AMBER

www.amber.international

Funded by the Horizon 2020 Framework Programme of the European Union

D6.3 Data Management Plan

This document is version 3.0 of a deliverable of the AMBER project, which has received funding from the European Union's Horizon 2020 Programme under Grant Agreement (GA) # 689682.
10 May 2018

History of changes

This document is version 3.0 and has been updated and modified following EC comments and the separation of D6.1 Project Management and Data Management into the separate D6.1 Project Management and D6.3 Data Management Plan. The presentation of data has been restructured (**Table 2**), data security issues have been addressed (Section 5), and data have now been linked to dissemination outputs (**Tables 5 and 6**).

DISCLAIMER

The opinion stated in this report reflects the opinion of the authors and not the opinion of the European Commission.

Executive Summary

This data management plan (D6.3) is one of three plans (D6.4, D6.2) which detail the data types, storage and how data are handled including data security. It follows the Horizon 2020 FAIR (Finable, Accessible, Interoperable, Reusable) Data Management Plan (DMP). AMBER produces diverse and large data sets related to four main outputs: (A) Atlas Data; (B) Tool data; (C) Case Study Data; (D) Dissemination data and metadata. The data outputs tie into dissemination (D5.6 Plan for Exploitation and Dissemination), with the links being illustrated at the end of this document. Datasets related to publications are also made public through open access and various data repositories.

Authors

Dr. Ian Dodkins, Prof. Carlos Garcia de Leaniz, Prof. Sonia Consuegra and Prof. Luca Börger, for the AMBER consortium.

Contents

1 INTRODUCTION

The Horizon2020 FAIR Data Management Plan (DMP) template is used in this report. 'FAIR' stands for: findable, accessible, interoperable and re-useable.

FAIR guidance
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Basis of the creation of the FAIR principles can be found here (Nature publication):
https://www.nature.com/articles/sdata201618

Three versions of the data management plan will be submitted (**Table 1**), but the third will be incorporated into deliverable D6.2 "Final Report with Legacy plan for updating and maintaining Barrier Atlas and other AMBER digital outputs" which was specifically created to ensure post project data management.

Table 1. Updates on the AMBER Data Management Plan.

Deliverable	Title	Submission Date (Month)	Content
D6.3	Data Management Plan v1, v2.	30 November 2016 (M6)	Initial data management plan outlining the intended approach for the project
D6.3 Updated	Data Management Plan v3	31 May 2018 (M24)	Update of the data management plan with modifications indicated by the EC and other changes generated by the project
Incorporated into D6.2	Final Report with Legacy plan for updating and maintaining Barrier Atlas and other AMBER digital outputs	31 May 2020 (M48)	Final data management plan, also covering post-AMBER legacy plan

2 DATA SUMMARY

2.1 Purpose of the data

In terms of data management, AMBER main data outputs include:

1. A pan-European Atlas of river barriers
2. A decision support tool for planning, removal and mitigation of barriers (dams, culverts, weirs) in European Rivers

Data collation (gathering pre-existing data) and data collection (new data obtained through the actions of the AMBER project) are important in producing databases for both of these objectives. The Barrier Atlas will also be used as a basis for creating other important pan-European data resources (maps) within AMBER, such as predicted fish community maps. The decision tool is comprised of individual tools that each contribute to the barrier planning/removal/mitigation and monitoring processes and have their associated data sets from tool development. Case Study data sets will result from testing of the tools, and finally there will be useful data resulting from dissemination activities and metadata associated with the project.

Table 2. Relationship between data collated/collected and objectives of the project

A: Barrier Atlas data and associated maps:

1. Barrier Atlas
   (a) Collation of currently available data held by regional and national authorities on barriers to produce the barrier base map
   (b) Data collected by the AMBER consortium to validate the Barrier Atlas
   (c) Citizen Science (public) data collected on barriers using a smart phone app to supplement the available barrier data
2. Fish community map
3. Atlantic salmon status map
4. Barrier impacts on river ecology map

B: Specific tools used to comprise the overall decision tool:

Monitoring:

1. eDNA tool for ecological monitoring
2. Rapid habitat assessment tool using drones

Barrier Passability:

3. Barrier Passability Tool: Fish and other aquatic organisms responses to barriers and hydrodynamics
4. Model of organism passability vs. hydropower generation
5. Model of fish movement through river networks

Conflict resolution:

6. Cost-benefit analysis of river infrastructure tool
7. Barrier management scenario tool (habitat stress days)
8. Ecosystem services evaluation tool
9. Social attitudes tool for conflict resolution

C: Case study data; outputs of testing the tools and mitigation techniques

D: Dissemination data and project metadata

The Barrier Atlas itself is the first pan-European barrier map and will have applications in scientific research, barrier planning, and policy making. The tools, both individually and the final decision tool, will have use within industry (hydropower), policy decisions, catchment management, regional planning, national planning, and also within scientific research.

Being 'adaptive' management, it is important that future researchers also have access to the data used to initially create the tools such that they can be iteratively improved. Consequently, the overall decision tool can improved, as scientific understanding progresses.
Additional data associated with dissemination activities and metadata relating to these data sets will also be created.

The following data descriptions list whether the data being used is pre-existing to AMBER i.e. fully (yes), partially (some) or completely collected within the scope of the AMBER project (no). It also details the origin of the data being used by AMBER, then the type of data output from the study (variables), the format of the output data, how the output can be further utilised, and which organisational bodies are likely to use that output.

Within AMBER there are Case Studies (WP4) which are used to test the various barrier management tools which AMBER produces. There are also Case Study sites (WP4) which are chosen to assess the tools in general, a separate 'test catchment' in Germany (River Neckar) has been selected for in-depth studies on specific socio-economic aspects (i.e. Ecosystem services and barrier cost evaluation).

What follows is a data summary, but more detailed data outputs associated with the specific tasks are listed in Appendix 1 (**Table 6** and **Table 7**).

2.2 Summary

A: Barrier Atlas data and associated maps (T1.2.1; D1.2; T1.2.2; D1.3)
BARRIER ATLAS

A1a.Collated barrier data

Data contact	POLIMI (SB)
Existing data?	Yes
Data origin	Collated barrier data from regional and national authorities throughout Europe
Data type	All available data, but focusing on: Source ID of barrier; url of data source; country; latitude; longitude; river name; basin name; barrier height; barrier type; year constructed; fishpass (y/n)
Data format	Original spreadsheet data, processed into databases and GIS themes: .xls .csv .mxd .shp .dbf
Expected size	10 GB
Data utility	Along with the Citizen Science and validation data, will create the pan-European Barrier Atlas.
Data Users	Will be used by the public; hydropower companies; educational establishments; scientists; municipalities; water authorities; NGOs and policy makers.

