Grant Agreement No.: 761488

D2.4: CPN Open Virtual Platform v3

CPN Platform v3 – Accompanying report

Work package	WP 2
Task	T2.1 – T2.2 - T2.3 – T2.4
Due date	31/12/2019
Submission date	08/01/2020
Deliverable lead	ENG
Version	1.0
Authors	Ferdinando Bosco (ENG), Vincenzo Croce (ENG)
Reviewers	Fulvio D'Antonio (LVT)
Keywords	CPN – Open Virtual Platform – Microservices - Integration

Document Revision History

Version	Date	Description of change	List of contributor(s)
V0.1	05/12/2019	1st version of the deliverable with table of contents	Ferdinando Bosco (ENG) Vincenzo Croce (ENG)
V0.2	15/12/2019	Draft version of the deliverable with integration of contributions from partners	Ferdinando Bosco (ENG) Vincenzo Croce (ENG) Thomas Sounapoglou (Blockachain)
V0.3	27/12/2019	Version of the deliverable ready for internal review	Ferdinando Bosco (ENG) Vincenzo Croce (ENG)
V0.4	08/01/2020	Reviewed Version	Fulvio D'Antonio (LiveTech)
V1.0	08/01/2020	Final Version	Ferdinando Bosco (ENG) Vincenzo Croce (ENG)

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.

Project co-funded by the European Commission in the H2020 Programme		
Nature of the deliverable:	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	

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

EXECUTIVE SUMMARY

The CPN project foresees three major releases of the CPN Open Virtual platform that include functionalities aligned with the requirements gathered for the three CPN pilots. Each platform release includes specific functionalities, selected and implemented after a process of evaluation and prioritization of the user requirements.

The third version of the CPN open virtual platform, delivered as final version, will be described in this document: the CPN Open Virtual Platform v3 accompanying report describes all the updates and improvements of the platform on the previous versions, following the specifications selected during the requirements analysis phase together with the feedback collected during pilot executions.

In particular, the first chapter introduces the scope of the deliverable and the process of implementation and delivery of the CPN platform.

The second chapter describes all the updates and improvements of the platform on the previous version, focussing on the platform core components and the list of integrated technology bricks.

The third chapter will focus on the requirements satisfied, as defined in “D3.4 CPN Technology Bricks v3”, and the list of functionalities implemented. It will describe the results accomplished in terms of requirements coverage and functionalities provided for the execution of the third pilot. In addition, this chapter will report the new functionalities introduced taking into account the feedback collected during pilots executions, project reviews and a final check on the platform functional and non-functional requirements

The fourth chapter describes the integration activities conducted by the technical partners with the external media companies and SMEs. There are mainly two kinds of different integration activities: the first one is focussed on exploiting the CPN platform services outside the consortium and on demonstrating the interoperability and innovativeness of the platform, while the second one allows to include new innovative features to the CPN platforms and demonstrates its flexibility and extensibility.

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

API	Application Programming Interface
ATC	Athens Technology Center
CPN	Content Personalisation Network
DCat	Digital Catapult
ENG	Engineering Ingegneria Informatica
GDPR	General Data Protection Regulation
IMEC	Interuniversity MicroElectronics Center
JSON	JavaScript Object Notation
JWT	JSON Web Token
PoC	Proof-of-Concept
SME	Small Medium Enterprise
SSO	Single Sign On

....

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

1 INTRODUCTION

The CPN project foresees three major releases of the CPN Open Virtual platform that include functionalities aligned with the requirements gathered for the three CPN pilots. Each platform release includes specific functionalities, chosen after a process of evaluation and prioritization of the user requirements.

The CPN Open Virtual Platform v3, described in this document, mainly focuses on the final pilot execution. It integrates all the new technology bricks delivered and implements a set of functionalities detailed on D1.4 Technical requirements (platform and service requirements)[1]. In particular, it focuses on the functionalities necessaries to satisfy the user requirements for the latest pilot.

The process of prioritization of the requirements, already described on D2.2 CPN Open Virtual Platform[2], was performed to plan the activities and define the technical tasks needed to deliver both the technology bricks and the new version of the platform.

In addition, this version of the platform includes some new functionalities, necessary for the integration of SMEs and external media companies, as well as introducing the improvements reported during the feedback collection of the pilot 2 and the second review of the CPN project.

[1] https://www.projectcpn.eu/s/CPN_D14_Technical_requirements_platform_and_service_requirements_20180830_v10.pdf

[2] https://www.projectcpn.eu/s/CPN_D22_CPN_Open_Virtual_Platform_v1_20180629_v10.pdf

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

2 PLATFORM IMPLEMENTATION – UPDATES

In this third version of the CPN Open Virtual Platform, new functionalities have been implemented and made available, through the integration of new modules and the evolution of existing ones, including the core components of the platform.

In order to have an overview of the status of the platform, below are reported all the updates of the core components and in particular of the API Gateway, as well as the list of the technology bricks integrated or updated in this version. The details of the features and improvements of technology bricks are instead reported in D3.4 CPN Technology Bricks v3.

2.1 CORE COMPONENTS

As already described into the D2.2 CPN Open Virtual Platform v1, three core components were deployed into the CPN platform to allow communication among the internal components and exploitation of the CPN functionalities: API gateway, message broker and orchestrator.

At this point of the project only the API gateway includes some additional features, needed for extending the flexibility of the client applications and integrating new specific features becoming from SMEs integration.

2.1.1 API Gateway

The updates in this version of the API Gateway mainly focused on the integration of new authentication mechanisms, working together with the JWT system implemented in the latest version.

In particular, the list of authentication methods has been extended for two reasons:

- to allow the client applications to provide a wider choice for users
- to integrate specific features as SMEs services and Google Assistant

Below the list of allowed authentication mechanism:

- Standard Login (Username and password)
- Facebook Login
- Google Login
- IDWARD login (provided by YOOP, one of the integrated SMEs)

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

The JWT based authentication system, introduced on the second version of the platform, allows to the CPN platform to manage all authorized requests in compliance with GDPR and with a solid and secure mechanism.

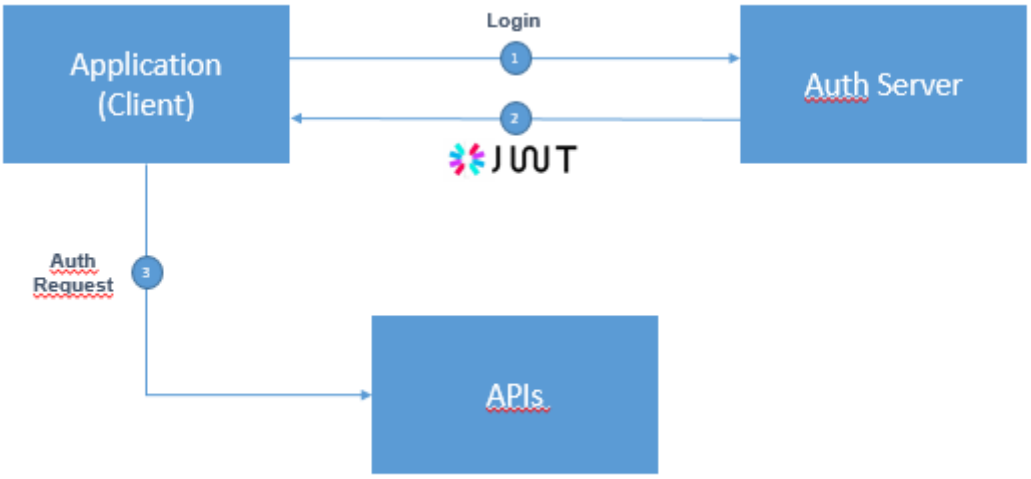


Figure 1: CPN Platform JWT based authorization system

The new authentication mechanisms have not repercussions on the JWT authorization system and are managed in a transparent way by the CPN platform (through the API gateway).

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

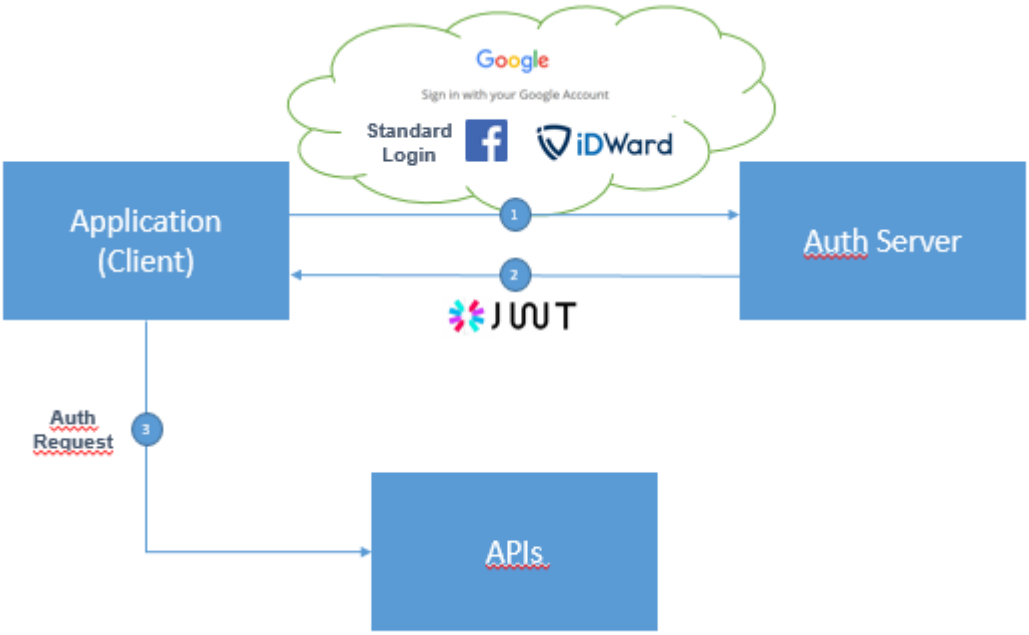


Figure 2: New authentication system with additional sign in and login mechanisms

Furthermore, a major security improvement has been made, including the HTTPS protocol in all the REST APIs exposed by the CPN platform.

2.2 CPN Technology Bricks

In the third version of CPN platform, two new technology bricks were introduced (deployed and integrated) and seven more were updated. A more detailed description of these changes, is provided by D3.4 CPN Technology Bricks v3 Deliverable.

In the following section the list of technology bricks that have been updated or delivered in the third prototype:

Table 1: CPN Technology Bricks for third prototype

Layer	Name	Partner	Status
Content Technology Bricks	Fine Grained Entity Recognition Module	Imec	New
	Topic Extractor	LiveTech	Updated
	Recommender AB-Testing	LiveTech	Updated

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

User Technology Bricks	User Modelling	LiveTech	Updated
	Reader's App	ATC	Updated
	Personal Data Receipt	DCat	Updated
Mapping technology Bricks	Distribution Framework	DCat	New
	Producer's App	ENG	Updated
	Recommender	LiveTech	Updated

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

3 CPN PLATFORM – 3RD PROTOTYPE DELIVERY

The SCRUM agile process for delivering the platform, adopted for the first two prototypes, was also applied for the implementation and delivery of the final one.

