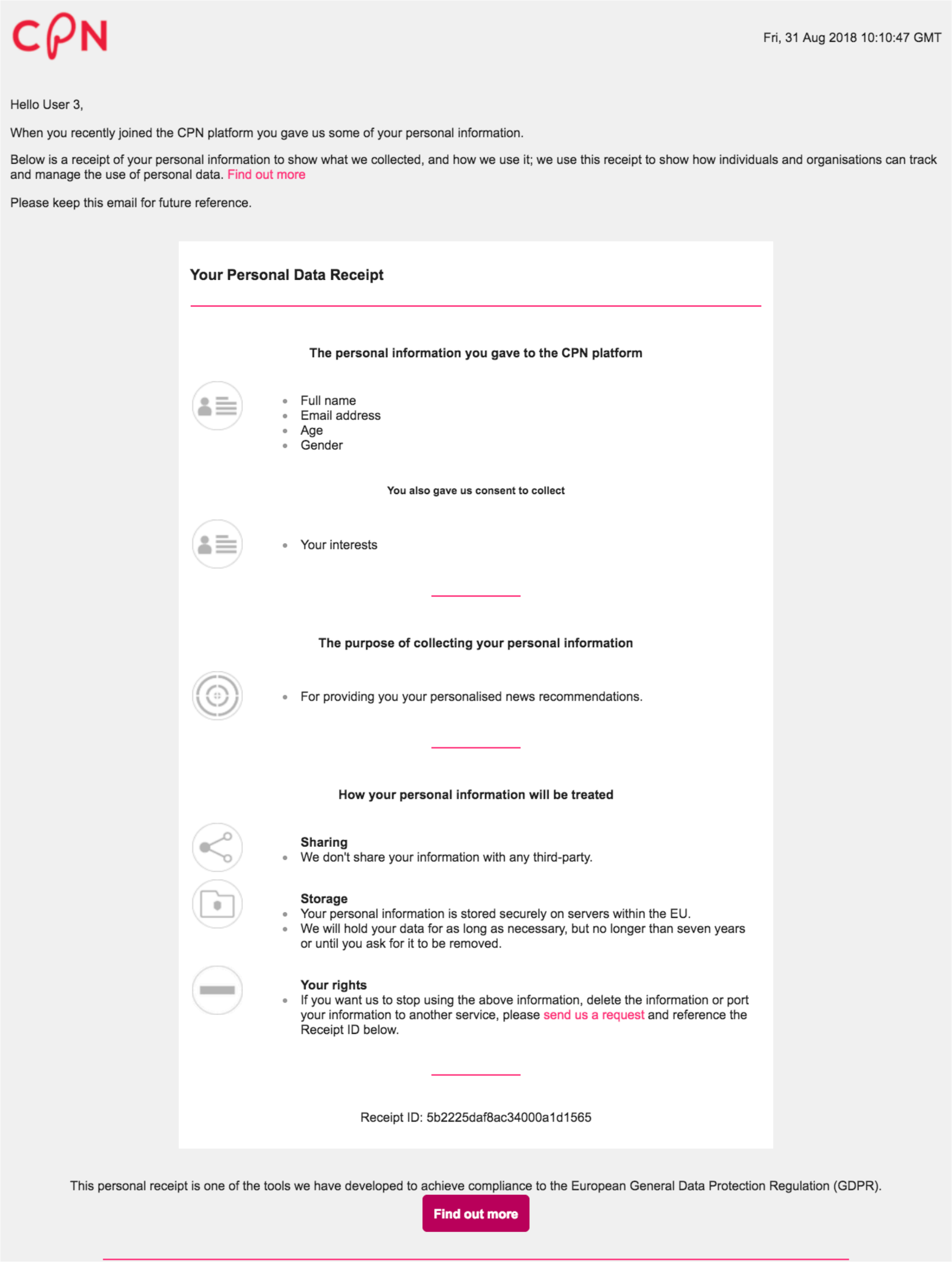
This personal receipt is one of the tools we have developed to achieve compliance to the European General Data Protection Regulation (GDPR).

Figure 2: PDR example 2

Example 3:

- Giving consent to use only interest data

Co-funded by the Horizon 2020
Framework Programme of the European Union



Fri, 31 Aug 2018 10:10:47 GMT

Hello User 3,

When you recently joined the CPN platform you gave us some of your personal information.

Below is a receipt of your personal information to show what we collected, and how we use it; we use this receipt to show how individuals and organisations can track and manage the use of personal data. Find out more

Please keep this email for future reference.

Your Personal Data Receipt

The personal information you gave to the CPN platform

- Full name
- Email address
- Age
- Gender

You also gave us consent to collect

- Your interests

The purpose of collecting your personal information

- For providing you your personalised news recommendations.

How your personal information will be treated

Sharing
- We don't share your information with any third-party.

Storage
- Your personal information is stored securely on servers within the EU.
- We will hold your data for as long as necessary, but no longer than seven years or until you ask for it to be removed.

Your rights
- If you want us to stop using the above information, delete the information or port your information to another service, please send us a request and reference the Receipt ID below.

