D2.4 – Data – Management Plan v3

Deliverable No.	D2.4	Due Date	30-SEP-2021
Type	Report	Dissemination Level	*Public (PU)*
Version	0.9	Status	Internal Review
Description	This deliverable is the final version of the Data Management Plan, and relates how the FAIR principles have been implemented in PIXEL over all the data generated by the research of the project. This deliverable includes information required in the Guidelines for Data Management in Horizon 2020		
Work Package	WP2		

Authors

Name	Partner	e-mail
Carlos E. Palau (PC)	UPV	cpalau@dcom.upv.es
Ignacio Lacalle	UPV	iglaub@upv.es

History

Date	Version	Change
15-JUL-2021	0.1	Template from D2.2 and D3.3 customized. New ToC created.
21-SEP-2021	0.2	Completed version forwarded to WP leaders for internal review
24-SEP-2021	0.9	Version submitted to Internal Review
30-SEP-2021	1.0	Version to be submitted to EC

Key Data

Keywords	FAIR, data set, DOI, result, protection
Lead Editor	Ignacio Lacalle, P01 UPV

Abstract

This deliverable has been created in the context of the Work Package 2 (*Work plan, coordination and document management*) of the H2020-funded project PIXEL (Grant No. 769355).

In this document the Consortium related all the provisions made over the data generated during the project research (M1-M41). This document updates the previous DMP versions delivered in M6 and M18.

Deliverable D2.4 compiles the information of all datasets generated in PIXEL. Drawing from the already identified data sets exposed in D2.2 and D2.3, it summarises all the previous and fulfils enough information to consider the DMP compliance in PIXEL finalised. The document includes technical details of all data sets of the project, organised by category as outlined in the previous DMP versions.

This DMP complies with Horizon 2020 FAIR Data Management Plan (DMP) Template.

Disclaimer

Table of contents

List of acronyms

Acronym	Explanation
AI	Artificial Intelligence
CA	Consortium Agreement
DMP	Data Management Plan
DPO	Data Protection Officer
EC	European Commission
EM	Ethics Mentor
ETA / ETD	Estimated Time of Arrival / Departure
GA	Grant Agreement
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
PC	Project Coordinator
PO	Project Officer
POPD	Protection Of Personal Data
ORDP	Open Research Data Pilot

List of tables

1. About this document

Following Data Management Guidelines for H2020 projects:

Data Management Plan should be updated - if appropriate - during the project lifetime (in the form of deliverables). Please create and submit a new version of the DMP whenever significant changes arise in your project such as:

o *new data sets*

o *changes in consortium policies*

o *external factors.*

This deliverable aims at updating D2.3 (PIXEL intermediate DMP) following the previous instructions and other indicated recommendations and commitments from the official current DMP of the project. In addition, this iteration of the DMP happens to be the final deliverable of the WP and the task, which means that all provisions made over data generated in PIXEL with regards to Data Management Plan are include hereby.

1.1. Deliverable context

Table 1. Deliverable context

Keywords	Lead Editor
Objectives	N/A
Exploitable results	Though not directly related with any Exploitable Result, updating and complying with the Data Management Plan will allow the project to develop a useful and regulatory compliant solution.
	This deliverable is specifically related to the final exploitation report of the project (D9.8 – to be delivered at the same time), as many indications on IPR protection, open and/or private publication decisions have resulted from a direct interaction between both.
Work plan	This deliverable, even considered transversal, has been of application mostly in the preparation and execution of pilots and every task in PIXEL in which data were involved.
Milestones	N/A
Deliverables	This deliverable is closely related to D1.5. While this document (D2.4) is rather generic for PIXEL data outcomes, D1.5 is focused primarily in the compliance of Data Management mechanisms with the current legislation, procedures and EC recommendations in relation to personal data.
Risks	Even if not related directly to any of the Risks identified in the proposal, not following this Data Management Plan could imply certain IPR risks or further troubles related with the use and dissemination of data.

1.2. Methodology

In month M6, the first release of the PIXEL Data Management Plan (DMP) was delivered. It outlined the measures that the PIXEL project has put in place in order to accommodate the requirements set for projects complying with FAIR data management and contributing to the Horizon 2020 pilot action on open access to research data.

The plan considered the protection of personal data and business confidential information. It identified the requirements for accessing existing datasets that form the basis of the work of the project. Pertaining to the data that the project will produce, the DMP initially identified the types of data sets that will be outcomes of the project, namely:

- Project (public) deliverables.
- Scientific publications.
- Contributions to standards.
- Software and Applications.
- Data (often traces) collected for analysis and evaluation.

However, according to the DMP, these data may evolve during the project, e.g. anonymized data traces from the transport and logistics use case.

Additionally, also according to the DMP, the practical details on the management of the data sets were to be provided later on through deliverables D2.3 and D2.4 (this one). Besides, it was made clear that, during the project lifetime, additional information on several aspects might be elaborated for all data sets on a case by case basis before making a final consortium decision on handling the particular data generated or collected.

In M18, an update of the initial Data Management Plan took place. As expected, deliverable D2.3 went deeper in the definition of technical details of already identified data sets. The template proposed in D2.2 was used to complete the information related to a limited amount of data sets that were known and used in the project. Those were: (i) Project deliverables, (ii) Scientific Publications and (ii) the initial description of software applications.

Furthermore, the scope of D2.3 also included the identification of new potential data sets to be generated by the project. In that exercise, data sets like online questionnaires, raw data from observations, historical data for training AI models, data predicted, pictures of PIXEL members were identified. A complete list can be found in the version submitted (D2.3).

During M41, directly drawing from the previous release, a complete analysis of the whole DMP generated by PIXEL was performed. In addition, it was considered relevant to slightly enhance the template for data sets related to technical description to include extra information. Therefore, the scope of this deliverable has been two-fold:

- Add practical details of already known data sets to be generated by the project. This consisted in completing all DMP-related provisions over those data sets.
- Identify other data sets (not realised before) that have been generated by the project and complete the DMP technical details for all of them.