The process consisted of 3 phases: two starting activities steps and one cyclic step for task development. In detail:

- Requirements prioritization for pilot 3
- Mapping Requirements/Technology Bricks
- Sprints (Implementation Results)

In the development of the final version of the components and of the CPN platform, a series of improvements and new functionalities have been taken into consideration, as well as the originally foreseen requirements.

These extensions were reported as an activity plan in the “D6.5 2nd Periodic Report” and monitored through a secondary Trello board.

Furthermore, since this is the final version of the platform, a check is reported below on all functional and non-functional requirements relating in a specific way to the CPN platform, as described in "D1.4 Technical requirements (platform and service requirements)".

3.1 USER REQUIREMENTS FOR PILOT 3

The first step of the process was the requirements prioritization for pilot 3. In this phase starting from the list of requirements expected for the third pilot iteration, as described in D1.4, the media partners (VRT, DW, DIAS), that in the case of the Scrum methodology assumed the role of Product Owners, gave a priority for each user requirement expected.

Being the latest version of the platform, in addition to prioritizing the requirements, a feasibility study was also carried out on these.

The result of these two activities is a list of requirements sorted by rank (where the rank represents the “development priority”), named “Pilot 3 Features”, which includes all the requirements to implement for the latest version of the platform and a list of discarded requirements, deemed no longer relevant or not technically implementable.

The “Pilot 3 Features” list is reported below, as input for the mapping of the requirements with the technology bricks, while the discarded requirements will be described in paragraph 3.4 (result) together with the final list of the implemented requirements.

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

Table 2: List of prioritised requirements for pilot 3

Priority	Requirement ID	Description
1	UR- UP7.1	The system should check on what device the user is consuming the content
2	UR-PS2.3	The system should allow producers contributions are used and distributed to readers
3	UR- UP6.4	The system should be able to offer insights and advice based on what it learn about what a user consumed in relation to a certain entity (e.g. a place)
4	UR- PS1.4	The system should be able to show these numbers during the creation process of the content
5	UR- UP6.5	The system should allow the user to delete part of the systems knowledge for specific time frames back in time from the moment of viewing
6	UR- PS2.2	The system should provide contract templates to allow freelancers to easily work together and with editors, to define and track the scope of individual contributions and expected revenues
7	UR- AF7.4	The system should be able to offer a feedback interaction to determine the ground level of personalisation based on mood, time and interest
8	UR- PS2.1	The system should allow for an easy integration into the producers workflow
9	UR- UP1.3	The system should be able to offer personalised content on the basis of the users mood or values
10	UR- AF1.3	The system should offer the user an easy overview of what content from which sources he has consumed over a certain period of time

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

11	UR- AF6.4	The system should be able to memorize where a user left off and restart at the same point
12	UR- AF3.2	The system should offer the user a short overview of all important headlines at a specific point in time with access to more details upon request
13	UR- AF7.1	The system should offer user feedback requests in a playful/entertaining way
14	UR- UP7.2	The system should adjust its content offering based on the type of device the user is using
15	UR-UP8.3	The system should be able to surprise the user with content, he/she would not have chosen themselves
16	UR-UP9.6	The system should allow the user to add external data to update their profile

The iterative evaluation process led to focus on some macro-functionalities to be implemented during the third phase:

- Cross-Channel
- Transparency
- UI and UX Improvements
- Breaking News
- Recommender Improvements

These macro-functionalities include improvements of the already existing ones and the definition of new requirements and technical tasks.

3.1.1 Mapping requirements/Technology bricks

After the definition of the prioritized backlog in the frame of the Scrum Framework, a mapping between the requirements for Pilot 3 and the Technology Bricks (new or updated) took place. This process was followed through the Scrum Team, involving both media and technical partners.

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

The input for the mapping has been the technical requirements defined on D1.4 (after the already mentioned prioritization process) together with the new macro-functionalities identified, the list of the affected Technology Bricks (section 2.2) and the technical tasks defined for this phase of the project and described in details on the project Trello board.

Below the list of technical task identified, linked with the respective user requirement, when expected:

Table 3: Mapping between technical tasks and technology bricks for pilot 3 implementation

Requirement ID	Task	Technology Brick
UP 1.3	TUP1.3.1 Classify articles with mood criteria	Uplifting/Depressing classifier
	TUP1.3.2 Provide articles following mood criteria (experimental)	Recommender
UP 6.4	TUP6.4.1 Show questions to the user	Reader's App
	TUP6.4.2 Update user model based on types of content	User Modelling
UP 6.5	TUP6.5.1 Provide history information on system knowledge	Reader's App
	TUP6.5.2 Update user model based on deletions of system knowledge	User Modelling
	TUP6.5.3 Add the functionality of deleting specific user activities in specific timeframes	Reader's App
UP 7.1	TUP7.1.1 Save user device info	Reader's App
	TUP7.1.2 Update user model based on device info	User Modelling

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

UP 7.2	TUP7.2.2 Provide content based on type of device	Recommender
UP 8.3	TUP8.3.1 Display news that might surprise the user	Reader's App
	TUP8.3.2 Implement news "surprise" algorithms	Recommender
UP 9.6	TUP9.6.1 Allow user to upload a file	Reader's App
	TUP9.6.2 Update user model based on external data	User Modelling
	TUP9.6.3 Provide additional information becoming from Twitter	Twitter Analytics
	TUP9.6.4 Provide additional information becoming from Facebook	Reader's App (integration with SMEs)
AF 1.3	TAF1.3.1 Provide content usage statistics	Reader's App
	TAF1.3.2 Integration of IDWARD login for user statistic dashboard	Reader's App (integration with SMEs)
AF 3.2	TAF3.2.1 Display a date filter headlines	Reader's App
	TAF3.2.2 Add a date filter for latest articles	Producer's App
AF 5.3	TAF5.3.1 Display justification based on item level	Reader's App
	TAF5.3.2 Provide justification based on item level	Recommender
AF 6.4	TAF6.4.1 Track left off/ start point	Reader's App
	TAF6.4.2 Measure how far a reader scrolls through the list of articles in each tab	Reader's App
	TAF6.4.3 Update user model with tracking actions	User Modelling

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

	TAF6.4.4 Include tracking actions on recommender	Recommender
AF 7.1	TAF7.1.1 Offer feedback requests in a playful way	Reader's App
AF 7.4	TAF7.4.1 Display feedback information	Reader's App
	TAF7.4.1 Integrate Qualtrics 3API for user experience management	Reader's App
PS 1.4	TPS1.4.1 - Producer's app must expose API to retrieve analytics	Producer's App
PS 2.1	TPS2.1.1 - Prodcuer's app must expose API to integrate media company articles	Producer's App
	TPS2.1.2 - Development of a Javascript module for tracking user actions and providing recommendation	Producer's App
	TPS2.1.3 - Development of a Javascript module for identifying user without signup (fingerprint and/or cookies based)	Producer's App
PS 2.2	TPS2.2.1 - Producer's Dashboard UI must include a section for freelancer integration	Producer's App
	TPS2.2.2 Provide rewards for contributors	Reward Framework
PS 2.3	TPS2.3.1 Provide producer's usage information	Reader's App
	TPS2.3.2 Provide usage statistics	Reward Framework
General	New Platform service for sending email	CPN Platform
General	Integration of Facebook Login	Reader's App CPN Platform
General	Integration of Google Login for Google Assistant	Reader's App CPN Platform

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

Breaking News	TP1.1 - Producer's app must to allow to set an article as "breaking news"	Producer's App
	TP1.2 - API Gateway must expose API to retrieve a list of breaking news	CPN Platform
	TP1.3 – Recommender should introduce breaking news in recommendation service if any	Recommender
Cross Channel	Google Assistant integration	Reader's App
	Smart TV app dedicated services	CPN Platform Recommender
UI Improvements	Switching streams only by tapping, not by swiping	Reader's App
	Improve display of videos, Links, interactive elements inside articles	Reader's App
	Make it clear to the user where they can see what they have already consumed	Reader's App
	Fighting FilterBubbles: Visualisation of what a user read throughout a specific period (e.g. a week)	SMEs integration (YOOP)
	Branding customization (colors, logo, etc.)	Reader's App
	Add label of the article's category	Reader's App Producer's App

[3] https://www.qualtrics.com/it/

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

3.2 RESULTS

The third step was the implementation of the tasks in an iterative way. During the third phase, we preferred to have a more frequent monitoring of the activities and to organize sprints of a week. After each sprint a technical conference call was arranged in order to review the results of the just terminated sprint and formulating a detailed plan for the following sprint execution.

As a result of these activities, the final version of the CPN platform was delivered. The following paragraphs show the list of user requirements completed, a check on platform requirements (functional and non-functional) and the final status of the platform.

A detailed description on the new macro-functionalities and in particular on the improvements expected for pilot 3 will be provided together with the pilot 3 evaluation.

3.2.1 User requirements

Table 4: Status of the user requirements for pilot 3

Requirement ID	Requirement Description	Status	Notes
UR-UP 1.3	The system should be able to offer personalized content on the basis of the users mood or values	Partially Completed	Uplifting/Depressing classification is available but user mood will be tested as experimental in the third pilot
UR-UP 6.4	The system should be able to offer insights and advice based on what it learns about what a user consumed in relation to a certain entity (e.g. a place)	Completed	
UR-UP 6.5	The system should allow the user to delete part of the systems knowledge for specific time frames back in time from the moment of viewing	Completed	
UR-UP 7.1	The system should check on what device the user is consuming the content	Completed	
UR-UP 7.2	The system should adjust its content offering based on the type of device the user is using	Partially Completed	The device of the user is tracked and contents will be presented in a different way in base of the used device (responsiveness). Not enough technical resources & capabilities available to go

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

			deeper in detail within this requirement (also impact on work for media companies).
UR-UP 8.3	The system should be able to surprise the user with content, he/she would not have chosen themselves	Completed	
UR-UP 9.6	The system should allow the user to add external data to update their profile	Completed	
UR-AF 3.2	The system should offer the user a short overview of all important headlines at a specific point in time with access to more details upon request	Completed	
UR-AF 5.3	The system must make it transparent to the users why they are shown certain content, based on an item level	Completed	The requirement was completed for pilot 2 but transparency macro-functionality was improved
UR-AF 6.4	The system should be able to memorize where a user left off and restart at the same point	Completed	
UR-AF 7.1	The system should offer user feedback requests in a playful/entertaining way	Completed	
UR-AF 7.4	The system should be able to offer a feedback interaction to determine the ground level of personalization based on mood, time and interest	Completed	
UR-PS 1.4	The system should be able to show these numbers during the creation process of the content	Completed	
UR-PS 2.1	The system should allow for an easy integration into the producers workflow	Completed	
UR-PS 2.2	The system should provide contract templates to allow freelancers to easily work together and with editors, to define and track the scope of individual contributions and expected revenues	Completed	
UR-PS 2.3	The system should allow producers to transparently see	Completed	

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

	how often their contributions are used and distributed to readers		

3.2.2 Platform Requirements

The identification of the technical requirements, starting from the analysis of user requirements, also led to a definition of functional and non-functional requirements for the CPN platform, intended as an administration and integration platform for microservices (technology bricks and core components).