Receipt ID: 5b2225daf8ac34000a1d1565

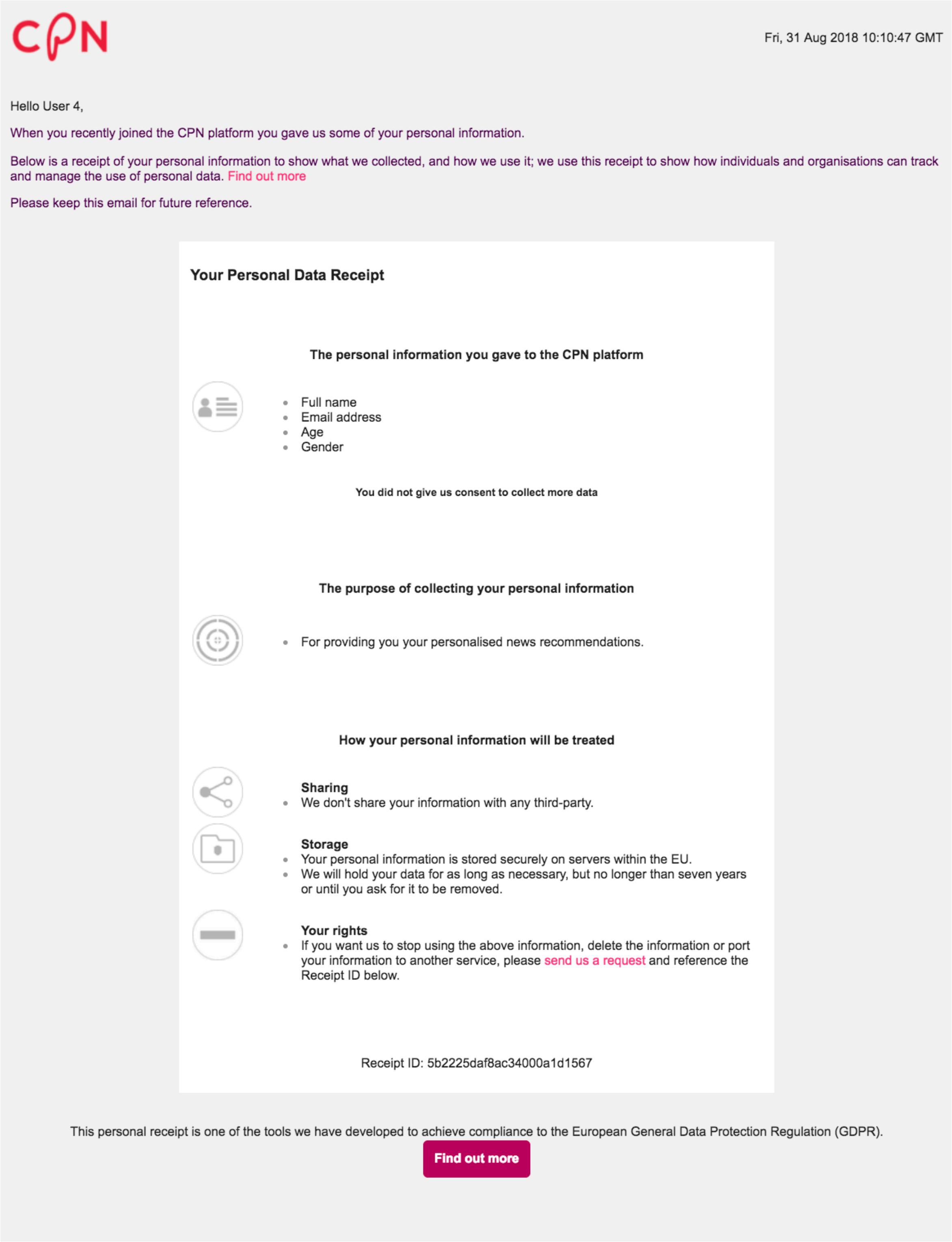
This personal receipt is one of the tools we have developed to achieve compliance to the European General Data Protection Regulation (GDPR).

Figure 3: PDR example 3

Example 4:

- Not giving consent to use any profiling data

Co-funded by the Horizon 2020 Framework Programme of the European Union



Fri, 31 Aug 2018 10:10:47 GMT

Hello User 4,

When you recently joined the CPN platform you gave us some of your personal information.

Below is a receipt of your personal information to show what we collected, and how we use it; we use this receipt to show how individuals and organisations can track and manage the use of personal data. Find out more

Please keep this email for future reference.

Your Personal Data Receipt

The personal information you gave to the CPN platform

- Full name
- Email address
- Age
- Gender

You did not give us consent to collect more data

The purpose of collecting your personal information

- For providing you your personalised news recommendations.

How your personal information will be treated

Sharing

- We don't share your information with any third-party.

Storage

- Your personal information is stored securely on servers within the EU.
- We will hold your data for as long as necessary, but no longer than seven years or until you ask for it to be removed.

Your rights

- If you want us to stop using the above information, delete the information or port your information to another service, please send us a request and reference the Receipt ID below.

Receipt ID: 5b2225daf8ac34000a1d1567

This personal receipt is one of the tools we have developed to achieve compliance to the European General Data Protection Regulation (GDPR).

Figure 4: PDR example 4

3.3 PDR INTEGRATION

This section describes how the initial implementation and integration of PDRs has been achieved in order to meet the requirements for the first pilot of CPN.

Co-funded by the Horizon 2020
Framework Programme of the European Union

The PDR module is a standalone application, released as a docker image, that depends on two external APIs: the CPN gateway and an email delivery platform (currently Mallgun) and executes the following actions, repeated at regular intervals:

1. Fetch the list of users from the CPN Gateway
2. For each new user (I.e. not part of the existing mailing list):
    - Generate a PDR from the appropriate template
    - Send an email to the user with the PDR attached
    - If the mail was sent, add the user to the mailing list (so that they don't receive an email more than once)

At a later stage, we expect to remove the dependency to the CPN Gateway and integrate the PDR module directly with the Kafka Streams API. This will allow us to send the PDR immediately after profile creation, but also to send a new PDR if the users decide to withdraw or alter their consent, as the PDR module will be triggered every time a user manages their profile.

This future embodiment will also remove the necessity to store personal information (the email address) in the email delivery platform.

The PDR component and the email delivery platform are designed as stateless services, without any personal data persisted after the PDR is sent.

3.3.1 Future work

The advantage of the Personal Data Receipts for consumers is that they provide a clear and actionable record of the terms under which they have provided personal data to an organisation.

A logical extension of the existing PDR would be a mechanism to allow all parties to trust the metadata and content of the receipts, to help ensure compliance, verify accountability and provide a tamper-proof public record in case of disputes.

This could be achieved by using a public ledger to store a suitably encrypted form of the receipts, potentially implemented using a distributed ledger platform such as a blockchain.