The original DMP was inherently compliant with the Horizon 2020 FAIR Data Management Plan (DMP) Template. It was stated that its future versions would as well be compliant with it, but some licenses have been taken for the writing of D2.4, similarly to was done for deliverable D2.3 (for the sake of readability):

- Section 2, Data Summary has been kept. In the initial DMP (D2.2) this section served for summarising project objectives, architecture, main outcomes and main data concerns. For the deliverable (D2.3), this section was used to introduce the main content needed: i) tables of detailed information on data sets, and ii) list of new identified data sets. In this deliverable (D2.4), section 2 has been moved to section 3 to be used to list and shortly summarise all the data sets (in a bulleted list style), that have subsequently been further developed (technical details) as Section 4 in this document.

- The rest of the sections of the FAIR DMP standard has been omitted since no change on the content needed to be clarified during these12 months.

2. Data Summary

After 41 months of execution of the project, all data sets generated have been identified. Prior to gather the technical details of those, here is the summary list:

- **Project deliverables** and **Other internal documentation.**
- **Scientific publications:** the scientific publications, mainly scientific papers, created by the consortium members, contain technical results from the PIXEL project.
- **Other publications and outputs:** Besides the scientific publications mentioned above, e.g., in journals or conference proceedings, the project has generated further publications and other project outcomes, such as: promotion material (brochures, flyers, posters, etc.), press releases and further project announcements, etc.
- **Contribution to standards:** standardisation activities to promote PIXEL's PEI as a potential standard. These actions also include the consideration of PIXEL data models as official FIWARE Data Models.
- **Software and applications:** the PIXEL project has developed and tested several applications. In addition to the source code and binaries, documentation of the developed applications, their specifications, and other related material have been made available through the project deliverables.
- **Online questionnaires for AB members:** sent to gather information about the advance of the project, feedback on the results and other collaborations that have taken place.
- **Online questionnaires for external people** *(new for D2.4)***:** that have been needed to complete the Budget Allocation for PEI weighting and also to gather information about potential PIXEL and PEI adoption in the context of task T5.4 and T9.4.
- **Data collected for analysis and evaluation:** results gathered from the online questionnaires above.
- **Literature review** *(new for D2.4)* on ICT and environmental fields for the development of the project. An initial literature review was performed in the first months of the project with the market analysis; in the last months future research lines were investigated and proposed as part of task T8.4.
- **Results generated from models and PEI calculation:** quantitative values of the PEI for different ports in different periods as well as the results of executing other models through the PIXEL platform (e.g., air pollution dispersion map).
- **Guidelines and recommendations related to the PEI adoption,** based on the current values of the environmental KPIs and aimed at optimising the value of the PEI for the port in future executions.
- **Historical data** collection and processing for model training (related to predictive algorithms).
- **Data predicted** generated after the execution of predictive algorithms.
- **Pilot deployments** in the four ports participating in PIXEL **(participation of port staff)** *(new for D2.4):* data sets generated after the interaction with port staff, either to collect information to feed the platform and to gather data about evaluation and assessment of the solution.
- **Data used in PIXEL pilot deployments** *(new for D2.4):* Data (in real time or not) that is used in the platform, either coming from NGSI agents or that is stored (as configurations, or result of the pilots) in the long-term storage database of the platform.
- **Pictures of PIXEL members presenting at events**

3. Enhancement of data sets' technical details template

The template for describing technical details of the data sets of PIXEL was designed in D2.2. It has been duly followed during all project execution, however it was deemed relevant to enhance its structure towards the final documentation of PIXEL DMP. The enhancements made are as follows:

- Explanation on how each data set is complying with FAIR principles.
- Indication of IPR or any additional protection.
- Access links (global or individual, per case).
- Relation with the commitment to ORDP.
- Metadata (if applying) and associated format.
- Other relevant comments.

4. Practical details of PIXEL data sets

This section aims at covering the committed statement: *During the project lifetime, additional information on the following aspects will be elaborated for all data sets on case by case basis before making consortium decision on handling of the particular data generated or collected*:

- Nature and scale of the data in consideration.
- To whom it could be useful, targeted audience and its size and level of interest.
- Information on the existence of similar data and possible synergies.
- Possibility for integration and reuse of the provided data by external users and researchers.
- Any further related issue.

These explanations aim at covering the Annex 1 of DMP template by the European Commission.

The internal management of the information and the data sets is being handled using the Project Management tools deployed for the project; however, because of the participation in ORDP, this procedure is ready to be accommodated to other different repositories.

With that aim, a template table has been used to embed the associated information in a handy and useful format.

4.1.1. Project deliverables

Table 2. Information of dataset: Project deliverables

Reference/name	Deliverables	Data sharing method		Public website (section deliverables)
Description	Public deliverables (according to GA) that have been generated during the execution of the project.	To whom it could be useful		Researchers, scientific community on the field
Type	Written documentation.	Targeted audience	Size	5-10 for official review by EC.
Related WPs and tasks	All WPs but WP1, which associated deliverables remain private.		Description	European Commission and technical reviewers
Format, standards	DX.Y – [Name]	Possibility of integration and reuse of this data by external users and researchers		Deliverables are not prone to integration but the information contained in them can be used for further researches (with the proper referencing and IPR compliance).
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or others.			
Estimated size	At the end of the project, the total size of public deliverables (aggregated) is: 250MB.	Possible synergies with similar data		N/A
Storage	1. In our secure repository OnlyOffice. 2. In a secure web server hosting the project's website (pixel-ports.eu)	Back-up		A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup).
Personal data included	Names of authors, Email of authors.	Means for personal data protection		Dealt with in CA.
ORDP	Officially, only D7.3. In reality, all of them are open as well.	IPR Owner / Data Owner		N/A
Link where it can be accessed	https://pixel-ports.eu/?page_id=30	Metadata	Yes/No	Yes
			Format	OpenAIRE and EC-Research
FAIR provisions	**Findable:** Specific naming has been used according to D2.2. Name "PIXEL" in the title. Structured upload. All deliverables have their assigned DOI. **Accessible:** Online on PIXEL's website. **Interoperable:** Documents uploaded and downloadable in PDF format.			

	Reusable: No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.
Other FINAL relevant information	PIXEL has finally made openly available a total of 43 deliverables, distributed as follows: • WP2: 9 deliverables • WP3: 4 deliverables • WP4: 4 deliverables • WP5: 4 deliverables • WP6: 5 deliverables • WP7: 3 deliverables • WP8: 5 deliverables • WP9: 6 deliverables All deliverables contain authors, history of changes, date of edition and contact details.