These requirements were listed on D1.4 and their status was checked and reported in this deliverable.

Functional requirements

Table 5: CPN Platform functional requirements status

ID	Description	Status
PR-1	The Platform should have a User friendly UI in order to manage the platform components and host	Done
PR-2	The Platform UI should allow creating and managing different User roles with different permissions (at least two: administrator and developer)	Done (4 different roles)
PR-3	The Platform UI should contain a section with CPN components catalog	Done
PR-4	The Platform UI should have a section with the API documentation	Done
PR-5	The administrator should be able to monitor the entire platform resources	Done
PR-6	The administrator should be able to manage each single component of the platform	Done
PR-7	The administrator should be able to add new component into the platform from a catalog	Partially Done (Not all the components

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

		are listed on the catalog)
PR-8	The developer should be able to add a new component into the CPN catalog	KO (Only the admin can add new components on the catalog)
PR-9	The developer should be able to view and test all the existing services	Done
PR-10	The developer should be able to use the existing services	Done

Non-functional requirements

Table 6: CPN Platform non-functional requirements status

ID	Description	Status	Notes
NFR-1	The system should be able to scale horizontally or vertically depending on the demands related to data ingestion, processing and storage.	Done	Microservices are scalable with replication and clusterisation. API Gateway can manage and redistribute on different hosts thousands of calls simultaneously (tested with 1000 requests, no noticeable delay)
NFR-2	It must be possible to decompose the solution in different microservices	Done	All the components (including the core components) are deployed as microservices
NFR-3	The system should enable a versioned and automated deploy mechanism for the microservices	Done	Private docker registry enabled

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

NFR-4	The system should include a centralized authentication and authorization mechanism	Done	Handled by the API Gateway
NFR-5	The system should enable monitoring and logging mechanisms	Done	The Rancher UI allow to monitoring and logging the entire host and each single microservice
NFR-6	The system should enable synchronous communication among the microservices, through orchestration mechanisms	Done	Handled by Orchestrator
NFR-7	The system should enable asynchronous communication among microservices, through messaging mechanisms	Done	Handled by Message Broker
NFR-8	The system should offer optimized endpoints in base of different kind of client's applications (mobile, web, smart tv...)	Done	Web, Mobile, SmartTV and Google Assistant supported
NFR-9	The system should be extensible with new functionalities and new components	Done	SMEs additional services integrated
NFR-10	The APIs exposed by the microservices should be documented in a standard format.	Done	OpenAPI 2.0[4]
NFR-11	The system should keep the data private and only accessible via authenticated APIs	Done	4 different roles handled. All the APIs are under authentication through JWT.
NFR-12	The data produced, consumed and transformed shall be documented in an information model which shall	Done	Data models included on Appendix1

[4] https://github.com/OAI/OpenAPI-Specification/blob/master/versions/2.0.md

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

	also include the relationships between information types.		
NFR-13	The system shall be able to exchange data with a great number of devices and, at the same time, preserving its computational capacity.	Done	

3.2.3 Status of the platform

The final version of the CPN platform was deployed and released in two twin environments (the development environment will be used as backup of the production one), ready to be exploited for the final pilot phase.

This final version includes:

- 12 Technology Bricks deployed as microservices (11 backend services + 1 web dashboard)
- 1 Technology Brick as client application (Android)
- 3 Core components
- 1 Smart speaker integration (Google Assistant)
- 1 external mobile app integration (VRT MyNews)
- 8 Additional services from SMEs

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

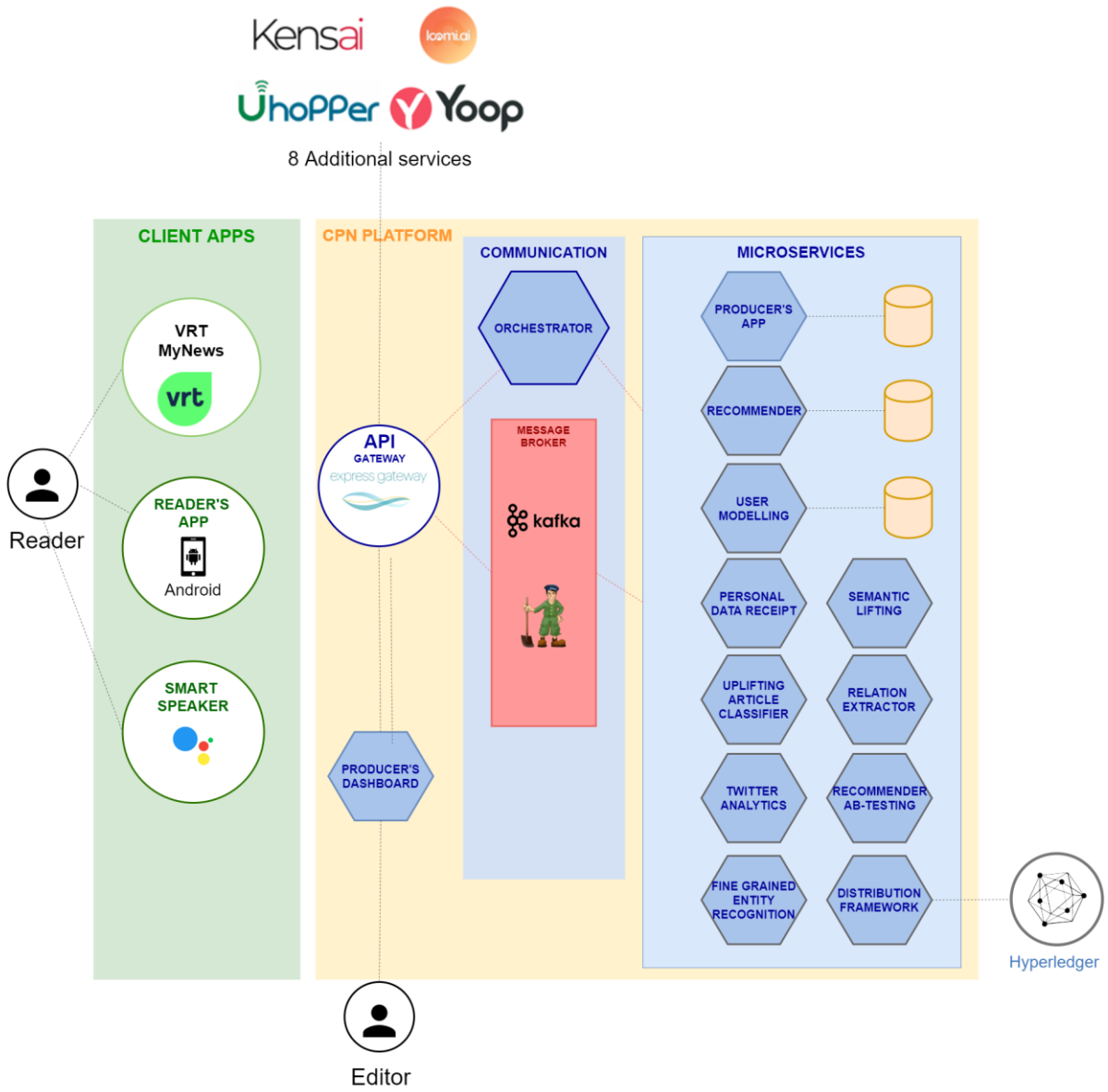


Figure 3:Final version of the CPN Platform

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

4 PLATFORM INTEGRATION

From the beginning, the main objectives of the CPN platform have included extensibility and reusability. The design of the CPN architecture, included in the D2.1 Reference Architecture[5], laid the foundations for the platform which gradually evolved into 3 prototypes up to the final one.

The extensibility process was enabled, in addition to increasing the number of technology bricks and features for each pilot, also planning the integration of some components and services offered by external partners, which were well integrated with those already existing. This made it possible to add a series of features to the platform or enrich existing ones.

The reusability one, which is very important for the exploitation of the platform, was carried out by meeting the needs of the media stakeholders and trying to create a process as simple as possible for the integration of CPN services on already existing media ecosystems.

4.1 SMEs

The integration of external services, which will be demonstrated in the third pilot, was done in collaboration with the SMEs selected during the two Hackathon phases carried out in February and June 2019.

This selection process, described in detail in D4.3: Cycle 2 Piloting Report[6], led to the identification of 4 SMEs that expanded the offer of the CPN platform with their tools and services: Loomi.ai, U-Hopper, YOOP and Kensai[7].

Figure 4:the 4SMEs selected from the Hackathons

After the SME selection process, 8 services were identified to be implemented through 4 PoCs (one for each SMEs) and integrated in the CPN ecosystem in collaboration with the CPN technical team.

The development phase of the PoCs ends at the end of the year and the integration will continue in the coming weeks in preparation for the pilots. The following paragraphs briefly describes from

[5] https://www.projectcpn.eu/s/D21-CPN-Reference-Architecture-v10.pdf

[6] https://www.projectcpn.eu/s/D43_2ndCycleEvaluationReport_v10.pdf

[7] https://www.projectcpn.eu/news-3/2019/6/17/highlighting-the-startups-working-with-cpn

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

a technical point of view the objectives of each PoC in terms of offered features and integration with CPN.

PoC1 – YOOP

Yoop is integrating and customizing for the CPN platform an already SSO extended system call ID ward. ID Ward is a unified user identity for content providers - a simple way for users to reap the benefits of news personalisation without losing control over their personal data.

In the CPN context, is building a proof-of-concept (PoC) software infrastructure for secure and transparent collection, aggregation and storage of user data. The innovation relies on a networked approach to user authentication and a bespoke data security and transparency tool for increased privacy control. The PoC will be built to support the CPN's machine learning technology using privacy by design principles.

The proof of concept will be used to validate the following assumptions:

- the proposed networked approach can provide a better experience at user sign in
- the networked approach can provide much richer user profile data;
- the proposed data security and transparency tool gives data subjects more control over their personal data and increases their default privacy.

The proof-of-concept is provided as a service and in particular three different services will be enabled:

- Single sign on (SSO): comprises SSO interface and API
- User data collector: comprises user profile data, data API (send and retrieve) and access control system
- Data privacy, security and transparency tool: comprises a user profile, permission and data permission interface, data encryption and communication tool, and access log.

The ID ward SSO system well integrates with the CPN ecosystem and in particular with the Reader's app, in which the ID ward login will be available and the user modelling component, that will exploit user profile data becoming from many sources to enrich the user profiles.

Furthermore, the topic of transparency will also obtain enormous benefits, as for users who decide to register via ID ward a dashboard will be made available which shows all the activities done by the user and allows them to obtain total control of their data.