The Atlas data will comprise of stream barrier location (Latitude; Longitude) and all other available information that is stored on barriers from regional and national authorities within all 31 European Economic Area (EEA) countries, as well as some Balkan countries (Albania, Bosnia and Herzegovena, Macedonia, Montenegro, Serbia). It will also include islands within these countries, e.g. Azores. This is data collation i.e. gathering pre-existing data. Deliverable D1.2 (Country specific reports containing the metadata) provides more details on the data being collected.

A1b. Atlas validation data

Data contact	POLIMI (SB)
Existing data?	No
Data origin	Collated barrier data from regional and national authorities throughout Europe
Data type	ID of barrier; photo; latitude; longitude; date recorded; barrier type; barrier height; extends across entire watercourse (y/n); in use (y/n); altitude; slope; river type; sinuosity; local land use
Data format	Original spreadsheet data, processed into databases and GIS themes: .xls .csv .mxd .shp .dbf
Expected size	5 GB
Data utility	Along with the Collated barrier data and Citizen Science data, will create the pan-European Barrier Atlas.
Data Users	Will be used by the public; hydropower companies; educational establishments; scientists; municipalities; water authorities; NGOs and policy makers.

Collated regional and national authority data will vary in types of barriers surveyed by different authorities and the minimum height surveyed as well as the survey methods. To allow comparability between Member States and to estimate the numbers of barriers of types of heights not monitored, AMBER consortium members will do a validation exercise. This will be a field exercise whereby selected locations are surveyed for all types and heights of barriers. Comparison to the collated data set for that region will allow upscaling of the data to provide better estimates of total barrier numbers, and estimated barrier numbers of each type, for Member States and across Europe. It will

also allow fair comparisons between regions which have been surveyed by authorities using different survey methods.

A1c. Citizen Science Data

Data contact WFMF (JD)
Existing data? No
Data origin Citizen Science: the European public will record barrier data using 'barrier tracker' app.
Data type ID of barrier; photo; latitude; longitude; date recorded; barrier type; barrier height; extends across entire watercourse (y/n); in use (y/n)
Data format Original spreadsheet data, processed into databases and GIS themes:
.xls .csv .mxd .shp .dbf
Expected size 10 GB
Data utility Along with the Collated barrier data and Citizen Science data, will create the pan-European Barrier Atlas.
Data Users Will be used by the public; hydropower companies; educational establishments; scientists; municipalities; water authorities; NGOs and policy makers.

Citizen Science (CS) data will be from the 'barrier tracker' app developed for AMBER. The app consists of tier 1 and tier 2 expertise levels. The majority of users will use the simple tier 1 app. Additional data fields to be used in tier 2, by expert users, is still being decided. A 3rd party contractor, 'Natural Apptitude' developed the app and will collect the data on their servers before sending it to beneficiary 19-JRC. Validation of the data (checking images, checking against other records) will be done before being utilised within the Barrier Atlas.

A2. Fish Community Map (T2.2.1)

Data contact SSIFI (PP)
Existing data? Yes (plus model outputs)
Data origin Pan-European fisheries and habitat data will be collated.
Data type Fish species; abundance; age category; river type; channel sediment type; channel width; channel depth discharge/flow rate
Data format Original spreadsheet data, processed into databases and GIS themes:
.xls .csv .mxd .shp .dbf
Expected size 10 GB
Data utility Assessing the effect of barriers on ecological habitats and fish communities; input into predictive planning.
Data Users Will be used by the hydropower companies; scientists; municipalities; NGOs

The map of fish communities in different water bodies will be created. This will be based on determining the ecological fish habitats in water bodies, taking into account barriers and the hydrologic regimes. Habitat models already developed by SSIFI and ERCE will be used to delineate these fish habitats. The fish communities will also be compared with expected reference conditions, and Restoration Alternative Analysis to examine the change in habitat structure and the change in 'habitat stress days'. This data will allow assessments of the available and optimal options for stream restoration.

A3. Atlantic salmon status map (T4.2.1)

Data contact SOTON (PK)
Existing data? Some (plus model outputs)
Data origin The AMBER Barrier Atlas; Barrier impacts on river ecology; national juvenile salmon stock assessments (from regional authorities)

Data type	Spreadsheet of predicted salmon stocks; map of Atlantic salmon status
Data format	Spreadsheet data and GIS themes: .xls .csv .mxd .shp .dbf
Expected size	1 GB
Data utility	Targeting specific barriers in Europe which could be removed to improve salmon stocks and socio-economics; rehabilitation schemes; strategic planning
Data Users	Will be used by: educational establishments; scientists; municipalities; water authorities; NGOs and policy makers.

Utilising the 'Barrier Impacts on River Ecology' output and the Barrier Atlas a pan-European, river by river assessment of the status of Atlantic salmon will be done; examining the effects of barriers on salmon communities. The model will be validated with data from national juvenile salmon stock assessments. This will include an assessment of Atlantic Salmon river habitats lost, at different spatial scales. This data will also be used to select barriers whose removal would most benefit Salmon populations and socio-economic return.

A4. Barrier impacts on river ecology map (T2.1)

Data contact	SU (LB)
Existing data?	Yes (plus model outputs)
Data origin	Water Framework Directive (WFD) fish/invertebrate/plant/phytoplankton data; Barrier Atlas
Data type	species and abundance data: invertebrates; fish; macrophytes; phytoplankton
Data format	Original spreadsheet data, processed into databases and GIS themes: .xls .csv .mxd .shp .dbf
Expected size	50 GB
Data utility	Determining how connectivity is affecting ecological status as defined within the WFD; targeting river restoration schemes i.e. seeing if connectivity likely to be a problem
Data Users	Will be used by scientists; municipalities; water authorities; NGOs and policy makers

Existing raw stream survey data of ecological assemblages (aquatic plants, benthic macroinvertebrates and fishes) will be collated from national WFD databases, into a single database. A novel modelling approach 'PREDICTS' will be used to examine the effects of barriers on ecological assemblages at a pan-European scale.