Further work could include services on the servers of *data controllers* (e.g. publishers) linked through from the PDR enabling the user to enact a choice to review, amend, or port their data in accordance with GDPR. Amending the permissions granted to the *data controller* by the *data subject* could involve a series of 'radio-button' like options accessed through a web page, or perhaps within the CPN application.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Also, on the server side, there could also be an event-driven stream provided by *data controllers* allowing users to subscribe to see when and how their personal data is being used. There may also be an option for users to store PDRs within the reader application, depending on implementation decisions.

3.42 GDPR COMPLIANCE

GDPR Article 4 states that consent should be freely given, unambiguous, as well as specific to the purpose, while Article 7 requires a proof of such consent to be maintained by both parties, the *data subject* and the *data controller*. Because PDRs are issued whenever a user joins a new digital service, this will typically fulfil the requirements for Article 7. However, the intention of PDRs is to increase transparency and eventually simplify management of individual digital rights, rather than just being a consent management tool for a data controller. The following table provides an explanation of how PDRs can achieve this.

Table 3: PDR compliance with GDPR

Article	What PDRs Offer
(12-14) Right to be informed	PDRs provide a standardised human-readable template designed by consumers for consumers. They allow data controllers to easily customise and deliver all the information required by the 'right to be informed'. Moreover, instead of showing this information on generic web-pages, they deliver it through personalised digital means (e.g. emails) and embedding data controller contact details. This opens up a direct channel between data subject and data controller. Using a unique ID simplifies the linkage of additional personal data to a specific data subject and opens the door to automatically update and notify the customers of the future acquisition of their additional personal data (e.g., after a Privacy Policy is revised).
(15) Right of access	PDRs provide a direct and convenient channel to automatically trigger an access request from a data subject.
(16) Right to rectification	Through the requirement to list third-party sharing, PDRs create good practice for linking data subject information to specific third parties receiving them, thus simplifying cascade updates.
(17) Right to erasure (18) Right to restrict processing (18) Right to data portability (21) Right to object	PDRs provide a direct channel that links data subjects and their personal data with data controllers, thus simplifying the notification and management of these rights. Potential future embodiments of the PDR would allow data subjects to interact digitally with data controllers in order to assert these rights through a web or app-based mechanism.

Co-funded by the Horizon 2020 Framework Programme of the European Union

4 DISTRIBUTION FRAMEWORK

The aim of the Distribution Framework is to democratize digital content distribution, enabling content producers to have direct control over how their content is distributed and monetized. The distribution framework simplifies the creation of licensing agreements for sharing content between content producers and the future network of publishers using the CPN platform, and to create an audit trail of agreements between content producers and publishers. These agreements will be technically verifiable by auditing immutable records.

This section describes the high-level architecture of the solution and related components.

4.1 ARCHITECTURE OVERVIEW

This section presents the *core abstractions* of the Distribution Framework architecture which provide the key concepts enabling interaction with the platform. The core abstractions are captured in the conceptual model of the Distribution Framework shown in Figure 5. At the core of the model is the notion of *agreement*. An agreement is related to a single item of *content*, a digital asset created by one or more *content producers*. An item of content has an *endpoint* used to access it. An endpoint could be centrally-held (e.g. on a cloud platform) or decentralised (e.g. IPFS). When an agreement is registered into the platform by a content creator, it becomes discoverable by *publishers*. The latter may purchase the access to the endpoint of the content that an agreement is related to, or obtain it for free if that refers to open content.

Several agreements may refer to the same content as they can provide different licensing models (*licenses*) and *price plans.* Licenses are separated into: (i) a *standard license* that can be chosen from a predefined set (e.g. the Creative Commons licenses), and (ii) a *personalized license*, according to which content creators can set additional restrictions to the access rights of their content, such as exclusivity, geography, duration, and purpose, along with the permissions regarding redistribution, adaptation, and resale.

Price plans are separated into the following options: (i) a *Pay per read* plan enables creators to charge publishers periodically according to the amount of times their users have actually "consumed" the content; (ii) a *one-time payment* plan only charges publishers at the time of sealing the agreement; (iii) finally, some agreements might be *free of charge*, which is mostly the case of agreements referring to open content.