During the evaluation phase, in addition to using ID ward in the CPN platform, it will also be integrated on some websites provided by SigmaLive, creating a real network of news websites that will allow us to increase the knowledge of the users.

PoC2 – U-Hopper

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

U-Hopper is customising for CPN an already existing solution, called Tapoi. Tapoi is a customer intelligence service that builds profiles based on customer online actions, allowing businesses to provide targeted and tailored services, personalising user experience and achieving higher conversion.

In the CPN context, Tapoi will provide additional information extracted from Facebook, to enrich the user profiles, inferring user's interests from their social activities. In this way it is possible to mitigate both the cold start and the filter bubble problems.

In detail:

- Cold start problem. Tapoi is able to generate a consumer's map of interests even for newcomers, at the exact moment of the first online interaction with CPN. Therefore, potentially unknown users can be targeted with the right content, at the right time.
- Filter bubble problem. Tapoi is able to help the management of the distribution (quantity, frequency) of interest-related news to consumers, by means of a customized algorithm. An appropriate choice of the parameters ensures that consumers are exposed occasionally to content that is out of their interest sphere, thereby mitigating the echo chamber effect.

The integration with the CPN ecosystem is very deep. Indeed, the TAPOI system will exploit a new authentication system provided by the reader's app (the Facebook Login) and will extract data from users that will decide to use Facebook as a login system.

All the data analysed by TAPOI will be accessible from the CPN recommender (via API) and will be used to enrich the user profiles and improve recommendations.

During the pilot 3, TAPOI will be evaluated with SIgmaLive users.

PoC3 – Kensai

Kensai is building a solution for newsrooms to monitor the narrative in real-time on Twitter around any news story to personalise it to their audience. This feature can also be used to address the filter bubble issue by providing contrasting views underneath the article to give a broader picture.

The PoC developed by Kensai, will include a dashboard for the editors, from which they can monitor up to 10 different topics through many analytics and statistical data extracted from Twitter.

In this case, no integration with CPN existing services is expected, so no API or integration services will be developed and only a web UI will be provided.

PoC4 – Loomi.ai

Loomi.ai is building a personalisation news focused ontology for the CPN project with the main objective of enhancing the accuracy of the personalisation recommender algorithms. The ontology can be used to formalise both the user and her/his interests, and the news content. The developed ontology allows matchmaking between user and content at different levels.

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

The service will have the capacity to help other news providers increase the accuracy of the personalisation they do through either the CPN project or accessing the service directly.

In the context of the PoC, the ontology will be only in English and for this reason the integration and evaluation phase will include only DW contents.

- access to get normalised named entities (entity alias matching)
- access to get news category for either news article URL or a set of extracted named entities from an article (news categorisation)

In particular the PoC will provide two different services via API.

4.2 EXTERNAL MEDIA COMPANIES

The integration of the platform with external systems is a very important aspect for the flexibility and reusability of the platform.

For the usage of the CPN platform by external media companies, the approach chosen was to make the CPN platform a Platform-as-a-service, in which a media company can choose which services and features to use (from individual components or from the platform itself) and these are offered in a standard and interoperable way through the API gateway.

Furthermore, all the client applications developed are available to external stakeholders, who can choose to use them or not together with the selected services.

All these choices were made also from the point of view of the stakeholders themselves (in particular the external media companies) who, after being contacted or having contacted us through our communication channels, were followed in the process of approach to the CPN ecosystem and whose feedback has been useful in preparing a simple and performing integration and exploitation plan.

The result of these activities, which involved the dissemination, communication and exploitation team but also the technical one, was:

- Develop a 3-step technical integration plan for media companies
- Propose 2 different approaches for testing and evaluating the services offered by the CPN platform, based on the needs of media companies
- Implement a convincing business offer for the exploitation of the results of the CPN platform

Below we describe what the platform technically offers for integration and for the evaluation phase. The exploitation topic will be covered in the final exploitation results of the project.

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

4.2.1 Technical integration

The technical integration proposed to the external media companies is described in the document attached as Appendix 2.

It follows an easy 3-steps process:

- Step 1: Contents injection/ingestion
- Step 2: User identification and profiling
- Step 3: Content recommendation

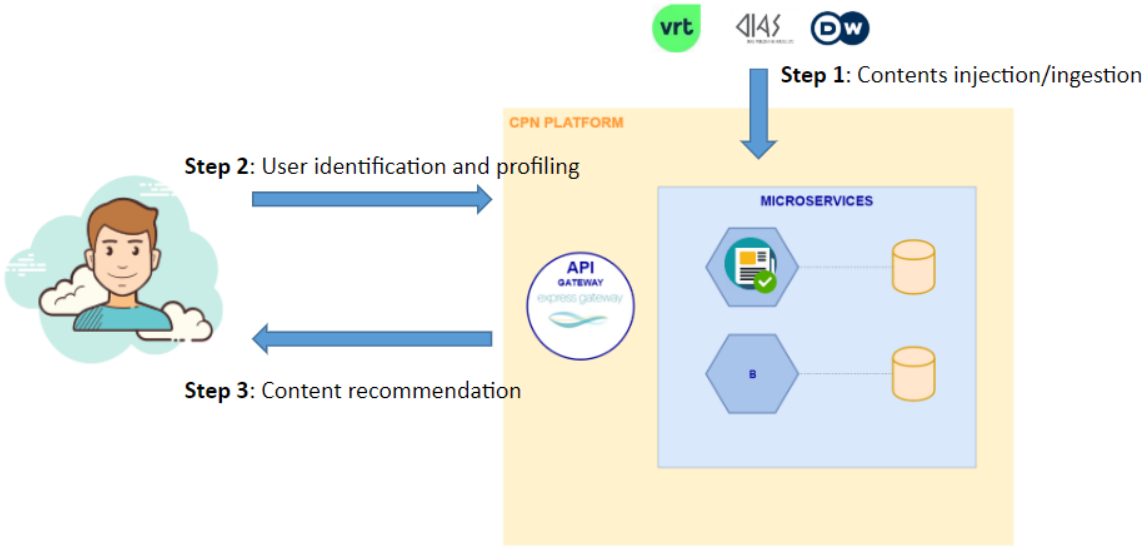


Figure 5:3-steps integration process for external media companies

The first step is mandatory for the integration of a new media company within the CPN platform. It is possible to push new content via API or provide us an already existing repository (RSS feed, REST APIs, etc.). In this case, the Producer's App component will implement a new connector in order to map the new contents with the CPN data model and introduce the new content within the platform. This is a very flexible approach that allows to the media companies to select the best practice for them and requires the minimum effort on the other part.

The first step is mandatory for the integration of a new media company into the CPN platform. It can introduce new content via API or provide us with an existing repository (RSS feed, REST API, etc.). In this case, the Producer's App (one of the technology bricks) will implement a new connector to map new contents with the CPN data model and introduce them into the platform. This is a very flexible approach that allows media companies to select best practices for them and requires very little effort.

The second and the third step are not mandatory. Their implementation depends on the testing approach that media companies selected (the testing and evaluation approaches are described in the next paragraph).

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

In case they want to implement these steps, in addition to providing the APIs necessary for the integration of all services, an alternative widget has been developed. This widget is based on HTML5 and Javascript technologies, so it is very flexible and pluggable in already existing websites. It provides:

- user identification through anonymous browser fingerprint[8]
- user profiling, collecting reading, like, dislike and share actions (if possible)
- content recommendation and visualisation

The integration of this widget is currently underway with one of the media companies interested in the integration of CPN services (RCS) and will be evaluated during the testing phase.

4.2.2 Testing and evaluation

For testing and evaluation phase, two different approaches have been proposed to media companies:

- ➔ Using the CPN Android App with their own contents
- ➔ Integrate all the services within their own infrastructure

In the first case, the integration process is very simple since it only requires to integrate their own contents within the CPN platform (the first step described in paragraph before) and allows a complete evaluation of all the services offered. On the other hand, it requires that users who use the app (Android) are recruited and the evaluation could be affected by UI or UX problems that do not reflect the choices of the media company.

In the second case, the integration requires some effort from the media company side for introducing the new services in their own infrastructure (CMS, Websites, apps, etc...). As already mentioned, for this type of integration, well detailed in the document provided to them, it is possible to use the APIs made available to them or the developed widget (which, at the moment, is still in the development and testing phase)

CONCLUSIONS

This document represents the report of all the activities conducted for the release of the final version of CPN Open Virtual platform.

The CPN platform v3 was packaged and it is available for the partners in internal repository.

In order to test the platform and verify the status of delivery as described in this report, the following software prototypes are available to internal partnership:

[8] https://github.com/Valve/fingerprintjs

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

- CPN Microservices Platform v3 – Two separated environments (One for production and one as backup)
- CPN API Gateway v3 - Two separated environments (One for production and one as backup)
- Producer's Dashboard v2
- Android Reader's App v2

The final version of the platform will be used both for the execution of the third pilot (6 weeks) and for the tests and evaluation of external media companies. These latest tests, scheduled for early 2020, will be more important than previous executions, as they are larger in terms of users involved and execution times. The results will demonstrate whether the platform is exploitable and ready to actually be placed on the market.

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

REFERENCES

[1] https://www.projectcpn.eu/s/CPN_D14_Technical_requirements_platform_and_service_requirements_20180830_v10.pdf

[2] https://www.projectcpn.eu/s/CPN_D22_CPN_Open_Virtual_Platform_v1_20180629_v10.pdf

[3] https://www.qualtrics.com/it/

[4] https://github.com/OAI/OpenAPI-Specification/blob/master/versions/2.0.md

[5] https://www.projectcpn.eu/s/D21-CPN-Reference-Architecture-v10.pdf

[6] https://www.projectcpn.eu/s/D43_2ndCycleEvaluationReport_v10.pdf

[7] https://www.projectcpn.eu/news-3/2019/6/17/highlighting-the-startups-working-with-cpn

[8] https://github.com/Valve/fingerprintjs

APPENDIX 1 – DATA SCHEMAS

Authentication

The API Gateway handle in a centralized way the authentication and authorization to the CPN platform.

Below the data schema used for authentication:

```
authSchema:
{
        email: String, //unique and required
        username: String,
        password: String, //Hashed password,
        origin: { type: String, enum : ['NEW','DIAS', 'VRT']},
        createdAt: {type: Date, default: new Date()},
        updatedAt: {type: Date, default: new Date()},
        role: { type: String, default: 'user' },
        registeredWithFacebook: { type: Boolean, default: 'false' },
        loggedInWithFacebook: { type: Boolean, default: 'false' },
        facebookUserId: String,
        registeredWithGoogle: { type: Boolean, default: 'false' },
        loggedInWithGoogle: { type: Boolean, default: 'false' },
        googleUserId: String,
}
```


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

PDRs - consumed/produced data format

Data consumed

Data consumed by this module includes:

- Categories of personal data collected from readers
- (this can be pre-filled: Reasons for collecting such data, How data are processed, Where data are stored)
- If and what type of such personal data are shared with third parties (if any)

Data produced

- A human-readable version of the PDR; this depends on the mean used to share PDR with the reader (e.g., email and HTML)
- A hashed version of this receipt to be stored in the blockchain (this is irrelevant as our module will directly interface with the blockchain)

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

DS4Biz-Recommender - consumed/produced data format

The module is in charge of computing the most suitable news recommendations for CPN users. It has to analyze the users' profiles and collected news to find the most "interesting" news items to be proposed by the app.