B. Specific tools used to comprise the overall decision tool
MONITORING

B1. eDNA tool

Data contact	UNIOVI (EGV)
Existing data?	Some
Data origin	Available primers; lab testing of primers; field testing of eDNA methods for detecting species
Data type	specific primers configurations; methodologies
Data format	Spreadsheet; publications .xls .csv .doc .pdf
Expected size	1 GB
Data utility	Will enable users to monitor species using eDNA and to make assessments on barrier effects based on eDNA differences up/down stream.
Data Users	Will be used by: hydropower companies; scientists; water authorities; NGOs

Environmental DNA (eDNA) is increasingly being used to do rapid detection of the presence of a suite of different species. A water sample provides DNA sequences from multiple species which can be analysed for presence/absence simultaneously. This method is still in development but testing above and below barriers is an excellent method to refine this technique. The tool will be useful for monitoring the effects of barriers on species passability.

B2. Rapid habitat assessment tool

Data contact	DU (PC)
Existing data?	No
Data origin	Photo and video footage from drone flights along river corridors done by AMBER consortium members. Development of a rapid habitat assessment methodology.
Data type	video; photo; report
Data format	video; photo; report .mov .avi .mp4 .jpg .doc .pdf
Expected size	20 GB
Data utility	Users will be able to assess river habitats rapidly using drone technology. This will have particular application for assessing hydromorphological change due to barriers.
Data Users	Will be used by: hydropower companies; scientists; water authorities; NGOs

BARRIER PASSABILITY

B3. Barrier Passability Tool

Data contact	SOTON (PK)
Existing data?	Some
Data origin	Published data on ability of aquatic organisms to pass different barriers types based on barrier heights, barrier structure and hydrodynamic conditions; experimental data in flumes on ability of weak swimmers to navigate different hydrodynamic conditions.
Data type	passability values for species based on variables such as water depth required to jump; swim velocity; jump height; behavioural responses etc. (TBC)
Data format	Spreadsheet data; report .xls .csv .doc .pdf
Expected size	1 GB
Data utility	Will enable barrier design and mitigation techniques to be optimized for different species; can be used to predict ecological effects of barrier construction; can be used for modelling ecological effects of barriers at a strategic (national, pan-European) scale.
Data Users	Will be used by the public; hydropower companies; educational establishments; scientists; municipalities; water authorities; NGOs and policy makers.

B4. Model of organism passability vs. hydropower generation

Data contact	SOTON (PK)
Existing data?	Some
Data origin	Data from the passability tool and from information relating to hydropower generation and flow (and seasonal migration patterns). Existing data comes from technical details relating to hydropower production and published information used in developing barrier passability tool
Data type	Passability values under different flow velocities and barrier heights; hydropower commitments (licensing) and flow-power generation relationships; technical hydraulic data relating to barrier types and mitigation types; temporal migration/movement patterns of different organisms (TBC).
Data format	Spreadsheet data .xls .csv
Expected size	1 GB
Data utility	Will assist in optimizing the management or mitigation strategies of individual barriers as well as feeding into strategic regional decision making.
Data Users	Will be used by hydropower companies; scientists; municipalities; water authorities; NGOs

Data from the passability tool and from information relating to hydropower generation and flow (and seasonal migration patterns) will be incorporated into a tool which can balance decisions on hydropower generation against the ability of different organisms to navigate different barriers under different flow/seasonal regimes. This will be validated in a specific field test catchment in Germany (Rivers Nehe or Neckar) where beneficiary 16-IBK have significant knowledge (T3.1.1).

B5. Model of fish movement through river networks (T3.2.3)

Data contact	SOTON (JK)
Existing data?	Some
Data origin	Data from the passability tool and from information relating to hydropower generation and flow (and seasonal migration patterns)
Data type	Passability values under different flow velocities and barrier heights; hydropower commitments (licensing) and flow-power generation relationships; technical hydraulic data relating to barrier types and mitigation types; temporal migration/movement patterns of different organisms (TBC).
Data format	Spreadsheet data .xls .csv
Expected size	1 GB
Data utility	Will assist in optimizing the management or mitigation strategies of individual barriers as well as feeding into strategic regional decision making.
Data Users	Will be used by hydropower companies; scientists; municipalities; water authorities; NGOs

The behavioural response of organisms to barriers and flow velocities will be modelled using an Agent Based Model (ABM). Data on swimming behaviour will also be obtained from experimental lab work done by AMBER in Swansea and Southampton. Information from other data sources created in AMBER will be used (Barrier Atlas; Barrier Passability Tool). Existing data has been used for both producing the Barrier Atlas and Barrier Passability datasets.

CONFLICT RESOLUTION

B6. Cost-benefit analysis of river infrastructure tool (T3.3.1)

Data Contact	DU (ML)
Existing data?	Some (within “Model of organism passability vs. hydropower generation” data)
Data origin	AMBER field studies in tests catchment and “Model of organism passability vs. hydropower generation” data
Data type	Hydrological variables: head difference; stream geometry; flow rate. Costings for different constructions of different dam and barrier types. Costing estimates for ecosystem services and economic benefits of barriers (e.g. of power production).
Data format	Spreadsheet data .xls .csv
Expected size	2 GB
Data utility	Will feed into decision tool and assist in strategic planning of barrier feasibility and location and provide objective information as a basis for stakeholder conflict resolution.
Data Users	hydropower companies; local government; municipalities; water authorities; NGOs

The field tests catchment in Germany (see “Model of organism passability vs. hydropower generation” data) will also undergo a comprehensive economic valuation of the effects of stream barriers on riverine goods and services.

B7. Barrier management scenario tool (D2.6, T2.2.1. T2.3)

Data contact	SSIFI (PP)
Existing data?	Some (WFD databases, fisheries and hydrological data)
Data origin	'Fish Community Map' (Fisheries and hydrological data) from above WFD data bases EC stream flow and climate records https://www.eea.europa.eu/data-and-maps/indicators/river-flow-3
Data type	Loss of habitat; change in habitat structure; change in number of habitat stress days; RAA diagrams
Data format	Spreadsheet data and diagrams (in reports) .xls .csv .doc .pdf
Expected size	2 GB
Data utility	Will feed into decision tool and assist in strategic planning of barrier feasibility and location and provide objective information as a basis for stakeholder conflict resolution.
Data Users	hydropower companies; scientists; water authorities; policy makers; NGOs

Using previous AMBER 'Fish Community Map' data (based on fisheries and hydrological data) the fish guilds and habitats will be assessed for deviation from expected reference conditions (WFD databases) within representative rivers of the EU. Restoration Alternatives Analysis (RAA), based on the MesoHABSIM model, will be used to assess loss of habitat, change of habitat structure and increase in the number of habitat stress days for different barrier management scenarios (planning, removal and various forms of mitigation).