Co-funded by the Horizon 2020
Framework Programme of the European Union

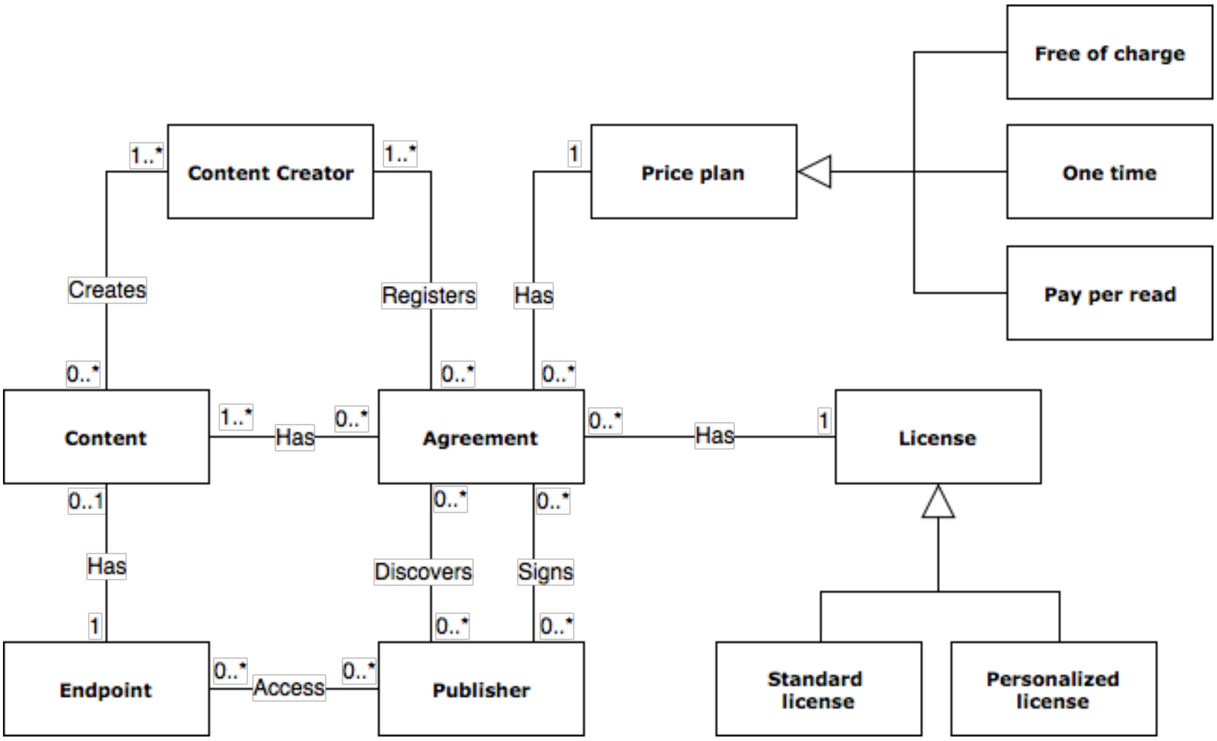


Figure 5: Conceptual model of the Distribution Framework

The high-level view of the Distribution Framework is shown in Figure 6 and consists of the following components:

- Distribution Framework API is the core element of the Distribution Framework and includes several sub-components that expose a set of functions to interact with that. Specifically, the Distribution Framework API allows content producers to register content agreements into the platform, and allows publishers to discover and purchase or license content.
- The CPN platform exposes functionalities providing a user interface through which content producers and publishers can interact with the platform, use the Distribution Framework API, and manage their accounts and information.
- Security includes three sub-components: (i) an *identity management system* to store and organize the identities of users; (ii) an *authentication component* that regulates the access to the platform; (iii) an *authorization component* that grants access to the functionalities and items of content according to predefined policies.

Co-funded by the Horizon 2020
Framework Programme of the European Union

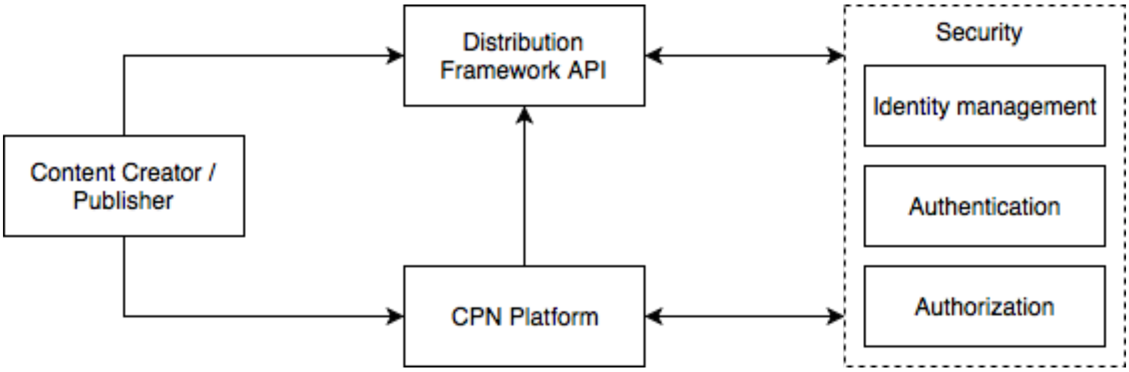


Figure 6: High-level view of the Distribution Framework

An overview of the *Distribution Framework API* is shown in Figure 7 and a description of its components is provided next.

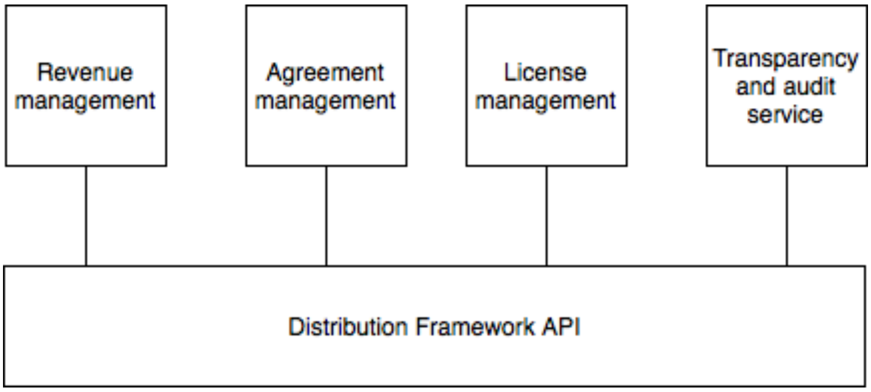


Figure 7: High-level view of the Distribution Framework API

License manager: this module allows content producers to set, define and customize different licenses for their content, thus enabling the creation of a dynamic ecosystem in which content creators can establish various business models. It provides an interface to retrieve pre-defined license templates so that content producers can link their content to a license instance selected from among the available templates. If the license templates do not fulfil the content producers' needs, this API allows for customisation in order for the license to better reflect the business model requirements. For instance, customisable templates allow to define: (i) business activity sectors for which the content may be used, (ii) purposes for purchasing and using the content, (iii) authorization to resell or sub-license the content, (iv) geographical territories in which the content may be used and, (v) the date after which the authorization period to use the content ends.

The main functionalities provided by this module are:

- License definition and customization: allows content producers to define different licenses templates based on standard licenses (e.g., Creative Commons) or based on custom models according to the specific business models chosen by the content producers.

Agreement manager: this module provides functionalities to register and search for different content agreements. Content agreements can be organized into groups/categories - in a hierarchical fashion when possible - to allow for an easy

Co-funded by the Horizon 2020 Framework Programme of the European Union

navigation and discovery of them. Attributes define characteristics and properties of content agreements. They may also be inherited from a higher level in a category hierarchy. The module allows content producers to define the description of the content they own as well as information related to the terms such as a price model, and license.