4.1.2. Other internal documentation

Table 3. Information of dataset: Other internal documentation

Reference/name	Other_internal_documentation	Data sharing method		Not to be shared publicly.
Description	Meetings minutes, agendas, internal documents of work, reporting docs…	**To whom it could be useful**		PIXEL partners.
Type	Written documentation.	**Targeted audience**	**Size**	70-80 (PIXEL partner members participating in the project).
Related WPs and tasks	All WPs		**Description**	People involved in the execution of the project.
Format, standards	*PPD_UC[No. use case]_Entity* and the rest of naming procedures detailed in deliverable D2.2	**Possibility of integration and reuse of this data by external users and researchers**		N/A
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or others			
Estimated size	Final size of internal repository (only documentation) is: 24GB.	**Possible synergies with similar data**		N/A
Storage	In our secure repository OnlyOffice.	**Back-up**		A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup).
Personal data included	Yes, names and emails of PIXEL partners people.	**Means for personal data protection**		Dealt with in CA:
ORDP	No.	**IPR Owner / Data Owner**		N/A.
Link where it can be accessed	No.	**Metadata**	**Yes/No**	No
			Format	N/A
FAIR provisions	N/A – Not public documents			
Other FINAL relevant information	N/A			

4.1.3.Scientific publications

Table 4. Information of dataset: Scientific publications

Reference/name	Scientific_publications	**Data sharing method**		Editorials/magazines/conferences proceedings RiuNet (UPV) Website (space "publications") Researchgate
Description	Scientific articles published at Scientific Journals, books or Conferences coming directly from technical work in PIXEL and authored by PIXEL participants	**To whom it could be useful**		Scientific Community
Type	Written documents.	**Targeted audience**	**Size**	-
Related WPs and tasks	WP4, WP5, WP6, WP7 and WP9		**Description**	Followers in ResearchGate, Current active researchers in the field of transport and IoT
Format, standards	Documents to be stored and shared in PDF. In the case of images, the default format should be svg if possible or png or jpg otherwise. Formats are depending on the journal/conference that they aim at being published. Namely, Springer LNCS or others.	**Possibility of integration and reuse of this data by external users and researchers**		Publications can be integrated in new papers and scientific articles with the common tools of automated/classic referencing. Furthermore, the information contained in them can be leveraged for SotA studies for further researches (with the proper referencing and IPR respect).
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or other			
Estimated size	Up to 6, 8, 10, 12 (or other) limit of pages set by the editorial. Average size ~4Mb per scientific article.	**Possible synergies with similar data**		Inner references in scientific articles can be used for cascade literature research.
Storage	To the accessible versions, in some articles we rely on the original publication website. For others, it is directly stored in the secure servers which host WordPress of PIXEL website. One copy of all articles is always stored in our secure repository. Own methods provided by the other data sharing means (check top right cell).	**Back-up**		Internal repository: A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup). Check policies of editors (IEEE, Springer, etc.) by other means.
Personal data included	Names, emails and affiliations of authors and co-authors.	**Means for personal data protection**		Informed consents and similar procedures established by the editorial prior to publication.
ORDP	Depending per article. Some of them are open access, other don't. Green access achieved via PIXEL.	**IPR Owner / Data Owner**		Depending per article.
Link where it can be accessed	Individually – one link per scientific contribution. Access point: https://pixel-ports.eu/?page_id=564	**Metadata**	**Yes/No**	No
			Format	No
FAIR provisions	**Findable:** All publications have DOI assigned. **Accessible:** Online on PIXEL's website and/or original Editor website (check table below). **Interoperable:** Documents uploaded and downloadable in PDF format. **Reusable:** Usual reusability of scientific papers.			
Other FINAL relevant information	A complete list of the scientific articles published can be found at deliverable D9.5 as well as in the reporting tool of the EC about Publications (Participant Portal – Sygma – Publications).			

4.1.4.Other publications and outputs:

Table 5. Information of dataset: Other publications and outputs

Reference/name	Other_material_dissemination	Data sharing method		Public website Social media channels Email to newsletter members
Description	Presentations of PIXEL used at conferences, press notes, posters, leaflets, supporting material for communication, newsletter.	To whom it could be useful		Followers, casual readers, PIXEL partners, event promoters, audience of certain events/fairs
Type	Images, presentations, designs, full-size materials, videos.	Targeted audience	Size	200/300 per year in conferences 8.000 in total through website 500 in total in social media
Related WPs and tasks	WP9		Description	Attendees of events, visitors of our website, followers in social networks
Format, standards	In the case of images, the default format should be svg if possible or png or jpg otherwise. The different files are expected to be below the 10 MB mark (M18) each. In the case of videos, YouTube standards..	Possibility of integration and reuse of this data by external users and researchers		This data is not prone to be integrated or reused by external users. It can be re-shared through social media by our followers. It can be used by event promoters to be included in publicity, promotion or event summaries
Software	These materials are produced by PIXEL partners using different software for edition: MS PowerPoint, MS Visio, draw.io, GIMP, Photoshop and/or others.			
Estimated size	Average size of each material is ~6 Mb. Total size expected is 120-150 MB (M18). The size of downloadable files has reached +120MB in M41 in total.	Possible synergies with similar data		N/A
Storage	In our secure repository OnlyOffice. YouTube.	Back-up		A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup).
Personal data included	No	Means for personal data protection		N/A
ORDP	Yes	IPR Owner / Data Owner		Public
Link where it can be accessed	https://pixel-ports.eu/ "PUBLICATIONS" (conferences, local media, marketing)	Metadata	Yes/No	No
			Format	No
FAIR provisions	**Findable:** All publications are under the same, easily identifiable space on PIXEL's official website. **Accessible:** All material includes a link to access the source of the note/article document, hosted either on PIXEL's servers or directly from the original source. Link to videos point to PIXEL's YouTube channel **Interoperable:** Documents uploaded and downloadable in PDF format. Logo uploaded in SVG format. **Reusable:** IPR must be respected, but no special license applied to these kinds of results.			
Other FINAL relevant information	The final distribution of additional material that has been made public via PIXEL website is: - Presentations of 44 conferences (where partners actively disseminated PIXEL). The website also includes PIXEL's webinars in the 'Conference' section 10 press notes about PIXEL 2 posters and a Final Closure related poster 2 versions of PIXEL leaflet 1 logo			