Habitat deficit, change of habitat structure and habitat stress days will also be calculated for barriers under different climate change scenarios, using a model based on EC stream flow and climate data. RAA diagrams will also be produced.

B8. Ecosystem Services Evaluation Tool (T2.6)

Data contact	ERCE (KK)
Existing data?	No (except some use of barrier atlas)
Data origin	Field studies in German test catchment; barrier atlas data; 'Cost-benefit analysis of river infrastructure tool'; additional cost-benefit valuations relating to ESS
Data type	from test catchment: categorization of ESS; cost-benefit valuations; diagrams and spreadsheet models of links between stakeholders
Data format	Spreadsheets and report .xls .csv .doc .pdf
Expected size	1 GB
Data utility	
Data Users	

Ecosystem Services (ESS) that rivers provide and the users of these services will be identified and defined, and the effects of barriers on delivery. ESS delivery rate will be determined in selected Case Studies (WP4). Interactions between stakeholders and how construction and removal of barriers affects and change in river status redistributes economic gains and losses (utilising data from German test catchment). Testing done in WP4. Consequences of management decisions under different temperature/flow conditions (due to climate change) will also be considered.

B9. Social Attitudes Tool

Data contact	UNIOVI (EDR)
Existing data?	no
Data origin	Questionnaire for public
Data type	database of public preferences and value placed on dams and services provided by dams and rivers (categorised by predictor variables; see description below).
Data format	Spreadsheet .xls .csv
Expected size	1 GB
Data utility	Will feed into decision tool and assist in strategic planning of barrier feasibility and location and provide objective information as a basis for stakeholder conflict resolution.
Data Users	hydropower companies; scientists; water authorities; policy makers; NGOs

Questionnaires will be used to collect data on public attitudes to dams and reservoirs and the financial value the public place on them. These will be done intensively for all Case Studies and additionally in AMBER beneficiary countries not represented in the Case Studies. This data will be used to construct a model of acceptability of dams given basic dam predictors (barrier height, type and age, as well as respondent education, gender, age, country).

C. Case Study data

Data contact	DTU (KA)
Existing data?	no
Data origin	Field and public surveys within Case Study catchments
Data type	(see data types for all tools above)
Data format	(see data formats for all tools above)
Expected size	28 GB
Data utility	(see data utility for all tools above)
Data Users	(see data users for all tools above)

The tools being developed within AMBER require testing and validation within a diverse range of catchments and situations e.g. barrier planning, removal and mitigation. Case Studies sites were chosen to be representative of this diversity. The data collected will inevitably be integrated into improving the functioning and accuracy of the tool, and is more appropriately stored with the tool for which it is being tested. However, during field studies centralised organisation of the collection and storage of the data will be organized.

D. Dissemination data and project metadata

Data contact	WFMF (JD) & SU (ID)
Existing data?	no
Data origin	databases used in assisting dissemination of the project e.g. stakeholders, Barrier Tracker app users, users of AMBER outputs, educational material, publications. Metadata of AMBER project.
Data type	contacts
Data format	databases and documents .xls .csv .doc .pdf
Expected size	6 GB
Data utility	contacting stakeholders, project organisation, promoting AMBER
Data Users	AMBER consortium (internal), and public; scientists

Databases of stakeholders, Barrier Tracker app users and users of AMBER outputs are maintained throughout the project. Promotional material relating to the project will also be maintained, such as Educational data, Newsletters etc. Scientific publications produced by the consortium are also referenced through OpenAIRE and stored at institutional or journal level (depending on Open Access copyright conditions).

2.3 Data not originating directly from AMBER beneficiaries

Data originates from various sources:

(i) Some members of AMBER (IFI, SSIFI, WFMF) brought data to the project prior to commencement and have associated IPR; specifically with an agreement that such data can only be used within AMBER. This was dealt with in the Consortium Agreement which was signed prior to AMBER commencement. Attachment 1 from the Consortium Agreement is included in this Data Management report (as Annex 1) in its full version, for reference.

(ii) Barrier data from Regional and National authorities within Europe is generally open to the public and free to use. However, some data is not and usage agreements have to be drawn up or agreed. Additionally there is some commercial data e.g. through hydropower companies, which require usage agreements although where data cannot be used within the Atlas (publically) or for research purposes it may not be worth collecting the data.

(iii) Citizen Science will be used to collect additional barrier data for the Atlas. This data will be open for use for research and by the public. A statement has been included in the app agreement which users have to actively tick to agree to before continuing. This is worded as such (subject to change prior to app launch):

Who will have access to the data and for what purpose

The AMBER project team and Natural Apptitude will have access to the data submitted. Data will be verified by staff at the World Fish Migration Foundation, before it is made available to JRC. Findings will be presented in a range of outputs, potentially including the Barrier Atlas, academic journals, magazines, project summaries, blog posts, infographics, leaflets, policy briefs and email newsletters. This will help improve scientific understanding of the impact of barriers across Europe. Members of the general public will also have access to records via the AMBER website, although record data will be summarized and will not include your personal information.

(iv) WP4 Case Studies will contain data from specific studies within catchments will include data collected by AMBER members and funded by the EC and thus will be freely available for use. Questionnaire data collected within the Case Studies is collected with a signed agreement for use of data (see ethics deliverable D7.2).

(v) AMBER members will also collect data for validation, which will be freely available.

Throughout the project additional data sources are likely to become available. It is important that signed agreements of Intellectual Property Rights are obtained, and that data is flagged within the data itself whether it has limitations or not on its use/reuse. AMBER members must also be aware of data protection law regarding the storage and use of personal data. This is covered in detail in section 5.1 of this report.

2.4 Data Size

NaturalApptitude, the company creating the CS app, has a server to collect the app data initially, which after pre-processing, will be transferred to JRC. The estimate of the barrier data collected for the barrier inventory and Atlas is **16GB**. This will be held and maintained by JRC (Ispra). At the end of AMBER (May 2020) the contract with NaturalApptitude ends and any data collected will go directly to JRC. The details of this change over will be in the final data management report, within D6.2 (month 48).