The main functionalities provided by this module are:

- Agreement specification: allows content producers to register a new agreement by detailing its description, pricing information, license terms.
- Agreement discovery: a list of available agreements can be retrieved and refined by specifying keywords and filters that match description, characteristics and properties of the desired content. As a result, publishers can easily discover what kind of agreements are available in the platform.

Revenue manager: this module allows content producers to generate revenue for their agreements by charging publishers for purchasing them. It exposes an interface to interact with external charging platforms (e.g., PayPal, Stripe). It collects all the information required for the charging process (price, publisher identifier, related parties etc.), which may differ according to the pricing model associated within the agreement.

The main functionalities provided by this module are:

- Charge management: provides the charging functionality for the system by interacting with one or multiple charging platforms (e.g., PayPal) and performing the required actions to charge the data consumers for purchasing data offerings provided by different data providers.
- Revenue sharing management: allows to define revenue sharing models to distribute revenues between the involved stakeholders (e.g., revenue shared among multiple content producers or between content producer and the platform provider as a transaction fee).
- Billing management: is in charge of sending invoices to publishers for their purchases. The invoicing process starts when a purchasing order is completed. In case of one-time payment model, a single invoice is sent to the publisher. Whereas, in case of pay per read model, invoicing can be done through time-triggered transactions.

Transparency and audit service: this module allows to register digitally signed agreements between content producers and publishers through immutable records. This enables the Distribution Framework to build and store a tamper-proof audit trail of all transactions. In case of potential disputes between parties, information stored within this module can be retrieved by auditors.
The main functionalities provided by this module are:

- Agreement registration: provides functionality to digitally sign an agreement and to store it on a distributed ledger.

Co-funded by the Horizon 2020
Framework Programme of the European Union

- Auditing: allows specific users (e.g., auditors) to retrieve signed agreements and to obtain their related details.

An overview of the Distribution Framework interaction among its internal components and other external ones is shown in Figure 8. Please note that only the bolded-line squares are part of the Distribution Framework, while dotted-line squares belong to external/common components (e.g., security, user management and external payment channels).

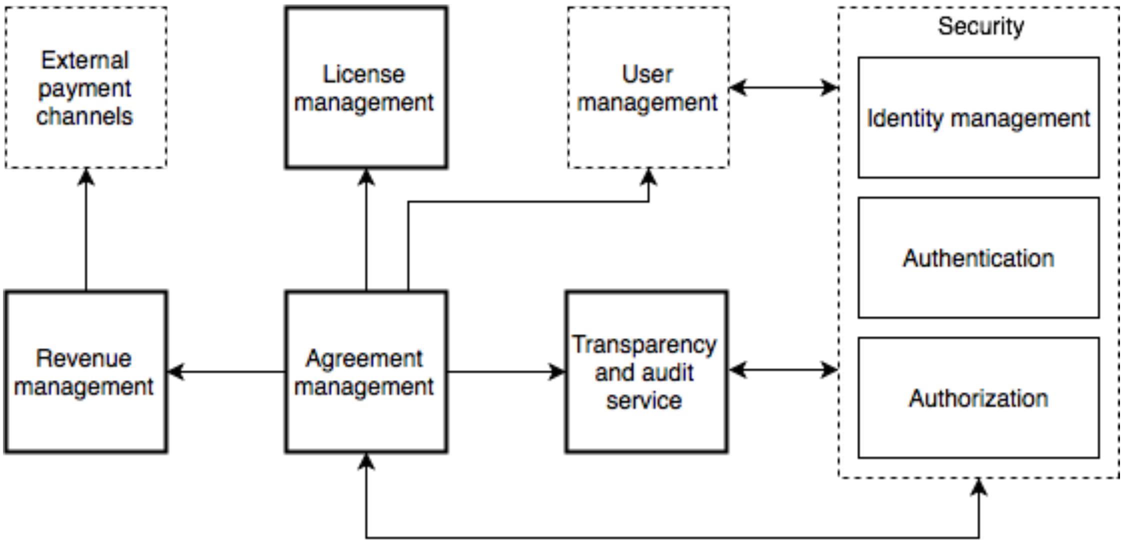


Figure 8: Interaction among components of the Distribution Framework

The *agreement manager* interacts with the *license manager* to bind a license to a new agreement. This component also interacts with the *revenue manager* component to enable monetization mechanisms, with the *transparency and audit service* to register and digitally sign agreements into a distributed ledger, and with the *security* components to update permissions of publishers when purchasing a content. It also interacts with the *customer manager* to retrieve publishers' billing information.

The *revenue manager* interacts with the *agreement manager* as well as with external payment channels to enable monetization mechanisms by exchanging transaction outcomes and charging information. The *customer manager* also interacts with the *identity manager* component to retrieve user's information.

Figure 9 summarises the overall content distribution process from the time of content creation by content producers, to the content acquisition by publishers.

In brief: One or more content producers (**CP**) creates content, say a photograph or an article. The CP optionally accesses a *license template library* to select a suitable license (e.g. Creative Commons) for their content, or they may create a new one based on their needs. A price model (e.g. free of charge, one-time payment, pay per read) is linked to the content along with the selected license. This information, together with additional metadata such as a description of the content and the

Co-funded by the Horizon 2020 Framework Programme of the European Union

intended licensee, create an agreement. The CPN platform stores the agreement in a distributed ledger through a digitally signed transaction. Publishers (**P**) can discover registered content and simply obtain access to the related content by agreeing to its terms and paying the CP if required. Again, a digitally signed transaction is stored into the distributed ledger, thus closing the loop for content distribution.