Data consumed

The data consumed by this module is the output of DS4Biz-UserModelling module and the output of the news and social media collector modules

Data produced

We will start with a very simple schema for the recommendations to be updated in the next releases.

Schema Version 0.1:

```
recommenderSchema:
[{
        id : String, // Internal Id of the recommendation
        user_id : String, // Internal Id of the user
        score : Number,   // The relevance score of this recommendation as computed by the
        recommender engine
        date : Date  // When this recommendation was computed
}]
```

Examples:


```
 {
   "id":"xxxxxx",
   "user_id":"yyyyyyy",
   "item_id":"zzzzzzz",
   "date":"Fri, 11 May 2018 08:40:10 +0000",
   "score":0.8
 },
 {
   "id":"xxxxxx",
   "user_id":"yyyyyyy",
   "item_id":"zzzzzzz",
```


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

```
    "date":"Fri, 11 May 2018 08:40:10 +0000",
    "score":0.5
  }
]
```


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

DS4Biz-UserModelling - consumed/produced data format

The module is in charge of maintaining users' data keeping track of his/her interests, history of click and news consumption and demographic data. This data will be mainly used by the recommender engine in order to select the most suitables news in according to a given user profile.

Data consumed

The data consumed by this module will be the list of events generated by the users (posted through the broker). The kind of events that the module will be processing are the following:

1. "News": when a news item is added to the broker it is retrieved by the module in order to tag, analyze it, extract topics,etc.
2. "Users_feedback": every action initiated by a customer that is collected by the platform it is collected, processed and stored in the user profile. Typical events are:
    a. User clicks
    b. Users ratings (explicit rates, thumbs up/down, etc)
    c. User profile update (e.g. change of name, location, age, etc)

Data produced

We will start with a very simple schema for the recommendations to be updated in the next releases.

Schema Version 0.1:

```
userModellingSchema:
{
  "user_id": "String",
  "demo": {
    "gender": "String",
    "age": "Number",
    "name": "String",
    "email": "String"
  },
  "interests": [
    {
      "id": "String",
      "label": "String",
```


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

```
      "score": "Number",

    }
  ],

  "activities": [
    {
      "item_id": "String",
      "event": "String"
    }
  ]
}
```

Examples:

```
{
  "user_id":"xxxxxxxx",
  "demo":{
    "name":"xxxxxx",
    "email":"xxxxxx",
    "age":"42",
    "gender":"M"
  },
  "activities":[
    {
      "event":"CLICK",
      "item_id":"xxxxxxxx"
    },
    {
      "event":"READING",
      "item_id":"xxxxxxxx"
    },
    {
      "event":"LIKE",
      "item_id":"xxxxxxxx"
    },
    {
      "event":"DISLIKE",
```


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

```
      "item_id":"xxxxxxxx"
    }
  ],
  "interests":[
    {
      "id":"xxxxx",
      "label":"politics",
      "score":"0.5"

    },
    {
      "id":"xxxxx",
      "label":"artificial intelligence",
      "score":"1.2"

    }
  ]
}
```


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

Relation Extraction Module- consumed/produced data format

http://surdeanu.info/kbp2014/TAC_KBP_2014_Slot_Descriptions.pdf

Input DW Example :

Example of expected Input:

```
{
        _id : String, // Internal Id
        originId : String, // Original Id from source
        origin : String, // Source
        url : String //original item url
        category : String, // Category of article
        title : String,
        text : String,
        language : String,
        author : String,
        date : Date, // date in Date format
        dateStr : String,     // date in String format
        timestamp : Number, // date in timestamp format
        location : { //location in geoJson format
                type: Object,
                index: '2dsphere',
                sparse: true },
        tags : [String] //list of tags
}
```

```
{
        "Url":" ",
        "date": "2018-05-08T15:41:57.140Z",
        "text": "Transgender student Gavin Grimm defeated the board of his old high
        schoigin": "DW",
        "_id": "5af1col before a federal court on Tuesday over the right to use the
        bathroom corresponding with his gender identity.  US District Judge Arenda
        Wright Allen in Norfolk rejected a bid by the Gloucester County School Board to
        dismiss the civil rights lawsuit filed by Grimm.",
        "originId": " ",
        "or61d3ecbb81a00bf0e8b",
```


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

```
        "tags": [],
        ]
}
```

Data produced

Data produced by this module includes **per sentence**,

- A ***entitymentions*** field which contains locations and tags of named entities
- A ***kbp*** field with relations detected from the predefined schema
- An ***index*** field of the sentence

Example of output:

```
{
        originId : String, // Original Id from source
        origin : String, // Source
        url : String //original item url
        title : String,
        language : String,
        author : String,
        date : Date, // date in Date format
        dateStr : String,    // date in String format
        timestamp : Number, // date in timestamp format
        sentences: { [ {
                        index: int,
                        entitymentions: [ {
                                        CharacterOffsetBegin: int,
                                        CharacterOffsetEnd: int,
                                        DocTokenBegin: int,
                                        docTokenEnd: int,
                                        text: text // Surfaceform of Entity, eg. Barrack Obama
                                        ner: text //Nertype, eg. PERSON
                                        tokenBegin: int,
                                        tokenEnd: int
                                        }
                                        ],
                        kbp: [ {
                                        Object: string, // Object Entity, eg. Barrack Obama
                                        objectSpan: [int,int],
                                        relation: string, // Object Entity, eg. per:title
```


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

```
                                relationSpan: int,
                                subject: text // Object Entity, eg. President
                                subjectSpan: [int,int]
                            } ]
                        } ] }

}
```

Example output from the relation extraction module:

```
{'sentences': [{'entitymentions': [{'characterOffsetBegin': 12,
                                    'characterOffsetEnd': 19,
                                    'docTokenBegin': 1,
                                    'docTokenEnd': 2,
                                    'ner': 'TITLE',
                                    'text': 'student',
                                    'tokenBegin': 1,
                                    'tokenEnd': 2},
                                   {'characterOffsetBegin': 20,
                                    'characterOffsetEnd': 31,
                                    'docTokenBegin': 2,
                                    'docTokenEnd': 4,
                                    'ner': 'PERSON',
                                    'text': 'Gavin Grimm',
                                    'tokenBegin': 2,
                                    'tokenEnd': 4},
                                   {'characterOffsetBegin': 100,
                                    'characterOffsetEnd': 107,
                                    'docTokenBegin': 17,
                                    'docTokenEnd': 18,
                                    'ner': 'DATE',
                                    'normalizedNER': 'XXXX-WXX-2',
                                    'text': 'Tuesday',
                                    'timex': {'tid': 't1',
                                              'type': 'DATE',
                                              'value': 'XXXX-WXX-2'},
                                    'tokenBegin': 17,
                                    'tokenEnd': 18},
                                   {'characterOffsetBegin': 54,
                                    'characterOffsetEnd': 57,
```


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

```
                        'docTokenBegin': 8,
                        'docTokenEnd': 9,
                        'ner': 'PERSON',
                        'text': 'his',
                        'tokenBegin': 8,
                        'tokenEnd': 9},
                       {'characterOffsetBegin': 162,
                        'characterOffsetEnd': 165,
                        'docTokenBegin': 27,
                        'docTokenEnd': 28,
                        'ner': 'PERSON',
                        'text': 'his',
                        'tokenBegin': 27,
                        'tokenEnd': 28}],
  'index': 0,
  'kbp': [{'object': 'student',
           'objectSpan': [1, 2],
           'relation': 'per:title',
           'relationSpan': [-2, -1],
           'subject': 'Gavin Grimm',
           'subjectSpan': [2, 4]}]},
 {'entitymentions': [{'characterOffsetBegin': 184,
                      'characterOffsetEnd': 186,
                      'docTokenBegin': 31,
                      'docTokenEnd': 32,
                      'ner': 'COUNTRY',
                      'text': 'US',
                      'tokenBegin': 0,
                      'tokenEnd': 1},
                     {'characterOffsetBegin': 196,
                      'characterOffsetEnd': 201,
                      'docTokenBegin': 33,
                      'docTokenEnd': 34,
                      'ner': 'TITLE',
                      'text': 'Judge',
                      'tokenBegin': 2,
                      'tokenEnd': 3},
                     {'characterOffsetBegin': 202,
                      'characterOffsetEnd': 221,
                      'docTokenBegin': 34,
```


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

```
                    'docTokenEnd': 37,
                    'ner': 'PERSON',
                    'text': 'Arenda Wright Allen',
                    'tokenBegin': 3,
                    'tokenEnd': 6},
                   {'characterOffsetBegin': 225,
                    'characterOffsetEnd': 232,
                    'docTokenBegin': 38,
                    'docTokenEnd': 39,
                    'ner': 'CITY',
                    'text': 'Norfolk',
                    'tokenBegin': 7,
                    'tokenEnd': 8},
                   {'characterOffsetBegin': 255,
                    'characterOffsetEnd': 285,
                    'docTokenBegin': 44,
                    'docTokenEnd': 48,
                    'ner': 'ORGANIZATION',
                    'text': 'Gloucester County School Board',
                    'tokenBegin': 13,
                    'tokenEnd': 17},
                   {'characterOffsetBegin': 331,
                    'characterOffsetEnd': 336,
                    'docTokenBegin': 56,
                    'docTokenEnd': 57,
                    'ner': 'PERSON',
                    'text': 'Grimm',
                    'tokenBegin': 25,
                    'tokenEnd': 26}],
 'index': 1,
 'kbp': [{'object': 'US',
          'objectSpan': [0, 1],
          'relation': 'per:countries_of_residence',
          'relationSpan': [-2, -1],
          'subject': 'Arenda Wright Allen',
          'subjectSpan': [3, 6]},
         {'object': 'Judge',
          'objectSpan': [2, 3],
          'relation': 'per:title',
          'relationSpan': [-2, -1],
```


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

```
                    'subject': 'Arenda Wright Allen',
                    'subjectSpan': [3, 6]}]}]
}
```


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

Produc'er's App - consumed/produced data format

Data Injector

Data consumed

Data consumed by this module depends on the source that it integrates.

This information is irrelevant for the integration of the module within the CPN platform.

Data produced

Data produced by this module is standardized regardless the source provider.