Other data is being held by the 20 individual participants in AMBER. However, a central Swansea Server has been made available with a current size of **4.4 TB**, expandable to 11TB if required. The Swansea Server has several roles: (i) document and allow data sharing between AMBER participants (having the guaranteed latest version) (ii) storing reference documentation for the running of AMBER, e.g. contact details, meeting minutes (iii) as a backup for data collected by beneficiaries. **Tables 2, 3 and 4** summarise the estimated sizes for different components of the data collected.

Publication datasets are linked to the publications themselves and stored at the Coordinator's (Swansea University) repository as well as the repositories of the Beneficiaries who generated them (see Section 3.1 below).

Deliverable D5.6 'Plan of Exploitation and Dissemination of Results' contains details of how the outputs from AMBER will be disseminated and the target audience. Table 2 in this document specifically relates to the data output from different tasks, some of which feed into or are combined to produce outputs from AMBER. A summary table (**Table 3**) has been created to show the relationship between the dissemination of data outputs in D5.6 and the different sets of data created within each task. This shows the data that come out of each task, the output created from the data, the method of dissemination and the target audience, thereby combining the information provided in **Table 10** of deliverable D5.6 and **Table 2** of this deliverable.

3 'FAIR' (FINDABLE, ACCESSIBLE, INTEROPERABLE AND RE-USEABLE) DATA

The 'FAIR' principles have the objective of making available data easy to find and access using modern computing methods and the internet. It is recommended that these documents are read by members of the AMBER consortium involved with creating open access databases.
https://www.force11.org/group/fairgroup/fairprinciples
http://www.nature.com/articles/sdata201618

3.1 Findable

AMBER will self-archive both publications and data in open access repositories commonly used by scientists, allowing easily findable and searchable access to this information on the servers where they are held e.g. CRONFA (Swansea University). Repositories can be searched for via:
https://www.openaire.eu/
http://www.opendoar.org/

A website front-end for the Atlas data is also being developed, principally as a public access interface for the barrier data. However, the barrier data will be accessible visually through the website, or in spreadsheet/csv form through JRC. Data which is not spatially associated with the Atlas will be held at Swansea University where possible, although some institutions will hold models and model output data (Open Access). The website front-end for the Atlas will also provide links to all the other data and publications associated with AMBER. Figure 1 shows how the repository lists (and website) are linked to the different data repositories and the types of data being held there. Digital Object Identifiers (DOIs) will be produced for final Open Access data sets https://www.doi.org/.

Some of the datasets being used are already collated in other repositories e.g. Water Framework Directive biological data. Such data sets will not be duplicated but links will be provided to these respositories. There is potential for data to be held in non-institutional repositories such as the Global Biodiversity Information Facility (GBIF) http://www.gbif.org/ or the free repository, Zenodo https://www.zenodo.org/, however it is easier to control the standards of data and data maintenance in institutional repositories if they are well established. For example, Swansea University has a specialized Institutional repository ('data hub') and this will be used for long-term storage of the AMBER datasets.

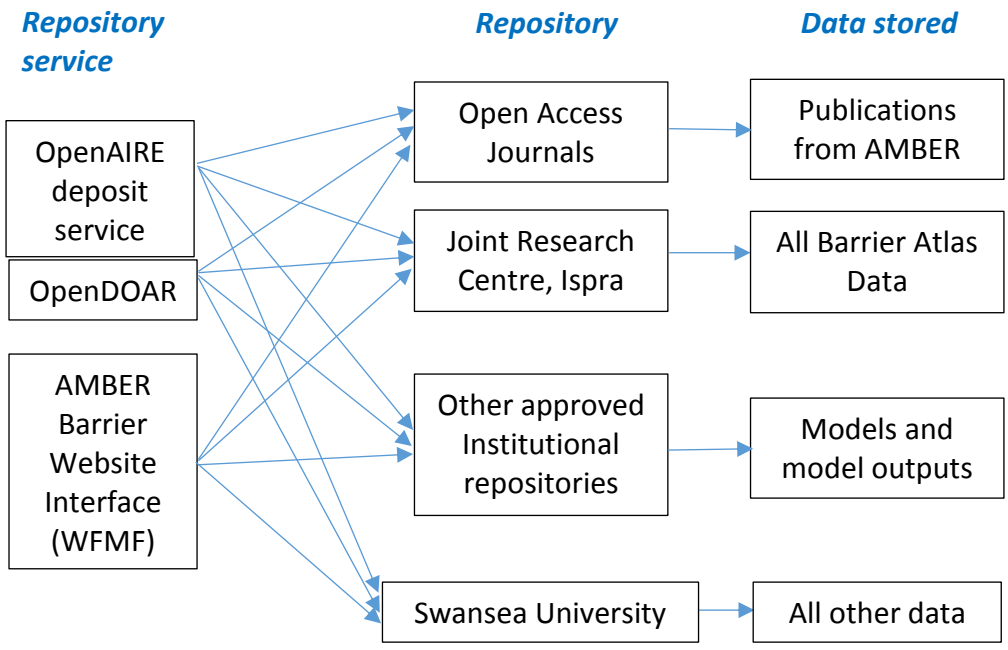

Figure 1. Relationship between repository lists, repositories and data stored

3.1.1 Open access to scientific publications generated by AMBER

Open Access is where the public, without subscription, can access publications. OpenAIRE is an Open Access project that can be used to find and link to Open Access publications.

AMBER will use two forms of Open Access publications:

- **Gold Open Access**: The final publication is freely available directly from the publisher
- **Green Open Access**: An author's version of the publication is made available through an institutional repository, a practice commonly referred to as "self-archiving". There is often an embargo period before the publication can be made available elsewhere.

Researchers within AMBER will ensure that all publications are either Gold or Green Open Access and that they include the terms: "European Union (EU)" and "Horizon 2020"; name of the action, acronym, grant number and duration i.e. **"European Union (EU) Horizon 2020 Programme. Adaptive Management of Barriers in European Rivers (AMBER) Grant#689682 (from 2016 to2020)"**

Each partner will self-archive via open access repositories in order to adhere to Article 29.2 of the GA. Institutional Repositories used by the consortium include:

- CRONFA – Swansea University
- RUO – Repository University of Oviedo
- DRO – Durham Research Online
- Orbit – Technical University of Denmark

In addition, all AMBER publications and associated data sets are stored in Swansea's Server

3.1.2 Costs associated with Open Access

As the AMBER project budget has been devolved, beneficiaries are responsible for forecasting and meeting publication costs, including any costs associated with Open Access.