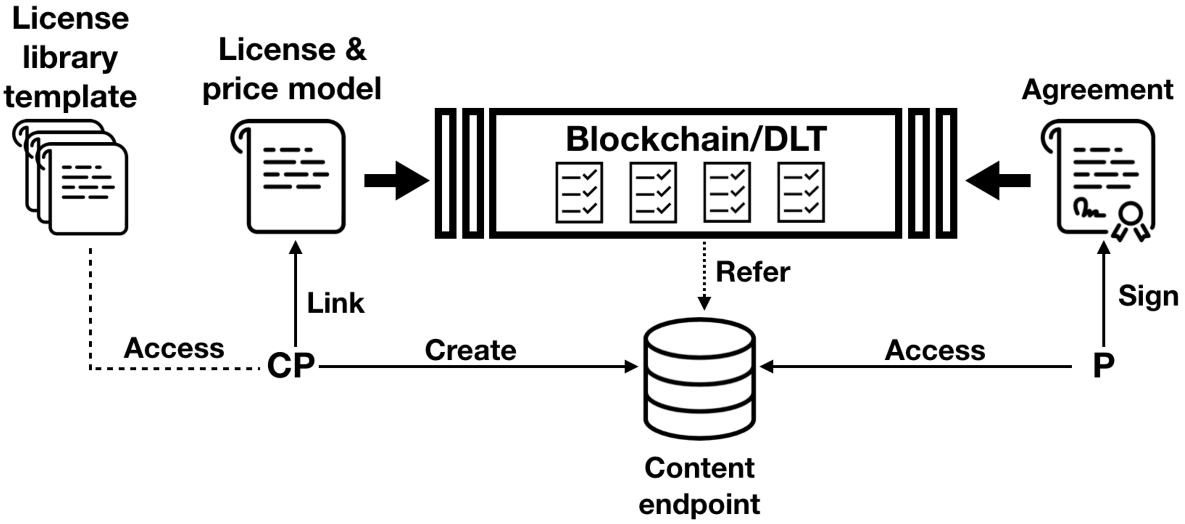


Figure 9: Content distribution process

4.2 SOTA ON EXISTING BLOCKCHAIN PROJECTS

This section provides a review of a range of existing blockchain projects trying to solve similar problems in the field of content licensing and distribution. Some of these projects are already successful and extremely well-funded.

- **Bloomen** - Bloomen [7] is a Horizon 2020-funded project (grant agreement No. 762091) which seeks to investigate how blockchain can be used to create a new media layer enabling users to copyright and monetise their own content. It aims to address topics such as: new content databases; proofs of ownership; payment methods; and quality metrics for media content. It is complementary to the work being undertaken in CPN, and also includes Deutsche Welle as a lead partner.

- **Flixxo** - Flixxo [8] is an ICO-funded project which promises to manage access to user created content streams using smart contracts on a blockchain. Users will pay for access using 'Flixx' tokens, which will unlock a bittorrent-based data stream to their device. The project is in early beta, and has raised US$2.5m in funding to date.

- **Unitalent -** Unitalent [9] aims to improve the connections between freelancers and corporate clients by using blockchain technology to manage contracts, payments, and worker reputation scores. The project has evolved

Co-funded by the Horizon 2020 Framework Programme of the European Union

from a traditional freelancer company launched in 2013, and will begin fundraising through an ICO in Q4 2018.

- **Steemit & d.tube** - Steemit [10] is a decentralised peer-to-peer blogging and social networking site resembling a mixture of Digg and Medium. It is also linked to a video streaming service called d.tube [11] which resembles an advert-free YouTube. Content on both systems is tracked and distributed using blockchain technology, as are the SMT (Smart Media Tokens) used by readers to reward content producers. In May 2018 it was the 1,788th most popular site on the internet according to Alexa, and the SMT economy was worth US$260m in August 2018.

- **Alexandria** - Aiming to be the 'Decentralized library of Alexandria', Alexandria [12] uses a combination of Bitcoin, IPFS, and their own cryptocurrency (the Florincoin) to distribute and track content and consumption, and to reward creators. It is still in alpha stage, and has the lofty goal of creating new protocols for content distribution, collaboration, and embedding. The project has raised US$7.6m through an ICO.

- **Brave (web browser) -** Brave [13] is a free, open-source web browser built on Chromium. It automatically blocks all adverts, but allows users to opt in to view targeted adverts in return for being paid in a cryptocurrency called the 'Basic Attention Token' or BAT. The intention is for content publishers to receive BAT in return for lost advertising revenue, and advertising companies to purchase BAT in order to place their adverts in front of viewers. The BAT ICO raised US$35m in May 2017.

Co-funded by the Horizon 2020
Framework Programme of the European Union

5 CPN DISTRIBUTION AGREEMENT

This section describes a minimum viable distribution agreement to fulfil the needs of the distribution framework (Section 4) for CPN. In the following we present a set of use cases first:

Use case 1 – single creator/single publisher

Michael is a freelance journalist who has created a new article and now wants to distribute it to publishers through CPN. Publisher P is interested in acquiring the exclusive rights to Michael's content, and is happy to pay Michael a one-time payment for the content. Michael uses the licensing tools within the distribution framework in order to attach a suitable exclusive license to his content, with an appropriate price, and Publisher P agrees with a few simple clicks.

Use case 2 – single creator/multiple publishers

Michael is a freelance journalist who wants to share his article through CPN using multiple publishers. Two publishers (P1 and P2) are interested in the article, but P1 wants to pay less in total than P2 because it expects a lower response among its readers base. Michael creates a license for his content which provides a payment-per-read price using the CPN platform. All three are satisfied with this arrangement.

Use case 3– multiple creators/single publisher

Michael and Maria are freelance journalists who want to share their article through CPN using publisher P. Michael provided image material while Maria wrote the article. They agreed to own 30% and 70% of the content respectively. Thus, they reward in proportion to their respective ownership. Publisher P wants a clear understanding of how the article can be shared and/or adapted to best fit hers/his needs. Using the CPN platform, Michael and Maria are able to attach a payment model which meets their partial contributions, and Publisher P can easily read the plain-language version of the license, as well a right to adapt or amend the content.