Below is the standard schema of the Producer's App articles, which could be extended.

```
articleSchema:
{
        _id : String, // Internal Id
        originId : String, // Original Id from source
        origin : String, // Source
        url : String //original item url
        category : String, // Category of article
        multimedia : [{ // list of media objects (images, video, etc..)
                type: String,
                url: String,
                name: String,
        }]
        title : String,
        text : String,   // Summary
        content: String, // Full text article
        language : String,
        author : String,
        date : Date, // date in Date format
        dateStr : String,    // date in String format
        timestamp : Number, // date in timestamp format
        location : { //location in geoJson format
                type: Object,
                index: '2dsphere',
                sparse: true },
        tags : [String] //list of tags,
        popular: Boolean,
```


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

```
}
```

Dashboard

Data consumed

Data consumed by the dashboard includes analytics about articles and user behaviour.

Data produced

This module provide data visualisation through a web UI. In addition, service that provide the data schemas could be provided.

Below the data formats:

```
User Actions Schema : {
            userId: String,
            date: Date, // date of the action
            itemId: Article ObjectId, // the reference article Id
            stream: String, // ['recommended', 'popular', 'latest']
            type: String, // ['read','like','interested','not interested']
}

App Visualisation Schema : { // How many times an article is shown on the client apps
        userId: String,
        date: Date, // date of the visualisation
        articleId: Article ObjectId, // the reference article Id
        stream: String, // ['recommended', 'popular', 'latest']'
}
```

```
Article Example :
{
        "_id" : ObjectId("5ba3986a7a43861e00c3553f"),
        "category" : "News",
        "date" : ISODate("2018-09-20T05:38:00Z"),
```


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

"text" : "The EU and the UK look no closer to a deal, as UK Prime Minister Theresa May has ruled out a second referendum and has urged the EU to \"evolve its position.\" A Brexit summit will be held in November, leaders announced.",
"url" : "http://www.dw.com/en/brexit-deadlock-theresa-may-urges-eu-to-compromise/a-45566322?maca=en-rss-en-top-1022-rdf",
"content" : "European leaders gathered in the Austrian city of Salzburg on Thursday for the second day of an informal summit. One of the main issues are the deadlocked negotiations on Brexit, the final agreement on which was supposed to be inked in October. On Thursday, however, EU leaders announced a special Brexit summit is to be held in November. Immigration, a major policy point for Austrian Chancellor Sebastian Kurz, whose government holds the rotating EU presidency, and security in the bloc are the other major talkingpoints. UK Prime Minister Theresa May told EU leaders on Wednesday that Britain has \"put forward serious and workable proposals\" and that it was now up to the EU to \"respond in kind\" and \"evolve its position.\" What the key players said at thesummit: Austrian Foreign Minister Karin Kneissl told DW that the EU member states disagreed on many things, there was \"no friction when it comes to Brexit. There is a high degree of cohesion.\" Irish Prime Minister Leo Varadkar: Ireland is a country that obviously wants to avoid a no-deal scenario...[but] we are preparing for that, we are hiring extra staff and officials, putting in IT systems, we're ready for that eventuality should it occur.\" European Commission President Jean-Claude Juncker: \"It was interesting, it was polite, it was not aggressive. She [May] is doing her job.\" EU Council President Donald Tusk: Despite progress in some areas, on the two thorniest issues of the Irish border and post-Brexit trade ties, \"the UK'sproposals will need to be reworked.\" He said a no-deal scenario was \"still quite possible.\" French President Emmanuel Macron: \"We [the EU] need to continue to act as a group, make sure we approach this as the EU-27. He stressed that a \"solution needs to be found, but it shouldn't jeopardize the coherence and the four freedoms of the single market.\" Brexit as it stands A major summit planned for October 18 was being treated as the last chance for a concrete deal for Brexit, which is supposed to go into effect on April 1. The November summit announced on Thursday is now seen as the deadline to reach a deal, which have to be ratified both by the EU legislature, all member states' Parliaments, and the UK Parliament. May has put paid to hints within her own government on a possible second referendum on Brexit, telling EU leaders that it is not an option. The two main sticking points

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

that remain unsolved are:how to regulate the border between Northern Ireland, which is part of theUK, and the Republic of Ireland, which is an independent EU member, as well as Britain's future trade negotiations with the EU. May told EU leaders that she won't accept an EU proposal for Northern Ireland to remain in the customs union while future trade ties are being negotiated. Migration to the EU EU leaders are also under pressure to come up with a compromise on immigration, after a joint summit in June produced a deal scant on details. While most countries agree on strengthening the border control agency Frontex, they still disagree on suggestions to redistribute refugees proportionally throughout the bloc. Countries like Poland, Czech Republic and Hungary do not look likely to change their hard-line stance against this plan. There are also disagreements across the bloc about which North African countries can be relied upon to set up schemes to stop migrants attempting the dangerous sea crossing to Europe, although many have praised Egypt for its efforts thus far. European Council President Donald Tusk accused member states of playing \"the migration blame game\" and urged them to create a bloc-wide solution to the issue. ng,es/sms (AFP, dpa, Reuters) Every evening at 1830 UTC, DW editors send out a selection of the day's hard news and quality feature journalism. You can sign up to receive it directly here.",

"title" : "Brexit deadlock: Theresa May urges EU to compromise",

"originId" : "popular_45566322",

"origin" : "DW",

"popular" : true,

"tags" : [ ],

"multimedia" : [

{

"type" : "Image",

"url" : "https://api.dw.com/image/45571119_301.jpg",

"name" : "Theresa May, Jean-Claude Juncker, Xavier Bettel (Reuters/L. Foeger)",

"_id" : ObjectId("5ba3986a7a43861e00c35543")

}

],

"__v" : 0

}

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

Uplifting Predictor - consumed/produced data format

Docker: imec/uplifting_predictor

Input DW Example :

Example of expected Input:

```
{
        _id : String, // Internal Id
        article: String, // Original text
        origin : String, // Source
        url : String //original item url
        category : String, // Category of article
        language : String,
        author : String,
        date : Date, // date in Date format
        dateStr : String,     // date in String format
        timestamp : Number, // date in timestamp format
        tags : [String] //list of tags
}
```

Data produced

Data produced by this module includes **per article**,

Example output from the uplifting predictor module:

```
{
        confidence: { //confidence of sentiment
                depressing: Float,
                neutral: Float,
                uplifting: Float },
        predicted: ['depressing', 'neutral', 'uplifting'], //list of tags
        language : String
}
```


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

Knowledge graph extraction (IDLab)

This is a Web API around the RMLMapper. You are able to execute RML rules on different data sources. Optionally, you can also request for the metadata that is generated during the execution of the rules.

1 Semantic Lifting

1.1 Data consumed

1.1.1 Rules

Content-type: text/turtle

Example:

```
@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix rml: <http://semweb.mmlab.be/ns/rml#>.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix ql: <http://semweb.mmlab.be/ns/ql#>.
@prefix map: <http://mapping.example.com/>.

map:map_person_0 rml:logicalSource map:source_0;
  a rr:TriplesMap;
  rdfs:label "person";
  rr:subjectMap map:s_0;
  rr:predicateObjectMap map:pom_0, map:pom_1.

map:om_0 a rr:ObjectMap;
  rr:constant "http://xmlns.com/foaf/0.1/Person";
  rr:termType rr:IRI. map:om_1 a rr:ObjectMap;
  rml:reference "firstname";     rr:termType rr:Literal.

map:pm_0 a rr:PredicateMap;
rr:constant rdf:type.

map:pm_1 a rr:PredicateMap;
rr:constant <http://example.com/name>.

map:pom_0 a rr:PredicateObjectMap;
```


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

```
  rr:predicateMap map:pm_0;
  rr:objectMap map:om_0.

map:pom_1 a rr:PredicateObjectMap;
  rr:predicateMap map:pm_1;
  rr:objectMap map:om_1.

map:s_0 a rr:SubjectMap;
  rr:template "http://example.com/{firstname}".

map:source_0 a rml:LogicalSource;
  rml:source "data.json";
  rml:iterator "$.persons[*]";
  rml:referenceFormulation ql:JSONPath.
```

1.1.2 Data

Content-type: application/json
Structure: JSON Object with attribute “sources” that describes the data.

Example:

```
{
  "sources": {
    "data.json": "{ \"persons\": [{\"firstname\": \"John\",\"lastname\":
\"Doe\"}, {\"firstname\": \"Jane\", \"lastname\": \"Smith\"}, {\"firstname\":
\"Sarah\", \"lastname\": \"Bladinck\"}] }"
  }
}
```

1.2 Data produced

1.2.1 Knowledge graph

See Section 2.2

1.2.2 Entity list

Content-type: application/json
Structure: array with terms (string)

Example:

```
[
```


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

```
  "Barack Obama",
  "United Stated of America"
]
```

2 RMLMapper Web API

2.1 Data consumed

```
{
  rml: string($text/turtle) // The RML rules used to extract the knowledge graph
                            // This parameter is required.
  sources: {}, // The sources used for the rules, where the key is the name of the
source
               // and the value the content of the source
               // default: {}
  generateMetadata: boolean // Metadata about the extraction process is generated
                            // when set to true
                            // default: false
}
```

Example:

```
{
  “rml”: "@prefix rr: <http://www.w3.org/ns/r2rml#>. @prefix rml:
<http://semweb.mmlab.be/ns/rml#>. @prefix rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>. @prefix rdfs:
<http://www.w3.org/2000/01/rdf-schema#>. @prefix ql: <http://semweb.mmlab.be/ns/ql#>.
@prefix map: <http://mapping.example.com/>.  map:map_person_0 rml:logicalSource
map:source_0;       a rr:TriplesMap;    rdfs:label \"person\";    rr:subjectMap
map:s_0;     rr:predicateObjectMap map:pom_0, map:pom_1. map:om_0 a rr:ObjectMap;
rr:constant \"http://xmlns.com/foaf/0.1/Person\";    rr:termType rr:IRI. map:om_1 a
rr:ObjectMap;       rml:reference \"firstname\";     rr:termType rr:Literal. map:pm_0
a rr:PredicateMap;  rr:constant rdf:type. map:pm_1 a rr:PredicateMap;    rr:constant
<http://example.com/name>. map:pom_0 a rr:PredicateObjectMap;     rr:predicateMap
map:pm_0;    rr:objectMap map:om_0. map:pom_1 a rr:PredicateObjectMap;
rr:predicateMap map:pm_1; rr:objectMap map:om_1. map:s_0 a rr:SubjectMap;
rr:template \"http://example.com/{firstname}\". map:source_0 a rml:LogicalSource;
rml:source \"data.json\"; rml:iterator \"$.persons[*]\";   rml:referenceFormulation
ql:JSONPath. ",
  “sources”: {
       "data.json": "{     \"persons\": [      {            \"firstname\": \"John\",
       \"lastname\": \"Doe\"     },          {            \"firstname\": \"Jane\",
       \"lastname\": \"Smith\"          },          {            \"firstname\":
\"Sarah\",          \"lastname\": \"Bladinck\"       }      ] }"
```


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

```
}
```

2.2 Data produced

```
{
  output: string($text/turtle) // The extracted knowledge graph in Turtle format.
  metadata: string($text/turtle) // The metadata of the extraction process.
}
```

Example:

```
{
  “output”: "<http://example.com/John>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://xmlns.com/foaf/0.1/Person>.\n<http://example.com/John>
<http://example.com/name> \"John\".\n<http://example.com/Jane>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://xmlns.com/foaf/0.1/Person>.\n<http://example.com/Jane>
<http://example.com/name> \"Jane\".\n<http://example.com/Sarah>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://xmlns.com/foaf/0.1/Person>.\n<http://example.com/Sarah>
<http://example.com/name> \"Sarah\".\n"
}
```


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

NewsFrames Module

Docker image: 109.232.32.194:5000/cpn-frames:v1.0-beta

Input data

To process a news article send a request to http://hostname:8080/process. The service expects two variables to be sent as form-data: text and format. Optionally an id variable can be specified for content identification purposes. Both HTTP GET and HTTP POST are allowed.