3.2 Accessibility

3.2.1 Data

Regarding data sets which will be collected as part of the AMBER project , there is no specific data set which cannot be shared i.e. :

- questionnaire data will not include identifiers of the individuals;
- much of the barrier data is publicly available and generated by national or regional agencies;
- case study and validation data collection is funded by the EC and will be publicly available.

However, there are some barrier and fish data used by AMBER that were collected by hydropower companies and Member Estates that are not all publicly available. Efforts are being made to make as much of these open access (detailed below). Where the data cannot be used for open access additional information regarding the restrictions will be kept in a database (that is not necessarily the same as the database containing scientific data) and documents of written agreements will be compiled and filed in a structured manner. Currently we have identified some potential data sets that will not be open access, but the data itself has not yet been collated.

The questionnaire examining the social aspect of barriers and asking opinions on barriers has some relevance under data protection law. Respondents are asked for informed consent for data sharing and long term preservation of the data within the survey, and are provided the details of how the

data will be used. The data from the questionnaire will not be stored with personal details (name, address etc) that could identify them. Data protection law was reformed in April 2016. More information is provided here:

http://ec.europa.eu/justice/data-protection/reform/index_en.htm

Conditions of use for some data collected prior to the Consortium Agreement has been agreed but within the AMBER project no beneficiaries have yet requested that the output of data collected by AMBER be restricted, except with regard to enabling time for AMBER researchers to publish articles based on this data before release.

3.2.2 Software for accessing the data

Through its citizen science portal ('AMBER Atlas website'), AMBER will permit users to access the data collated in the Atlas and, later also the Case Studies. This will also be linked to JRC barrier inventory at the end of the project. In addition to tools for visualising the data (principally a barrier map) there will also be the ability to download different data sets from the website. Currently 'csv' files are considered the best file type for file download as it can be opened directly in a variety of spreadsheet packages (e.g. Microsoft Excel, LibraOffice) as well as utilized directly in a range of statistical and analytical software packages. In addition, the file size tends to be minimal since there is no formatting of the text and data.

Thus, data can be accessed with any common internet browser and opened with an extensive range of software types over different platforms, e.g. Microsoft Office suite on the PC, Linux operating systems, Unix, Apple OS X.

3.3 Interoperable

Since the barrier inventory database will be the first comprehensive barrier inventory in Europe, it is hoped the data structure detailed in D1.1 Guidance on Stream Barrier Surveying and Reporting (Part B) will become the standard for barrier inventories. The database will already be highly interoperable, as it will combine data sets from different origins, comprising a base set of 20 core variables, but will also not discard any additional data collated on single barrier datasets. In order for this to be achieved, the following procedures will be adopted:

- the data will utilise English as a common language and the International System of Units (SI) for measurements
- any categorical data will refer to documentation on what the categories represent and how the categories were created (method), the data collection method will also be referred to within the database.
- scientifically accurate and non-ambiguous vocabularies will be used where possible, or the most commonly accepted terms (in English) if there is no specific scientific definition of the variable.
- words within the data and within column/row titles will be kept to a minimum to make it clear which columns/rows contain the same type of data (until further data has been collected throughout Europe, the specifics of this cannot be detailed).

3.4 Reusable

Most of the data (generated from AMBER and from National databases for the barrier inventory) will be open access and will not have restricted use. There may also be options to link into national databases to get automated updates of barrier data, although the ability to do this on a large scale has not been assessed and is likely to vary greatly depending on the structure of the database from

which the data is obtained and the permissions given by the data owners.

Restrictions on the release of data to open access, to allow time to publish, will be in place. It is estimated that 6 months between collection of the whole data set and allowing open access would be a guideline. However, many data collection activities create outputs that feed into other analysis and models, so there may be cases where the data will be withheld for longer prior to publishing due to the dependence on a data set later within the project timeline.

4 ALLOCATION OF RESOURCE

Costs of making data FAIR within the project are integrated within the specific tasks, particularly WP5 (dissemination) and not separately costed.

5 DATA SECURITY

The Swansea Server is a SFTP (Secure File Transfer Protocol) server, backed up every evening. Copies of data on this server from beneficiaries are also kept by the beneficiaries.

The JRC server will operate indefinitely under the auspices of the EC. Funding to sustain this will be applied for through specific grants in the last 2 years of AMBER. The Swansea Server will retain the data for at least 10 years.

5.1 Data Protection

In May 2018 the General Data Protection Regulation (GDPR) comes into force (Regulation EU 2016/679). http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679. AMBER will comply with current European and national data protection law. Full details of conformity to data protection will be provided in deliverable D7.2 (POPD - H - Requirement No.2). Data Controllers and Data Processers have been designated (as is required in the GDPR) and a legal agreement between Swansea University and the other institutions potentially involved with personal data is currently being written. The basic structure of data protection is detailed below.

Data protection concerns

There are three main areas where data protection is a concern due to the collection of personal data:

1. Collection of drone data over river basins where faces or people may be inadvertently recorded on drone film footage.
2. Audio recordings and opinions taken during questionnaires on river barriers (dams; weirs etc).
3. Emails from voluntary registration on the AMBER app or website.

Personal data that is being collected is this:

- Images of the public (potential to identify them or invade privacy)
- Audio recordings (potential to identify people through the recordings)
- Emails

However, in (1) and (2) there is no intention to retain this personal data and thus there is no processing of personal data, but there is a small risk that personal data will inadvertently be retained. Images of people within footage will be blurred in drone footage (except drone operators,

who will be asked to sign an agreement that their image can be used). Audio recordings will be destroyed after (timely) transcription.

In (3) we are retaining personal data (emails) for the duration of the project (ends 31 May 2020). This is also more complex since the beneficiary (partner) who will be using the data employs a 3rd party outside the consortium to collect the data. Potentially there may be another 3rd party to host the website, and another agreement will need to be drawn up in such a case. Also, at the end of the project and EC body (JRC) will retain data collected by the AMBER app. It is intended that we write to the registered users by email, asking if they wish to continue to be registered. If they do, these emails will then pass to this EC body (JRC), however if they do not, or they do not respond to our request, we will destroy that individual's personal data (emails).

Thus, there are two aspects to protecting personal data within this project. The first is ensuring in drone and questionnaire work, that personal data is not retained. The second is ensuring that personal data from app/website registration is properly controlled. As such a legal agreement needs to be drawn up between the parties involved.