Use case 4 – multiple creators/multiple publishers

Michael and Maria are freelance journalists that want to share their article through CPN using multiple publishers (P1 and P2). Michael produced image material while Maria wrote the article. They agreed to own 30% and 70% of the content respectively. Thus, they expect in return rewards in proportion to their ownership. Publisher P1 likes the article and would like to send payments based on the number of times it is read, while Publisher P2 prefers to send a one-time payment with rights to adapt the article. Michael and Maria are able to use CPN to produce two versions of the content with different license agreement attached - one to match the requirements of P1, and another on a single payment basis as desired by P2. Both licenses include plain language explaining regionality, and rights to re-use or adapt.

Co-funded by the Horizon 2020
Framework Programme of the European Union

5.1 Distribution agreement requirements

This section extracts requirements the CPN distribution agreements should have according to the use cases above.

CPN distribution agreements must contain 4 sections defining the following elements:

- **Parties involved**: A distribution agreement must provide the set of licensor and licensee. If the work referred to a particular distribution agreement has been created by a collaboration among different creators, it must be clearly defined the proportional ownership among creators.
- **Ownership claim**: An agreement must provide a tamper-resistant, authenticated and verifiable proof of the assert claiming the ownership of the related work as well as its date of creation.
- **License specification**: a license should include rights and terms by clearly define what rights the licensor is granting to the licensee and under what terms these rights are valid. The license should include a legal text to have legal value. Moreover, it should include a human-readable and a machine-readable version, to easy the comprehension of its meaning by non-technical persons and to foster tracking and processing automation by electronic means respectively.
- **Reward specification**: reward types should be defined along with the license specification and support different models such as free-of-charge, one-type payment, or pay-per-read, thus, providing a flexible set able to fulfil the needs of the different stakeholder presented in the above use-cases.

Table 4 summarize requirements by listing for each requirement its identifier, title, description and type (functional/non-functional):

Table 4: Distribution agreement requirements

ID	Title	Description	Type
CPN-DF-DP	Defined parties	The set of parties (e.g., licensor and licensee) involved in the distribution agreement must be defined	Functional
CPN-DF-OP	Ownership proof	Proof of ownership must be well defined, clear and easy to verify	Functional
CPN-DF-PO	Proportional ownership	Collaboration among different licensors must be defined in the distribution agreement by including the proportional ownership of each licensor	Functional
CPN-DF-RT	Reward type	A distribution agreement must give the creator/s the opportunity to specify the rewarding modality	Functional
CPN-DF-LR	License rights	A distribution agreement must clearly define what rights are granted to the licensee	Functional
CPN-DF-LT	License terms	A distribution agreement must clearly define what	Functional

Co-funded by the Horizon 2020 Framework Programme of the European Union

		terms (obligations and/or restrictions) are requested to the licensee	
CPN-DF-LF	Legal form	A distribution agreement should include a legal text of the related license	Functional
CPN-DF-EU	Easy to understand	A distribution agreement should include an easy-to-understand version of the license for non-technical people	Functional
CPN-DF-MR	Machine readable	A distribution agreement should include a machine-readable version of the related license	Functional

5.2 SOTA on Licenses

This section provides a review of existing licensing model for content distribution. More specifically, it presents the Creative Commons licensing framework.

Creative Commons (CC) is a global non-profit organisation that provides free tools, including Creative Commons licenses and software, to enable authors, researchers, artists and educators to easily mark their creative works with the specific intellectual property rights they wish their creative works to carry. CC encourages legal sharing, remixing, and reuse of creative work, and provides a legal platform to spread and build digitally enabled creative culture by incorporating distributive and legal mechanisms at the same time. The mission of CC is to build a system of balanced intellectual property rights by advocating a 'some rights reserved' alternative to the traditional 'all rights reserved' system [6].CC Licenses are built around the following terms:

- **Attribution (BY)**: all CC licenses require that others who use owner's work in any way must give owner credit the way owner requests, but not in a way that suggests owner endorses them or their use. If they want to use owner's work without giving credit or for endorsement purposes, they must get owner's permission first.
- **ShareAlike (SA)**: owner lets others copy, distribute, display, perform, and modify owner's work, as long as they distribute any modified work on the same terms. If they want to distribute modified works under other terms, they must get owner's permission first.
- **NonCommercial (NC)**: owner lets others copy, distribute, display, perform, and (unless owner have chosen NoDerivatives) modify and use owner's work for any purpose other than commercially unless they get owner permission first.

- **NoDerivatives (ND)**: Owner lets others copy, distribute, display and perform only original copies of her/his work. If they want to modify owner's work, they must get owner permission first.

As also described in detail on the CC web site [6], by combining these terms, the following license types are available:

- **CC0 1.0 Universal (CC0 1.0):** the person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.

- **Attribution 3.0 Unported (CC BY 3.0):**
  - **Share** — copy and redistribute the material in any medium or format
  - **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

- **Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0)**
  - **Share** — copy and redistribute the material in any medium or format
  - **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

- **Attribution-NoDerivs 3.0 Unported (CC BY-ND 3.0)**
  - **Share** — copy and redistribute the material in any medium or format for any purpose, even commercially.
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any

reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NoDerivatives** — If you remix, transform, or build upon the material, you may not distribute the modified material.

- **Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0)**
  - **Share** — copy and redistribute the material in any medium or format
  - **Adapt** — remix, transform, and build upon the material
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NonCommercial** — You may not use the material for commercial purposes.

- **Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)**
  - **Share** — copy and redistribute the material in any medium or format
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NonCommercial** — You may not use the material for commercial purposes.
  - **NoDerivatives** — If you remix, transform, or build upon the material, you may not distribute the modified material.