Variables:

- The `id` variable (optional) can contain any string value. It will be copied back in the output.
- The `text` variable must contain the content of the article. It is recommended to concatenate title and body together, separated by a newline character.
- The `format` variable specified the requested output format. Allowed values are: html, simplified-json and json (advanced output).

Output data

First an example of the simplified json format (`format=simplified-json`) is given. This format is recommended as tries to summarize the more exhaustive advanced output. The output structure returns a list of keywords, the main countries involved, the main topics (of the entities) and the main entity types.


```
{
   "keywords": [
       "US",
       "Donald Trump",
       "Ukraine",
       "Joe Biden"
   ],
   "locations": {
       "US": 12,
       "Ukraine": 8
   },
   "topics": {
       "media": 1,
       "politics": 14
```


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

```
  },
  "types": {
      "entity": 17,
      "gpe": 2,
      "gpe0": 2,
      "head_of_state": 1,
      "igo": 1,
      "location": 2,
      "minister": 1,
      "organization": 7,
      "party": 1,
      "person": 8,
      "politician": 6,
      "politics_institution": 4,
      "politics_per": 2,
      "role": 2,
      "so": 1,
      "time": 2,
      "value": 5
    }
}
```


Next we show the full output (`format=json`) for the same news article. The output contains a copy of the input data (id and content), but also a list of recognized mentions, a list of concepts and a list of relations. Each mention refers to the concept it is associated with through it's concept attribute (index in concept array). Concepts have the following attributes: text (the longest surface form), count (how many times they occur in the text), a list cpn types, a list of topics, slot types (if relevant), iptc codes (if relevant) and a list of wikidata instance types (Q codes). Q codes can be looked up on wikipedia and wikidata, they enable you to filter on very specific entity types: e.g. Q1520223 refers to a *constitutional republic*. Relations are expressed as subject, predicate, object triples. Both subject and object are indices into the concept array.


```
{
  "id": null,
  "content": "A top US diplomat has told an impeachment inquiry that he
followed President Donald Trump's orders to put pressure on Ukraine to
investigate his Democratic rival, Joe Biden.\r\n\r\nThe instruction came
from Mr Trump's personal lawyer, Rudy Giuliani, Ambassador Gordon Sondland
said.\r\n\r\nThe inquiry is assessing if Mr Trump withheld military aid to
```


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

```
Ukraine as a precondition. He denies any wrongdoing.\r\n\r\nIt is illegal
in the US to seek foreign help to gain electoral advantage.\r\n\r\nMr
Biden is one of the top contenders for the Democratic nomination for the
2020 presidential election.\r\n\r\nMr Sondland, the US ambassador to the
EU, told the latest hearing in the US House of Representatives that Mr
Giuliani had sought a public statement from Ukraine's leader, Volodymyr
Zelensky, announcing an inquiry into \"corruption issues\".\r\n\r\nMr
Giuliani specifically mentioned the company Burisma - which had the son of
Democratic presidential candidate Mr Biden, Hunter, as a board member -
and issues surrounding the 2016 US presidential election, he
said.\r\n\r\nSondland's testimony and reaction\r\nFive key moments from
impeachment hearing\r\nWho's who in Trump-Ukraine story?\r\nWhy Ukraine is
so important to the US\r\nIf found guilty in a majority vote in the House,
the Republican president will face an impeachment trial in the Senate. But
two-thirds of members of that Republican-controlled chamber would then
need to vote for Mr Trump to be removed from office.\r\n",
   "mentions": [
       {
       "begin": 6,
       "concept": 0,
       "end": 8,
       "text": "US"
       },
       {
       "begin": 67,
       "concept": 1,
       "end": 76,
       "text": "President"
       },
       ...
   ],
   "concepts": [
       {
       "text": "US",
       "count": 12,
       "type": [ "location", "entity", "gpe", "gpe0" ],
       "topic": [ "none" ],
```


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

```
      "slot": [ "keyword" ],
      "iptc": [],
      "auto": [ "Q1646605", "Q1520223", ... ],
      "tag": [ "type::location", "type::entity", "topic::none",
"type::gpe", "slot::keyword", "type::gpe0" ],
      },
      ...
  ],
  "relations": [
      [ 21, "member_of", 4 ],
      [ 15, "institution_of", 0 ],
      [ 16, "institution_of", 0 ],
      [ 23, "institution_of", 0 ]
  ]
}
```


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

APPENDIX 2 – HOWTO MEDIA COMPANIES INTEGRATION V2

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

HowTo - Media Company Integration

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

1. Step 1 - Articles collection

The CPN platform offers personalization and recommendation functionalities in as-a-service mode. In order to do this, it needs to analyze the content coming from media sources that want to access these functionalities.
The first step to follow for the integration of a media company within the platform is to provide access to its contents.

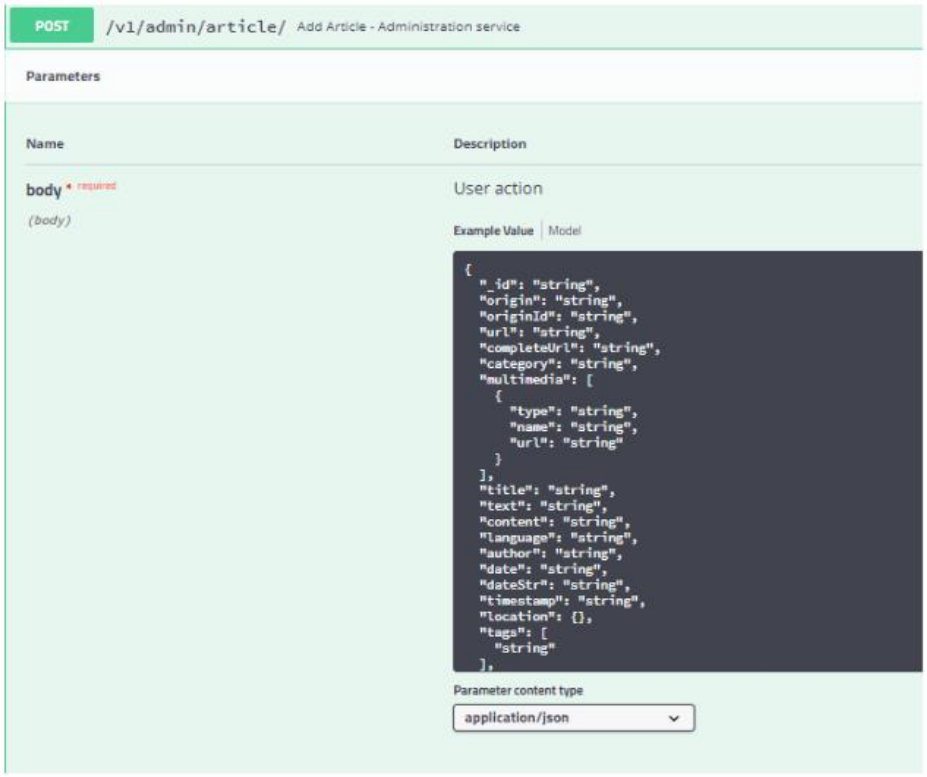
The CPN platform and in particular the Technology Brick Producer's App, takes care of integrating the contents within the platform, normalizing them and making them available to all the other components through the Apache Message broker. The media company can provide its own contents in two different ways:

1. **Providing a data repository (RSS or REST APIs)**
2. **Using CPN API Gateway**

In the first case, the CPN technical team must have access to article repository and takes care of mapping all the contents in the platform

In the second case, the media company itself can push its own articles exploiting the API, documented in the figure below:

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



At the moment the CPN platform provides, in addition to the recommended contents, also the most recent and the most popular contents. If you want to provide your users with these types of content, at least two data sources are required:

- **A list with all the articles** (Always mandatory)
- **A list with the most popular articles** (Optional if you want to provide more content lists)

The following input fields are mandatory or strongly recommended:

```
originId : String, // A unique ID for identify the article Mandatory
url : String //The article URL Mandatory
title : String,    Mandatory
text : String,   // Summary  Strongly recommended
```

Projectcpn.eu

3

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

```
content: String, // Full text article     Strongly recommended
date : Date, // date in Date format Mandatory
tags : [String] //list of tags, Strongly recommended
category : String, // Category or Topic of the article Strongly recommended
```

The output format of normalized articles will be the following:

2. Step 2 - Client-side Integration

The second step is focused on the integration of the CPN platform services within the client-side application of the media company.
The integration for the client applications is totally API based. This allows the integration of recommendation services in any type of client application. Now the CPN platform has been successfully tested with web, mobile and smart Tv apps, but can support any type of them.

This second step is divided into three parts:

- User identification
- Integration of recommendation services
- Integration of user profiling services

1.2 User Identification

A mandatory step of this process is the identification of the user within the CPN platform. In fact, all the personalisation services offered by CPN are strictly related to a specific user.

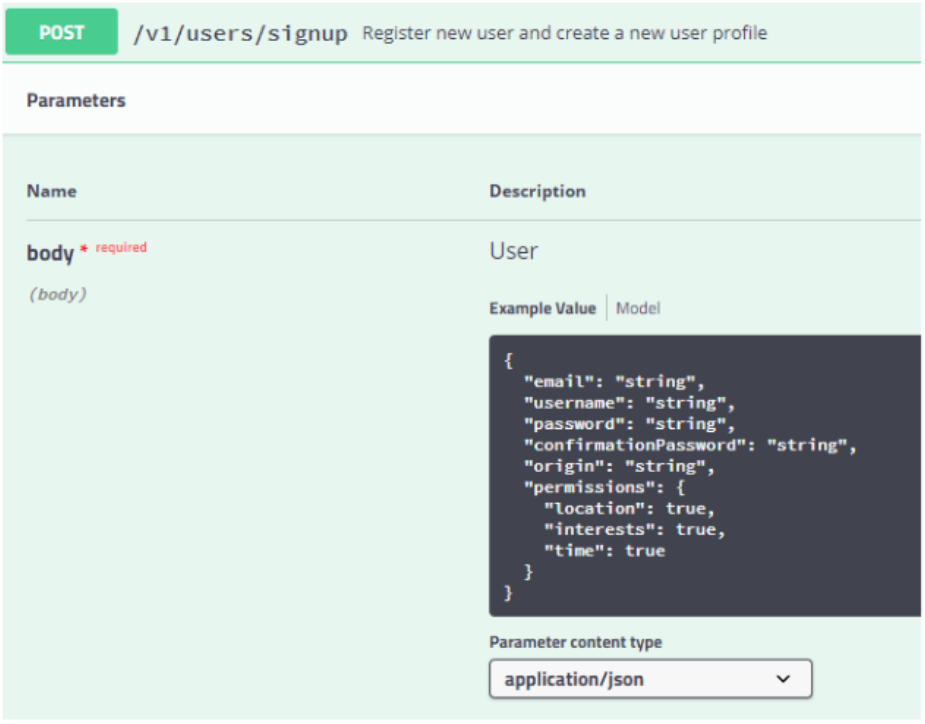
The CPN platform provide signup and login APIs for the registration of the user or, in alternative, we can provide a javascript tool for anonymous user identification based on browser-fingerprint.