- 1 roll-up
- 5 newsletters
- 38 videos uploaded to YouTube channel

4.1.5. Contribution to standards

At the end of the project, several trials to contribute to standards have been performed. While details on the actions taken are included in deliverable D9.5, the summary is the following:

- Actions with WPSP: resulting datasets: teleconferences and a couple of press notes on their site.
- Actions with AIVP: resulting datasets: teleconferences and a couple of press notes on their site.
- Actions with ESPO: resulting datasets: teleconferences and a couple of press notes on their site.
- Actions with FIWARE: resulting datasets: DataModels (format JSON)

Contacts with relevant entities and research of the current solutions and standardisation have been established but with no final delivery of a standard coming out of PIXEL. This must not be considered a shortcoming of the project but a rather expected final status for a RIA project. More details about this can be found at D9.5.

Only the interactions that have actually generated data sets are reported below.

Table 6. Information of dataset: Contribution to standards

Reference/name	FIWARE_PIXEL_datamodels	Data sharing method		GitHub
Description	FIWARE has confirmed interest on including PIXEL-custom data models (based on NGSI) in their brochure of recommended data formats as soon as those will be delivered under the NGSI-LD paradigm.	To whom it could be useful		Open source community, FIWARE community, IoT developers, SMEs in the realm of IoT, European Commission as FIWARE is considered a founding block of CEF.
Type	Software/data semantics	Targeted audience	Size	-
Related WPs and tasks	WP6, WP7		Description	Same as above.
Format, standards	JSON (examples and schema)	Possibility of integration and reuse of this data by external users and researchers		JSON is considered a de-facto standard in the software development. It is very used for HTTP REST APIs and it is widely adopted in the IoT realm as the main data piece format to exchange.
Software	Any software able to read JSON (e.g., notepad, development IDEs) will be able to inspect the data models.			
Estimated size	1 confirmed data models. 3 data models will also be requested (whenever translated to LD-semantics).	Possible synergies with similar data		FIWARE Data Models (current) format (schema, example) has been followed. Any project/Action compliant with those will be ready to adopt PIXEL data models.
Storage	PIXEL's own Git repository, PIXEL's GitHub account and a copy in the secure OnlyOffice repository hosted by Project Coordination.	Back-up		Periodic backup on OnlyOffice.
Personal data included	No	Means for personal data protection		N/A
ORDP	Yes	IPR Owner / Data Owner		Public.
Link where it can be accessed	On-going	Metadata	Yes/No	Yes
			Format	OpenAIRE, EC-Research
FAIR provisions	**Findable:** Data models will have a DOI assigned. A structured naming following FIWARE conventions are followed. **Accessible:** Online on PIXEL's GitHub account.			

	Interoperable: Following JSON format and data models conventions.
	Reusable: Total reusability, downloadable and prepared to be used in IoT-FIWARE compliant deployments.
Other FINAL relevant information	The Data Models that has been agreed to be part of FIWARE DataModels is VesselCall.
	Other three are under consideration and will be explored (via PIXEL Association) as long as they will be converted to the LD-semantics format.
	More details can be found in deliverable D9.5.

4.1.6.Software and applications

Table 7. Information of dataset: Software and applications

Reference/name	Software_components	Data sharing method	GitHub	
Description	Software modules developed in the project that can be made public.	**To whom it could be useful**	EC, developers, open source community	
Type	Software code with associated methodology	**Targeted audience**	**Size** | -	
Related WPs and tasks	WP4, WP5, WP6 and WP7		**Description** | EC, developers, open source community	
Format, standards	Containerisation practices have been followed, especially for all NGSI agents and all models. Common languages and libraries have been used. License Apache 2.0 is most predominantly applying, while others do exist for specific, protected results.	**Possibility of integration and reuse of this data by external users and researchers**	All components have been created in a modular way including an interface for integration (API or others). All developments have been made public with associated markdown-type documentation (Readthedocs documentation and individual README.md)	
Software	The software modules are being generated after PIXEL partners' coding activities. The IDEs for development, frameworks and languages are diverse, such as Python, Java, Vue.js, IDE Eclipse, IDE VS Code… among others. Required hardware for deploying PIXEL: at least 2 VMs with 4 cores, 16GB RAM and 300-500GB HD each.			
Estimated size	The size of the software is about 15 GB, mainly because they are provided as Docker instances and it increases the final size. Just considering the code in Github, without any Docker image, the size could be probably reduced up to a 60%.	**Possible synergies with similar data**	Some synergies have been identified: 1) Federation and SSO: if the target port includes some delegation support (e.g. SAML), there is no need to even store user profiles in the PIXEL platform 2) Storage: if the target port is also using Elasticsearch as database, both can be probably merged to reduce infrastructure costs	
Storage and backup	The use of Docker instances allows to just focus on the data for back-up purposes. The IH allows to create clusters redundancy for scaling and safety purposes. It is also possible to dump/export data from Elasticsearch or other databases (E.g. Mongo)			
Personal data included	Apart from user profiles, no personal data is stored in the system. Enough means have been put in place.	**Means for personal data protection**	The architecture includes a security module based on FIWARE security components (Wilma, Keyrock) to add OAuth2-based authentication and authorization security to PIXEL services and applications. No personal data is stored in plain in the system.	