Structure of data management agreement:

Overall data controller (1)
Institution: **Swansea University** (SU), UK

General Responsibilities:
Ensuring that the legal agreement and the EC law is abided by through contact with data processors and a co-data controller. Overall authority of data control.

Specific responsibilities:
Routine contact with data controllers to ensure personal data is not being circulated outside the signatories to this agreement and to ensure that data controllers' responsibilities are being followed.

Co data controller and data processer (2) – app/website data
Institution: **World Fish Migration Foundation** (WFMF), Netherlands

General Responsibilities:
To be co-data controller regarding the emails (i.e. collaborate with SU to ensure email data is correctly controlled) and to utilise this data for sending emails to registered users (as a data processor).

Specific responsibilities:
To collaborate with (1) in determining rules for controlling email data. To ensure email data does not go beyond Natural Apptitude or WFMF. To deal with email data at project end (destruction or change over of responsibilities to JRC).

Data Processor (3) – drone data
Institution: **Durham University** (DU)

General Responsibilities:
To follow the guidelines in the legal agreement (determined and regulated by data controller (1) and to follow the EU and UK law in regard to data protection).

Specific responsibilities:

To coordinate the assurance of data protection for all the drone work.

Data Processor (4) – drone data
Institution: **Instytut Rybactwa Srodladowego Im Stanistawa Sakowicza** (SSIFI), Poland

General Responsibilities:
To follow the guidelines in the legal agreement (determined and regulated by data controller (1) and to follow the EU and UK law in regard to data protection).

Specific responsibilities:
To coordinate the assurance of data protection for all the drone work.

Data Processor (5) – questionnaire audio recordings
Institution: **University of Oviedo** (UNIOVI), Spain

General Responsibilities:
To follow the guidelines in the legal agreement (determined and regulated by data controller (1) and to follow the EU and UK law in regard to data protection).

Specific responsibilities:
To destroy audio data on opinions after it has been transcribed (in a timely manner). Also to ensure questionnaire data collected does not contain personal data.

Note:
The Joint Research Centre (Ispra) may receive personal data (public emails) during a transfer process, following the conclusion of the AMBER project, but they will not hold or process personal data within AMBER prior to this. This hand-over will be detailed in D6.2.

No one except WFMF and Natural Apptitude (app developer) are permitted to hold or be given personal data relating to AMBER during the lifetime of the AMBER project. Personal data of those working directly on the AMBER project (such as beneficiaries' emails and addresses) can be held and circulated, following relevant EU and national data protection law.

In summary, only WFMF and Natural Apptitude will be handling personal data during the AMBER project(emails). However, some institutions have been given specific authority to ensure that personal data is not inadvertently collected in the questionnaire (University of Oviedo); and in the drone work (Durham University and the Instytut Rybactwa Srodladowego Im Stanistawa Sakowicza) through destroying audio data and smudging faces in video footage (respectively).

6 APPENDIX 1 – LINKS BETWEEN TASKS, DATA, DELIVERABLES AND DISSEMINATION

Table 3. Summary of data sources and types collected within the AMBER project WP1,2,3. 'Further utility' are applications for the data outside the direct scope of the project. (internal) refers to file type used whilst being worked on within AMBER and (external) is the file type as it will be presented to the public for open access.

Task	Data	Origin	Uses	Format	Estimated Size	Further utility
T1.2.1	Collation of barrier data throughout Europe	National and Regional authorities were data exists; specific river studies	To create the Barrier Atlas; (i) to inform policy decisions;(ii) strategic decision making; (iii) numerous models with further data output (listed here)	.xls (internal use) .csv (external use)	10GB	Scientific investigations
D1.2	Metadata on the barrier inventory (T1.2.1)	Created by AMBER based on the type of data collected in T1.2.1	Overview of data within the barrier inventory	.xls (internal use) .csv (external use)	1GB	Understanding the barrier data; procuring further barrier data within Europe
T1.2.2	Validation data	AMBER experts collecting field data on barriers	To allow comparability between survey methods and countries in T1.2.1; to give more realistic estimates of total barrier numbers in Europe, and within Member States; to be included as data within the European barrier inventory (D1.3)	.xls (internal use) .csv (external use)	5GB	Representative of intense and comprehensive barrier surveys
D1.3	Barrier inventory	Combination of the data obtained from collated European barrier data (T1.2.1); validation data (T.1.2.2) and Case Study data	The basis of the European barrier map (the 'Atlas')	.csv GIS theme (.shp; .shx; .dbf)	16GB	Research; shaping policy; promotion of project

T2.1	Europe wide connectivity and biodiversity data	Compilation of stream surveys of plant/invertebrate/fish from national WFD databases within Europe	To produce a predictive model of barrier effects on ecology	.csv	50GB	-data already available-
T2.2.1	Fish guilds predicted from habitat and barrier data	Fisheries; barrier; hydrological and stratified habitat data for selected rivers (pre-existing). Prediction of expected ecological guilds based on this data (generated with AMBER).	Assessing the effectivenss of Restoration Alternatives Analysis	.xls (internal use) .csv (external use) GIS theme (.shp; .shx; .dbf)	10GB	
T2.2.2	Drone generated river habitat data	Drone flight film and photos in selected river catchments	Developing rapid habitat assessment methodology through image interpretation.	.mp4 (video) .jpg (photo) .xls (predicted habitats) .csv (predicted habitats)	20GB	Research for improving image interpretation; promotional media; examining the catchments from the air
T2.2.3	European sediment connectivity data	Barrier Inventory data (D1.3); available hydrological data-> with output of sediment connectivity (movement) in rivers throughout Europe.	Creating sediment connectivity (movement) map for Europe, based on barriers.	.xls (internal) .csv (external) GIS theme (.shp; .shx; .dbf)	10GB	Widespread research applications
T2.3	Effect of climate change on river connectivity	Stream flow & climate data for 441 catchments in 15 countries (European Environment Agency data); National WFD data bases of catchment topography/size. -> output of habitat deficit; stress days; habitat change	Illustrate predictive model of analyzing effect of climate change on connectivity (based on barriers)	.xls (internal) .csv (external)	20GB	Research; example for strategic planning of climate change scenarios for environment agencies
T2.5.1	eDNA detection thresholds	AMBER eDNA research for metabarcoding protocols	Thresholds to develop the metabarcoding toolkit	.xls (internal) .csv (external)	1GB	Widespread use for application of