The CPN platform provides two different APIs for Signup and Login:

Signup

It registers a specific user to the CPN platform. The specification are shown in the figure below:

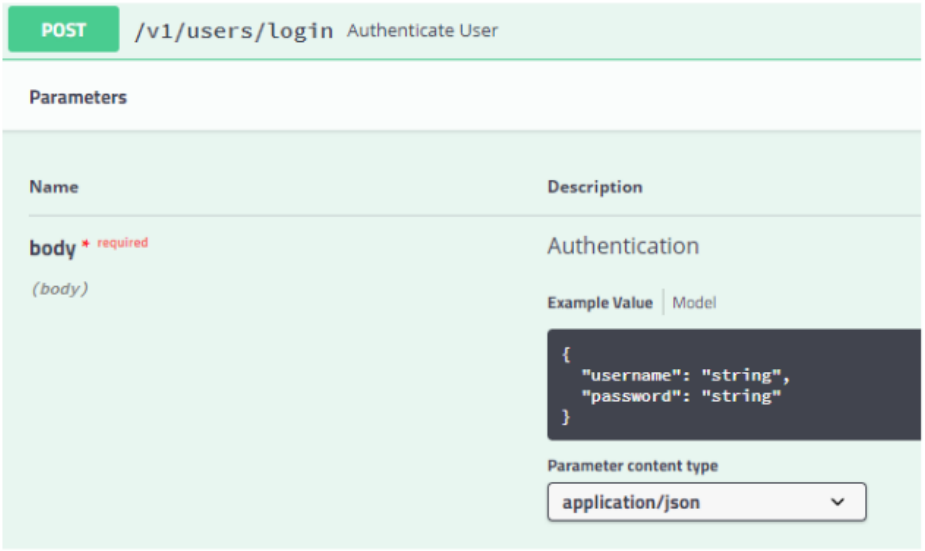
Co-funded by the Horizon 2020 Framework Programme of the European Union



Login

It identifies a specific user and returns a **JSON Web Token**. The **token is required** to access all the personalised services. The specifications are shown in the figure below:

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



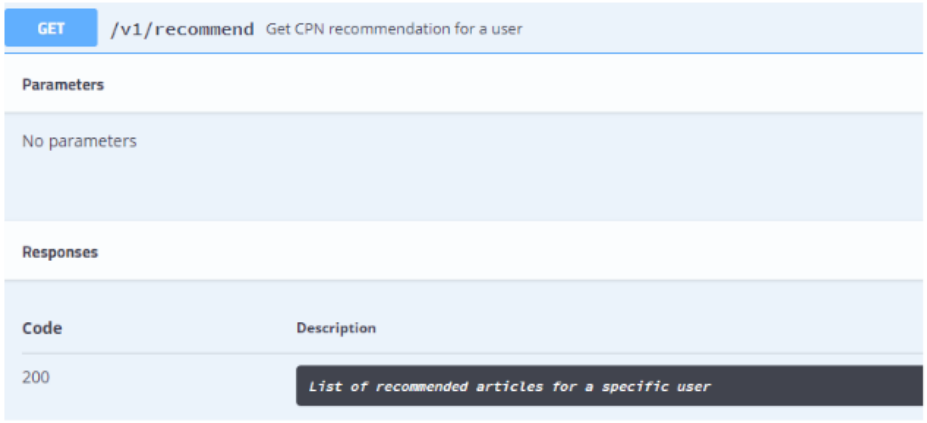
All the CPN APIs (apart the signup and login ones) needs a JWT for user identification. In order to access a protected API, the user agent should send the JWT in the Authorization header using the Bearer schema. The content of the header should look like the following:

```
Authorization: Bearer <token>
```

1.3 RECOMMENDATIONS

The CPN platform provide a protected API for recommendation service. This API returns a list of articles recommended for a specific user.

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



The API **doesn't need input parameter** but as mentioned before the **JWT is mandatory** in order to identify the user.

Below an example of a recommended item:

```
recommendedArticle:
{
        _id : String, // Internal Id
        originId : String, // Original Id from media source
        origin : String, // Source
        url : String //original item url
        completeUrl : String //complete original item url (in case of minified URL)
        category : String, // Category of article
        multimedia : [{ // list of media objects (images, video, etc..)
                            type: String,
                            url: String,
                            name: String,
        }]
        title : String,
        text : String,    // Summary
        content: String, // Full text article
        language : String,
        author : String,
        date : Date, // date in Date format
        dateStr : String,    // date in String format
```


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

```
        timestamp : Number, // date in timestamp format
        location : { //location in geoJson format
                type: Object,
                index: '2dsphere',
                sparse: true },
        tags : [String] //list of tags,
        popular: Boolean, // If this article become from popular feed,
        read: Boolean, // If the user already read the article
        score: Number, // The score of the recommendation of the article
        recommender_id: String, //The id of the recommender that produces this recommendation
        Description: String //Human readable description of the recommendation (for
        transparency)
}
```

1.4 User Profiling

In order to provide more precise personalized content for a specific user, the CPN platform needs to create a profile for each user and enrich these profiles while them using the CPN platform.

The users' behaviour is tracked by the CPN platform and now many "user actions" are collected:

1. "Interested"
2. "Not interested"
3. "Read"
4. "Rate"
5. "Share"
6. "Interest feedback"

1. "Interested" action

```
{
 "event": "Favorite item from stream: PERSONALIZED",
 "userId": "5b8fe613e7cffb000a11bafa",
 "itemId": "5c545800c13b2f1e00310b91",
 "timestamp": "1549443058160",
```


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

```
  "info": {
    "latitude": "38.0347891",
    "longitude": "23.7930435",
    "accuracy": "951"
    }
}
```


The /v1/add-action method will parse the event string and extract the action and stream. We could split the string to two fields (e.g. event: “Favorite item” and stream: “PERSONALIZED”) but currently this is how it works.

The location information inside the info section appear because this user enables the geolocation. Otherwise, the response would be empty:


```
{
  "event": "Favorite item from stream: PERSONALIZED",
  "userId": "5b8fe613e7cffb000a11bafa",
  "itemId": "5c545800c13b2f1e00310b91",
  "timestamp": "1549443058160",
  "info": { }
}
```


2. "Not interested"


```
{
  "event": "Remove item from stream: PERSONALIZED",
  "userId": "5b8fe613e7cffb000a11bafa",
  "itemId": "5c57439cc13b2f1e0032d88c",
  "timestamp": "1549444489459",
  "info": {
    "latitude": "38.0347891",
    "longitude": "23.7930435",
    "accuracy": "951"
  }
}
```


3. "Read"


```
{
```


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

```
 "event": "Read item from stream: PERSONALIZED",
 "userId": "5b8fe613e7cffb000a11bafa",
 "itemId": "5c574146c13b2f1e0032d712",
 "timestamp": "1549444639401",
 "info": {
  "readTimeMillis": "6235",
  "fullRead": true
 }
}
```


The info section contains the fields readTimeMillis, which corresponds to the (total) read time of this article in milliseconds and fullRead (boolean), if the user has scrolled to the bottom of the article.

4. "Rate"


```
{
 "event": "Favorite item from stream: PERSONALIZED",
 "userId": "5b8fe613e7cffb000a11bafa",
 "itemId": "5c55a26fc13b2f1e0031cf71",
 "timestamp": "1549448219381",
 "info": {
  "rating": 1,
  "ratingComment": ""
 }
}
```


The rating is an integer from 1 (Not at all relevant) to 5 (Very relevant) and rating comment is a string field (containing the user comments), as requested by VRT.

5. "Share"


```
{
 "event": "Share item to Twitter from stream: PERSONALIZED",
 "userId": "5b8fe613e7cffb000a11bafa",
 "itemId": "5c51ca0ec13b2f1e002f7443",
 "timestamp": "1549448912325",
 "info": {
  "latitude": "38.0347891",
```


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

```
    "longitude": "23.7930435",
    "accuracy": "951"
  }
}
```


The event can also be "Share item to Facebook", "Share item to LinkedIn", "Share item to Google+".

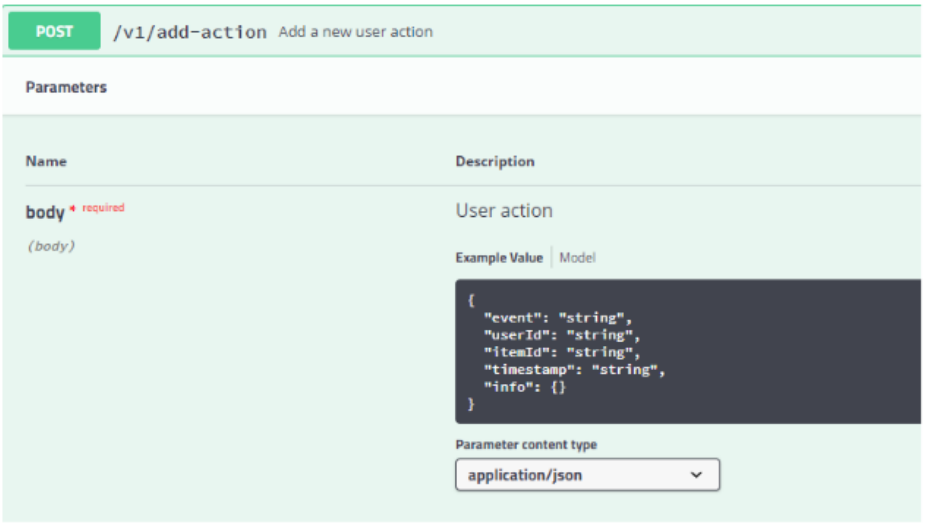
6. "Interest feedback"


```
{
  "event": "Interest feedback",
  "userId": "5b7ea4fce7cffb000a11baf8",
  "itemId": "5c3f0c64c13b2f1e00240fa5",
  "timestamp": "1549451022047",
  "info": {
    "interestFeedback": "No more football news!"
  }
}
```


The interestFeedback text contains the user-entered text. This type of action (feedback) is subject to change.

The specification are shown in the figure below:

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



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