ORDP	Yes	IPR Owner / Data Owner	Diverse. Although the main PIXEL code for the skeleton of the infrastructure has Apache2.0 license. Other modules are differently protected according to the partner holding the IPR. Deliverable D9.8 should be consulted to analyse such ownership and licensing.
Link where it can be accessed	Code at https://github.com/pixel-ports Doc at https://pixel-ports.readthedocs.io/en/latest/ Docker instances at https://hub.docker.com/orgs/pixelh2020/repositories	Metadata Yes/No	No
FAIR provisions	**Findable:** Source code and Docker instances are currently findable through GitHub and DockerHub search engines. **Accessible:** Online on PIXEL's GitHub, readthedocs and DockerHub. **Interoperable:** Following Docker instances and OT Framework specification. **Reusable:** Total reusability via Docker instances and total adaptable by downloading source code.		
Other FINAL relevant information	• The PIXEL main GitHub repo has 22 repositories • The PIXEL DockerHub main repo includes 55 Docker individual repositories • ReadtheDocs includes 8 different documented projects for the different components of PIXEL		

4.1.7. Literature review on ICT and environmental fields for the development of the solution

Table 8. Information of dataset: Literature review on ICT and environmental fields

Reference/name	Literature_review	Data sharing method		Public deliverables PIXEL website
Description	At the beginning of the project, several documents and data sets (graphs, statistics) were generated about the state of the art and literature review on those fields.	To whom it could be useful		Researchers Marketing departments Port Innovation and Development departments Public community
Type	Documentation	Targeted audience	Size	-
Related WPs and tasks	WP3, WP4; WP5; WP8		Description	Research community, maritime ports sector, SMEs…
Format, standards	PDF documents.	Possibility of integration and reuse of this data by external users and researchers		The state of the art studies are integrated in PIXEL deliverables. Part of them have been used to build some PIXEL papers on. Those will be integrable to the extent that PIXEL deliverables containing them will be.
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or other			
Estimated size	Over 100 pages.	Possible synergies with similar data		N/A
Storage	Deliverables containing the state-of-the-art studies are stored in secure web servers hosted by the Project Coordinator.	Back-up		A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup).
Personal data included	No	Means for personal data protection		N/A
ORDP	Yes/No	IPR Owner / Data Owner		N/A
Link where it can be accessed	See below (Other FINAL relevant information)	Metadata	Yes/No	No
			Format	No

FAIR provisions	**Findable:** Specific naming being followed according to D2.2. Name "PIXEL" in the title. Structured upload.
	Accessible: Online on PIXEL's website.
	Interoperable: Documents uploaded and downloadable in PDF format.
	Reusable: No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.
Other FINAL relevant information	The pieces of the literature review have been embedded within some PIXEL deliverables, in particular (and the link where they can be found): • Analysis of competitors, trending business models, market as a whole (D3.1) • Analysis of current environmental impact monitoring systems/methods (D5.1) • Analysis of mathematical approach and methods towards composite indicators (D5.2) • Analysis of current models and approaches for estimating air pollution (D4.1 and D4.2) • Analysis of current models and approaches for estimating noise pollution (D4.1 and D4.2) • Analysis of current approaches for water and soil pollution (D4.2) • Analysis of current approaches for estimating ETA and ETD in maritime ports (D4.3) • Analysis of predictive models over satellite imagery (D4.3) • Analysis of temporal series regression and other methods for predicting traffic congestion (D4.3). • Future Research Directions analysing past, present and future research areas and topics (D8.5)

4.1.8.Online questionnaires for AB members and external people

Table 9. Information of dataset: Questionnaires

Reference/name	AB_questionnaires	Data sharing method		PIXEL website
Description	Questionnaires created to let AB members provide feedback on PIXEL results. Questionnaires created to gather knowledge on various aspects addressed to people external to the project. Procedure: generation of an online form (using Google forms) with enough ethical/data privacy disclaimers, request of information and gathering of data that will be further used.	To whom it could be useful		PIXEL partners
Type	Documents, fillable forms	Targeted audience	Size	~70 people
Related WPs and tasks	WP3, WP5, WP9		Description	PIXEL Partners
Format, standards	.docx	Possibility of integration and reuse of this data by external users and researchers		These questionnaires can be re-used as long as any reference to any responsibility from PIXEL is removed.
Software	Google forms was used to circulate the forms, MS Word for editing and creating the original questionnaires. Any text processor software will open them for edition.			
Estimated size	7 editable documents up to 2 Mb.	Possible synergies with similar data		The questionnaires can be used to extend previous researches on the field as well as to serve as basis for similar studies in the future
Storage	Stored in the secure web server hosting PIXEL's WordPress website.	Back-up		Backup is not performed for the questionnaires apart from its storage in OO (automatic backup).
Personal data included	No. All questionnaires have been anonymised prior to be included in the deliverables.	Means for personal data protection		Anonymisation.

ORDP	Yes	IPR Owner / Data Owner		Public.
Link where it can be accessed	See below.	Metadata	Yes/No	No
			Format	No
FAIR provisions	**Findable:** Questionnaires are uploaded to open repositories – PIXEL website. **Accessible:** Online on PIXEL's website. **Interoperable:** Documents uploaded and downloadable in .docx format. **Reusable:** No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.			

Other FINAL relevant information	The questionnaires that have been generated for AB members are:
	
	   o Feedback from PIXEL (link to be avalaible during October 2021)
	   o Questionnaire for Budget Allocation method (link to be avalaible during October 2021)
	
	The questionnaires that have been gathered for external people are:
	
	   o Environmental performance metrics used by TEN-T ports (link)
	   o Exploitation survey for providers (link to be avalaible during October 2021)
	   o Exploitation survey for users (link to be avalaible during October 2021)
	   o Value survey for maritime stakeholders (link to be avalaible during October 2021)
	   o Questionnaire for Budget Allocation Method (link to be avalaible during October 2021)			

4.1.9. Data collected for analysis and evaluation:

The questionnaires indicated in 4.1.9 drove PIXEL partners to generate a knowledge database that was used to deliver documentation, software and final results of the project. Those questionnaires led to graphs, statistics, conclusions, reports that have been made available to the community via PIXEL deliverables.