- **Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0)**
  - **Share** — copy and redistribute the material in any medium or format.
  - **Adapt** — remix, transform, and build upon the material
  - **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NonCommercial** — You may not use the material for commercial purposes.
  - **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

Co-funded by the Horizon 2020 Framework Programme of the European Union

The following table summarise CC license types:

Table 5: CC license types overview

License Type	Free to:		Under the following terms:				No Copyright
	share	adapt	Attribution	ShareAlike	Non Commercial	No Derivatives	
CC0	X	X					X
CC BY	X	X	X				
CC BY-SA	X	X	X	X			
CC BY-ND	X		X			X	
CC BY-NC	X	X	X		X		
CC-BY-NC-ND	X		X		X	X	
CC BY-NC-SA	X	X	X	X	X		

5.3 CPN Minimum Viable distribution agreement

This section describes a minimum viable distribution agreement valid for the CPN and presents its digital form.

The following JSON structure presents the data model specification of a distribution agreement:

```
{
   licensor: [{
      party: object,
      proportionalOwnership: number
   }],
   licensee: [
      party: object
   ],
   creationDate: date,
   contentDigest: {
      value: string,
      metadata: string
   },
   license: {
```

Co-funded by the Horizon 2020
Framework Programme of the European Union

```
        id: string,
        title: string,
        href: string,
        legalText: string,
        commonsDeed : string,
        ccRel: string,
    },
    pricingModel: {
        type: string,
        description: string,
        value: number,
        currency: string,
    }
}
```

Licensor is defined as an array of *party* object along with the proportional ownership as described by requirement CPN-DF-PO. A *party* object is a JSON object defining the account details of a legal entity (e.g., individual, organization). An example of such an object defining an individual ***might*** include:

- **id** - Id of the party. Typically, corresponds with the username of the user in the system
- **href** - URL pointing to the party info. Each party info contains the following fields:
    - **gender** - Gender of the individual owner of the account
    - **placeOfBirth** - Place where the owner of the account was born
    - **countryOfBirth** - Country where the owner of the account was born
    - **nationality** - Nationality of the owner of the account
    - **maritalStatus** - Marital status (married, divorced, widow, etc)
    - **birthDate** - Date when the owner of the account was born
    - **title** - Preferred title of the user (Mr., Dr., etc)
    - **givenName** - First name of the user owner of the account
    - **familyName** - Family name of the user owner of the account
    - **contactMedium** - List of mediums that can be used to contact the user. Each medium contains the following fields:
        - **type** - Type of the contact medium. It could be *Email*, *TelephoneNumber*, or *PostalAddress*
        - **preferred** - If true, indicates that is the preferred contact medium
        - **emailAddress** - Full email address in standard format. This field is only used when the type is *Email*

- **number** - Phone number. This field is only used when the type is *TelephoneNumber*
- **street1** - Describes the street. This field is only used when the type is *PostalAddress*
- **street2** - Complementary street description. This field is only used when the type is *PostalAddress*
- **city** - City of the medium. This field is only used when the type is *PostalAddress*
- **postCode** - PostCode of the medium. This field is only used when the type is *PostalAddress*
- **stateOrProvince** - State or province of the medium. This field is only used when the type is *PostalAddress*
- **country** - Country of the medium. This field is only used when the type is *PostalAddress*

Licensee is defined as an array of party objects. Together with the *Licensee* field, they fulfill the CPN-DF-DP requirement.

Claim of the ownership is defined by the *creationDate* and *contentDigest* fields. The first is automatically set on the agreement creation and, once set, cannot be altered. *contentDigest* is defined as a string value representing the digest of the digital work (text, picture, video, etc.) by means of a cryptographic hash function, and a *metadata* field containing required information to verify the correctness of the digest (e.g., hash function algorithm used). These fields, along with the use of Public Key Cryptography used to sign the agreement creation transaction, fulfill the CPN-DF-OP requirement.

The *license* field allows to specify the license, including a *license ID, title, href* and three fields containing the legal text, the human-readable license version and the machine-readable one (legalText, commonsDeed and ccRel respectively). These fields fulfill the requirements CPN-DF-LR, CPN-DF-LT, CPN-DF-LF, CPN-DF-EU, CPN-DF-MR.

Finally, the *pricingModel* field defines the type and amount of rewarding. It is specified as a JSON object containing the *type, description,* the *value* along with the relative *currency*. This last field fulfills the CPN-DF-RT requirement.

Co-funded by the Horizon 2020
Framework Programme of the European Union

6 Conclusions

This deliverable has successfully completed a framework analysis of CPN in light of the General Data Protection Regulations. This has concluded that the platform can work within the law provided that users are given rights as mandated. In response to this, the deliverable has produced the first working version of Personal Data Receipts suitable for running the initial pilot of CPN. PDRs can be extended in the future with improved selection mechanism allowing users to tailor the rights they grant the platform, possibly with a server-side module integrating with publisher databases to implement any live rights changes enacted by the user.

The deliverable has also completed the architecture and outline for the Distribution Framework, with sample licenses derived from the publicly available Creative Commons templates. Future work will involve production of the module in a manner which can be easily integrated with the rest of the CPN platform, in a manner which allows for future piloting.

Co-funded by the Horizon 2020
Framework Programme of the European Union

REFERENCES

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Official Journal of the European Union, Vol. L119 (4 May 2016), pp. 1-88

[2] Global Consumer Trust Report 2017. https:// mobileecosystemforum.com/programmes/consumer-trust/global- consumer-trust-survey-2017/

[3] Unlocking the power of data in the UK economy and improving public confidence in its use. https://www.gov.uk/government/publications/uk-digital-strategy/7-data-unlocking-the-power-of-data-in-the-uk-economy-and-improving-public-confidence-in-its-use

[4] Data Protection Eurobarometer report. http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_eurobarometer_240615_en.pdf

[5] Nati M. Personal Data Receipts: How Transparency Increases Consumer Trust. Digital Catapult Centre; 2018.

[6] Creative Commons: http://creativecommons.org

[7] Bloomen: http://bloomen.io

[8] Flixxo: https://www.flixxo.com

[9] Unitalent: https://unitalent.io

[10] Steemit: https://steemit.com

[11] DTube: https://d.tube

[12] Alexandria: https://www.alexandria.io

[13] Brave: https://brave.com

Co-funded by the Horizon 2020
Framework Programme of the European Union