Table 10. Information of dataset: Results of questionnaires, statistics for project usage

Reference/name	Statistics_questionnaires	Data sharing method		PIXEL deliverables (except for D9.8)
Description	Statistics, graphs, conclusions, reports or simply plain results of the filling of questionnaires from people answering abovementioned questionnaires.	To whom it could be useful		PIXEL partners, researchers of the field, consulting firms, SMEs in environmental for maritime forms, European Commission.
Type	Documentation	Targeted audience	Size	-
Related WPs and tasks	WP5, WP9		Description	Same as above
Format, standards	*DX.Y – [Name]*	Possibility of integration and reuse of this data by external users and researchers		Deliverables are not prone to integration but the information contained in them can be used for further researches (with the proper referencing and IPR compliance).
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or others.			
Estimated size	Up to 25 pages.	Possible synergies with similar data		The results can be used to extend previous researches on the field as well as to serve as basis for similar studies in the future.
Storage	Deliverables including those results are stored both in PIXEL's secure repository online and on website hosting server.	Back-up		Periodic backup of deliverable files (see before).
Personal data included	No. All questionnaire results have been anonymised prior to be included in the deliverables.	Means for personal data protection		Anonymisation.
ORDP	Yes	IPR Owner / Data Owner		Partner P0X YYY

Link where it can be accessed	See deliverables below	Metadata	Yes/No	Yes
			Format	OpenAIRE, EC-Research

FAIR provisions	**Findable:** Specific naming being followed according to D2.2. Name "PIXEL" in the title. Structured upload. All deliverables have their assigned DOI. **Accessible:** Online on PIXEL's website. **Interoperable:** Documents uploaded and downloadable in PDF format. **Reusable:** No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.

Other FINAL relevant information	The results of the previous questionnaires are included in diverse formats (graphs, statistics, reports, conclusions) in the following deliverables:
	
	 o Budget Allocation method results: D5.4
	
	 o Results of Environmental performance metrics used by TEN-T ports: D5.4
	
	 o The following results are available **but are confidential (per Grant Agreement)**. Those might be provided under request:
	
	      o Exploitation survey for providers: D9.8
	
	      o Exploitation survey for users: D9.8
	
	      o Value survey for maritime stakeholders: D9.8
	
	      o Questionnaire for Budget Allocation Method: D9.8

4.1.10. Guidelines and recommendations related with the PEI adoption

Table 11. Information of dataset: Project deliverables

Reference/name	PEI_recommendations	Data sharing method	Public website (section deliverables) and private (recommendations per port)
Description	Recommendations to the ports to optimise their PEI (closer to 1), indicating suggested actions to improve the number.	To whom it could be useful	Maritime ports, environmental departments of logistic nodes, researchers on the ports environment field.
Type	• Documentation • Software (rules engine)	Targeted audience — Size	• N/A • 3-4 people per port usually.
Related WPs and tasks	WP5, WP7 (task T7.5)	Targeted audience — Description	Port Authorities, Terminal Operators, Policy Makers, Industry of maritime ports, SMEs on Energy, SMEs on Logistics.
Format, standards	• Reports in PDF. • j-easy rules[1]	Possibility of integration and reuse of this data by external users and researchers	Recommendations published can be re-used and converted into any format to feed rules engines according to the exposed in deliverable D5.4.
Software	PDF-enabled reader, j-easy library and Java-based rules engine.		
Estimated size	200 rules. Up to 5 pages of recommendations.	Possible synergies with similar data	The recommendations of PIXEL can be added to other databases of the sector or public whitepapers aiming at improving ports' green performance.
Storage	Reports in PDF: secure web server where PIXEL's website WordPress is hosted.	Back-up	OnlyOffice server has weekly backup versions of the reports and the rules. Rules in PIXEL deployments in ports are not backed up.

---

[1] https://github.com/j-easy/easy-rules

	Rules and engine: within the long term storage database of each PIXEL's deployment (currently, in 4 ports).			
Personal data included	No	Means for personal data protection		No
ORDP	Yes and No.	IPR Owner / Data Owner		Public (and associated pilot port in each case)
Link where it can be accessed	D5.4	Metadata	Yes/No	No
			Format	No
FAIR provisions	**Findable:** Specific naming being followed according to D2.2. Name "PIXEL" in the title. Structured upload (when proceeding). **Accessible:** Online on PIXEL's website (when proceeding). **Interoperable:** Documents uploaded and downloadable in PDF format. **Reusable:** Only the public part of recommendations: no specific license included. Usable by the community including proper references. Will be kept online on the website for 5 years after the end of the project.			
Other FINAL relevant information	These data sets have been generated by PIXEL three-fold: • As global recommendations drawing from partners' research: in deliverable D5.4 – Public. • Set of recommendations (baseline) and set of rules (including a software rule engine) – Not public. • Individually to each port as an auto-generated report drawing from the results of eKPI and PEI calculations for a specific period – Not public.			

4.1.11. Historical data collection and processing for model training

A variety of **historic data has been needed to train the predictive algorithms** that have been developed in PIXEL and integrated in the pilots:

Table 12. Information of dataset: Historical data collected for model training

Reference/name	historic_dataset	Data sharing method		GitHub (when applying)
Description	Data sets of historic data (previous to the moment where the task was executed in PIXEL) to train ML/AI models that have been used in PIXEL pilot deployments.	To whom it could be useful		Researchers in the fields of maritime ports
Type	Software	Targeted audience	Size	-
Related WPs and tasks	WP4, WP7		Description	SMEs, European Commission, IT departments of ports.
Format, standards	.CSV or JSON	Possibility of integration and reuse of this data by external users and researchers		Format of the historic data (whenever made available) is explained in README.md files in the repository where those are stored and can be accessed and re-used accordingly (as long as LICENSE and NOTICE disclaimers are respected).
Software	The data sets have been used by different models in PIXEL to train ML models (normally, in format HDF5 or as Jupyter Notebooks - .ipynb) that have been integrated within models' source code (Python, Java or .NET).			
Estimated size	Publicly available, 3MB.	Possible synergies with similar data		
Storage	Data is stored in specific folders at OnlyOffice and on the official GitHub repository of the project.	Back-up		Usual backup procedure in OO and relying on GitHub servers for public availability.
Personal data included	No	Means for personal data protection		N/A

ORDP	No	IPR Owner / Data Owner		LICENSE and NOTICE files accordingly (whenever public)
Link where it can be accessed	See below	Metadata	Yes/No	No
			Format	No
Other FINAL relevant information	For predicting renewable energy – energy consumption balance: o Historic of data in region of GPMB with determined azimut based on PVPROD by EC and PVoutput.org – aggregated historic publicly offered by PIXEL – Accesibility link. For training ETA and ETD models: o AIS data captured from AIS receptor and subscription to AIShub – privately owned data. o Vessel calls information of GPMB - private due to data protection concerns - a sample and statistics are uploaded here For training traffic models: o Historic of traffic at certain points (gates, points in a map, number of vehicles, speed) o ASPM not public, THPA public here, PPA not public o OpenWeatherMap historic information (public) - https://openweathermap.org/api o Vessel calls information of (particular per each port) – historic has not been made public. o DarkSky API – public - https://darksky.net/dev o TrafficThess historic information – (public) -(traffic in Thessaloniki, interesting at 5 nearest roads to the port) - https://www.trafficthess.imet.gr/exporter.aspx			

4.1.12. Data used in PIXEL pilot deployments

A variety of **(close-to) real time data sets** are necessary to produce tailored information which has been used in WP7 for pilots' execution.

Table 13. Information of dataset: Close-to real time data used in pilots

Reference/name	real_time_data	Data sharing method		Not to be shared publicly.
Description	Data that is currently being used to feed the functionality of PIXEL models and predictive algorithms.	To whom it could be useful		PIXEL partners, (those public to) researchers, EC-funded projects, other maritime ports
Type	Software-like files containing data	Targeted audience	Size	-
Related WPs and tasks	WP7		Description	Same as above.
Format, standards	Heterogeneous. In general, JSON files but also HTML tables, .CSV files or other messaging format.	Possibility of integration and reuse of this data by external users and researchers		All data accessible follows a specified format that can be consulted accessing to the entrypoint. Connection to those data should always be noticed, notified and properly referenced.
Software	Compatible software to receive and process each type of data. In general, all programming languages can use libraries to manage such information.			
Estimated size	N/A	Possible synergies with similar data		
Storage	N/A	Back-up		N/A
Personal data included	No	Means for personal data protection		N/A
ORDP	No	IPR Owner / Data Owner		Ports of PIXEL.
Link where it can be accessed	See below (if applying).	Metadata	Yes/No	No
			Format	No

| Other FINAL relevant information | The data is not shared publicly due to privacy concerns expressed from the stakeholders. Raw data from own sensors – globally not available publicly, generated (associated to each port) by the sensors acquired in the project:

    o Noise, Light and odour
    o CO_2, NOx and SOx
    o Weather (temperature, wind speed, humidity, dew point).
        o In the case of THPA, the data is public:

Observational data from own web services (APIs, PCSs, TOSs, etc.) - detailed technical table below.

    o Vessel calls information (particular per each port) – historic has not been made public.
        o APT – public – gathered from: http://www.porto.monfalcone.gorizia.it/sailinglist.asp
        o THPA – public – gathered from: https://download.thpa.gr/pixel/ (might not be online after the finalisation of PIXEL).
        o GPMB and PPA – private due to data protection concerns.
    o Waste, wastewater data – private due to data protection concerns
    o PAS model's definition of parameters – kept internal within ports as it may disclose information on the inner (protected) functioning of port terminal and machinery data
    o Data from SILI, trucks in APT and SDAG - kept internal within ports as it may disclose information on the inner (protected) functioning of the port and the logistics node

Observational data from external services (weather, etc.)

    o OpenWeatherMap - https://openweathermap.org/api
    o HERE (Traffic at specific points) - https://developer.here.com/
    o DarkSky API – public - https://darksky.net/dev
    o TrafficThess information – (public) -(traffic in Thessaloniki, interesting at 5 nearest roads to the port) - https://www.trafficthess.imet.gr/ |

4.1.13. Results generated from models, predictive algorithms and PEI calculation

Table 14. Information of dataset: Project deliverables

Reference/name	Model_execution_results	Data sharing method		Not to be shared publicly.
Description	The outputs of the models once they are executed using the data provided by the ports. All the process is performed within PIXEL infrastructure deployed in ports.	To whom it could be useful		PIXEL ports.
Type	Software and (less amount) reports.	Targeted audience	Size	4-5 people per port. Currently, a total of 16-20 people.
Related WPs and tasks	WP7		Description	Environmental department of ports, Port Authority management, Terminal Operators in ports.
Format, standards	JSON	Possibility of integration and reuse of this data by external users and researchers		This data is not shared with external researchers due to protection (IPR and ownership) from the stakeholders which execute the software on their premises.

Software	For visualising the results, two ways are possible:
	
	    o Via the Dashboard of PIXEL (Views -> Custom table, or others).
	    o Via Kibana (connected to PIXEL's Information Hub) as part of the ELK stack.			

Estimated size	Varying per port and depending on the periodicity of execution/scheduling of models	Possible synergies with similar data		N/A

	in the port. It is estimated that (in average) each model execution generates 20Kb of data.		
Storage	Results are stored in the Information Hub of PIXEL deployment in each case.	**Back-up**	PIXEL's Information Hub is not backed up automatically although it offers the possibility to the ports.
Personal data included	No.	**Means for personal data protection**	N/A
ORDP	Yes/No	**IPR Owner / Data Owner**	Partner P13 GPMB, P16 APT, P10 THPA and P11 PPA correspondingly. Each port owns the results of the models applied over their provided data.
Link where it can be accessed	N/A	**Metadata** **Yes/No**	No
		Format	No

Other FINAL relevant information	The following data sets are **generated** from the **models** applied to the ports: o **GPMB:** o Energy (PAS line): vessel name, vessel identifier, time of start of work (operation), time of end of work (operation), energy expected to be consumed to operate the vessel. o Renewable balance: photovoltaic energy estimated for one specific period of time, energy to be consumed by the port in the same period of time. o Energy (Gantt): vessel name and identifier, time of start and end of work (operation) o PEI (software): eKPIs about air emissions, waste, waste water, noise and light on the specific period, result of sub-indices, result of the PEI (0 to 1) and result of reliability rating per source and in total. o PEI (report on guidelines and recommendation) – data set described in 4.1.10. o Predictive algorithm; ETA and ETD (drawing from PAS and AIS): vessel name, vessel identifier, vessel arrival time, estimated vessel departure time, actual vessel departure time. o Predictive algorithm: PV production: energy generated, potential PV renewable energy to be generated. o **APT:** o Energy (PAS line): vessel name, vessel identifier, time of start of work (operation), time of end of work (operation), energy expected to be consumed to operate the vessel. o COVID-19 pilot model: vessel name, vessel identifier, areas occupancy associated to the operation of those vessels, evolution in time of occupancy, evolution in time of energy to be consumed in such operations. o Hinterland multimodal model: o Renewable balance: photovoltaic energy estimated for one specific period of time, energy to be consumed by the port in the same period of time. o Energy (Gantt): vessel name and identifier, time of start and end of work (operation). o PEI (software): eKPIs about air emissions, waste, waste water, noise and light on the specific period, result of sub-indices, result of the PEI (0 to 1) and result of reliability rating per source and in total. o PEI (report on guidelines and recommendation) – data set described in 4.1.10. o Predictive algorithm: Traffic prediction: estimated traffic at diverse points of the port and in the nearby roads. o **THPA:** o Air pollution: map with points indicating the dispersion and concentration of CO_2, NO_x, SO. o Noise pollution: heatmap representing noise pollution dispersion in the map of the port. o COVID-19 pilot model: vessel name, vessel identifier, areas occupancy associated to the operation of those vessels, evolution in time of occupancy, evolution in time of energy to be consumed in such operations. o Energy (Gantt): vessel name and identifier, time of start and end of work (operation) o PEI (software): eKPIs about air emissions, waste, waste water, noise and light on the specific period, result of sub-indices, result of the PEI (0 to 1) and result of reliability rating per source and in total. o PEI (report on guidelines and recommendation) – data set described in 4.1.10. o Predictive algorithm: Traffic prediction: estimated traffic at two gates of the port (10A and 16).

	o **PPA:**
	o Air pollution: map with points indicating the dispersion and concentration of CO_2, NOx, SO.
	o Noise pollution: heatmap representing noise pollution dispersion in the map of the port.
	o Energy (Gantt): vessel name and identifier, time of start and end of work (operation)
	o PEI (software): eKPIs about air emissions, waste, waste water, noise and light on the specific period, result of sub-indices, result of the PEI (0 to 1) and result of reliability rating per source and in total.
	o PEI (report on guidelines and recommendation) – data set described in 4.1.10.
	o Predictive algorithm: Traffic prediction: estimated traffic at several points in Piraeus.

Although the result datasets are owned by ports and will be kept within their PIXEL deployments, a summary of those results, the conclusions obtained and the technical and business evaluation are available at project deliverables.

The latter is part of the ORDP and here are some relevant details:

Findable: Specific naming being followed according to D2.2. Deliverables where the results are inserted do have a DOI.

Accessible: Results of usage and execution of the models can be found in deliverables D7.2 and D7.3.

Interoperable: PDF format.

Reusable:. No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.

4.1.14. Participation of port staff in pilot deployments

Table 15. Information of dataset: Project deliverables

Reference/name	participation_staff	Data sharing method		PIXEL deliverables
Description	Feedback on the usage of the platform, studies, conclusions, statistics or full reports.	To whom it could be useful		Maritime ports, PIXEL partners, European Commission.
Type	Documentation	Targeted audience	Size	~50 people initially.
Related WPs and tasks	WP7, WP8		Description	Same as above.
Format, standards	*DX.Y – [Name]*	Possibility of integration and reuse of this data by external users and researchers		Deliverables are not prone to integration but the information contained in them can be used for further researches (with the proper referencing and IPR compliance).
Software	Generated after the writing of PIXEL partners. Software used can be text editors: MS Word, OnlyOffice or others.			
Estimated size	D7.2, D7.3 and D8.3 sum up to 10 Mb.	Possible synergies with similar data		N/A
Storage	1. In our secure repository OnlyOffice. 2. In a secure web server in UPV premises hosting the WordPress of pixel-ports.eu	Back-up		A mirror server for backup. Backup is done every day (delta changes) and each Sunday (full backup).
Personal data included	Names of authors, name of staff in partners that have participated in the evaluation (when applying), Email of authors.	Means for personal data protection		Dealt with in CA.
ORDP	Yes	IPR Owner / Data Owner		N/A
Link where it can be accessed	https://pixel-ports.eu/?page_id=30	Metadata	Yes/No	Yes
			Format	OpenAIRE

FAIR provisions	**Findable:** Specific naming being followed according to D2.2. Name "PIXEL" in the title. Structured upload. All deliverables have their assigned DOI.
	Accessible: Online on PIXEL's website.
	Interoperable: Documents uploaded and downloadable in PDF format.
	Reusable: No specific license included. Usable by the community including proper references. Will be kept online in the website for 5 years after the end of the project.

Other FINAL relevant information	Participation of port staff:
	
	   o Usage of the platform: D7.2 and D7.3
	   o Evaluation of the platform: D8.3

4.1.15. Pictures of PIXEL members presenting at events:

Table 16. Information of dataset: Project deliverables

Reference/name	PIXEL_members_pics	Data sharing method		Social media Presentations on website
Description	Photos taken of members of PIXEL partners in meetings, fairs, congresses and other events where PIXEL is actively disseminated.	To whom it could be useful		Social media followers of PIXEL, PIXEL partners, maritime ports community.
Type	Multimedia material	Targeted audience	Size	-
Related WPs and tasks	WP2, WP9		Description	PIXEL partners, Research entities in Europe.
Format, standards	.jpg, .png, .jpeg, .bmp.	Possibility of integration and reuse of this data by external users and researchers		This data is not prone to be reused as it consists in images of people attending events in the context of PIXEL. No research value associated.
Software	To use the images, any image-opener software is valid. No special software used for its generation.			
Estimated size	2 Gbs approx.	Possible synergies with similar data		N/A
Storage	The source images are stored in the OnlyOffice repository in a secure server managed by Project Coordination.	Back-up		OnlyOffice is backed up weekly since the beginning of the project. For older backups, a monthly or yearly backups are available.

Personal data included	Yes:
	
	  • Names of partners' members.
	  • Photos / pics of those members.
	  • (sometimes) Email addresses of members.
		Means for personal data protection		Dealt with in CA.
				
				In case new members or additional people were included, written consent was requested and obtained prior to the publication.

ORDP	No	IPR Owner / Data Owner		Individual protection of anyone's own images.
Link where it can be accessed	N/A	Metadata	Yes/No	No
			Format	No

Other FINAL relevant information	The images have been published (enough consent provided) via:
	
	  • PIXEL Twitter account
	  • PIXEL LinkedIn account
	  • PIXEL website
	  • Own partners' websites accordingly			