						metabarcoding; improvement/research to further develop metabarcoding
T2.5.2	Presence/absence of aquatic biota based on eDNA sampling in test catchments	AMBER field collection and analysis of eDNA, processed with metabarcoding toolkit (T2.5.1) and barrier data collected in Case Studies	Illustrate use of eDNA toolkit to determine species presence/absence	.xls (internal) .csv (external)	1GB	Example for other metabarcoding field exercises
T2.6	Ecosystem services and interaction with stakeholders	Ecosystem Services evaluated in the Case Studies; and stakeholders/stakeholder interests identified in the Case Studies.	Data to inform model development	.xls (internal) .csv (external) **NB. Data protection considerations**	1GB	Example of relationships between ESS, barriers, and stakeholders
T3.1.1/ T3.1.2	Hydropower potential and passability	Structural and hydrological and passability data collected by AMBER from test catchment in Germany -> output of hydropower generation potential; dam construction costs at different locations	For prioritization in the barrier mitigation and hydropower placement decision tool	.xls (internal) .csv (external)	1GB	Example of assessing hydropower potential (though data likely to be combined with other data within decision tool)
T3.2.1 (D3.1)	Hydrodynamic conditions at river infrastructures	Flow velocities, shear and turbulence values associated with barriers and fishways; hydrodynamics for key biological species	Determines hydrodynamic parameters/thresholds for species and how structures thus permit/prevent passage: For Agent Based Model	.xls (internal) .csv (external)	1GB	Useful data for research, regulatory bodies and hydropower industry
T3.2.2	Behaviour and locomotory performance of weak swimmers	Behaviour and locomotory performance of weak swimming species (e.g. crayfish) under conditions found at barriers (AMBER experiment at SOTON labs). Also, similar data for inverts and macrophytes collated.	Used to develop response criteria for range of organisms in Agent Based Model. **NB. Data collection has ethics considerations (working with animals)**	.xls (internal) .csv (external)	1GB	Research

T3.2.4	Field data of passability of species (focus on non-salmonids)	Movement data of non-salmonid spp. Including weak swimmers; invertebrates and macrophytes. From surveys and tagging exercises in the Case Study sites.	For testing Agent Based Model. **NB. Data collection has ethics considerations (tagging)**	.xls (internal) .csv (external)	2GB	Important information for regulatory bodies: informing EU habitats directive and Convention on Biological Diversity and movement of invasives.
T3.3 (D3.2, MS3)	Cost-Benefit of restoring stream connectivity	Data collected in Case Studies assessing cost-benefit (including non-market benefits/costs) and data from non-market benefit inventories, of various restoration options; includes MS3 Evaluation of Natural Capital data.	Will feed in to barrier planning and decision tool.	.xls (internal) .csv (external)	2GB	Regulators/Public: data to assist conflict resolution in barrier management; Research
T3.3 (D3.5; MS10)	Social attitudes to dams in rivers	Questionnaire on social attitudes to dams in rivers	Will feed in to barrier planning and decision tool.	.xls (internal) .csv (external) **NB. Data protection considerations**	1GB	Regulatory bodies/public: for understanding and informing conflict resolution
D3.3	Inventory of barriers and river structures within German test catchment	Data collected on location and properties of barriers within the German test catchment	**NB. Likely to be integrated into the validation data** (T1.2.2)	.xls (internal) .csv (external)	3GB	Representative of intense and comprehensive barrier surveys

Table 4. Data collected during WP4 (Case Studies). NB. Much of this field data is collated for use in specific tasks in WP1,2 and 3 (**Table 2**).

Task	Data	Origin	Uses	Format	Estimated Size	Further utility
T4.1.1	River Nalon field data (Spain)	Field work in Case Study areas	Data feeds into tasks in WPs 1 to 3.	.mp4 .jpg .xls .csv GIS theme (.shp; .shx; .dbf)	4GB	Case study examples for public/regulators; catchment management within the specific catchments; publicity
T4.1.2	River Allier (France) field data				4GB	
T4.1.3	River Munster (Ireland) field data				4GB	
T4.1.4	River Gary (Scotland) field data				4GB	
T4.1.5	River Vistula (Poland) field data				4GB	
T4.1.6	Lowland river (various countries) field data				4GB	
T4.1.7	River Guardalhorce (Spain) field data				4GB	
T4.2.1 (D4.3)	Trans-European Status of Atlantic Salmon	A trans-European river by river GIS map showing status of Salmon derived from barrier inventory (D1.3) and connectivity and biodiversity data (T2.1)	[Output directly for external (policy shaping/AMBER promotional use)]	.xls (internal) .csv (external) GIS theme (.shp; .shx; .dbf)	5GB	Informing policy decisions; national and local conservation/ restoration efforts; promotion of AMBER

Table 5. Data from WP5, 6, 7.

Task	Data	Origin	Uses	Format	Estimated Size	Further utility
T5.1/T6.1	Parallel Projects database	Collated during project	Linking AMBER to other projects	.xls **NB. Data protection considerations**	1GB	For future projects
T5.1/T6.1	AMBER member details and contacts	Collated during project	Communication within project	.xls **NB. Data protection considerations**	1GB	For future projects and for contact regarding further information on AMBER or future collaborations
T5.2.1	Stakeholder database	Collated during project	Feedback pre-output; dissemination of results and information	.xls **NB. Data protection considerations**	1GB	Links stakeholders into outputs of project to ensure maximum use
T5.3.2	Registered app users	Database with details of number, location and other information relating to app use (including non-registered and registered users)	Monitor uptake and use of app; feed into improvements to app; managing citizen science activity and improving website which presents the data.	.xls (internal only) **NB. Database itself not to be made public. Data protection issues.**	1GB	**Complete database not to be released**; some analysed data may be presented for publicity (eg number of users)
D7.1/7.2	Ethics documentation	Database of ethics documentation for the AMBER project	Keep track of ethics documentation for beneficiaries. Internal (AMBER) use only.	.xls (internal only)	1GB	**-Confidential: Internal use only-**
T 6.2 (D6.2)	AMBER project metadata	Data base detailing all the data collected and produced by AMBER (listed within **Table 1,2 and 3**) including details of data that is not publically available (with source contacts)	For external users	.xls	1GB	Allows ease of access and understanding of available AMBER data to all external user types.

Table 6. Summary of how data leads to deliverables and is then disseminated to specific audiences WP 1, 2, 3, 4.

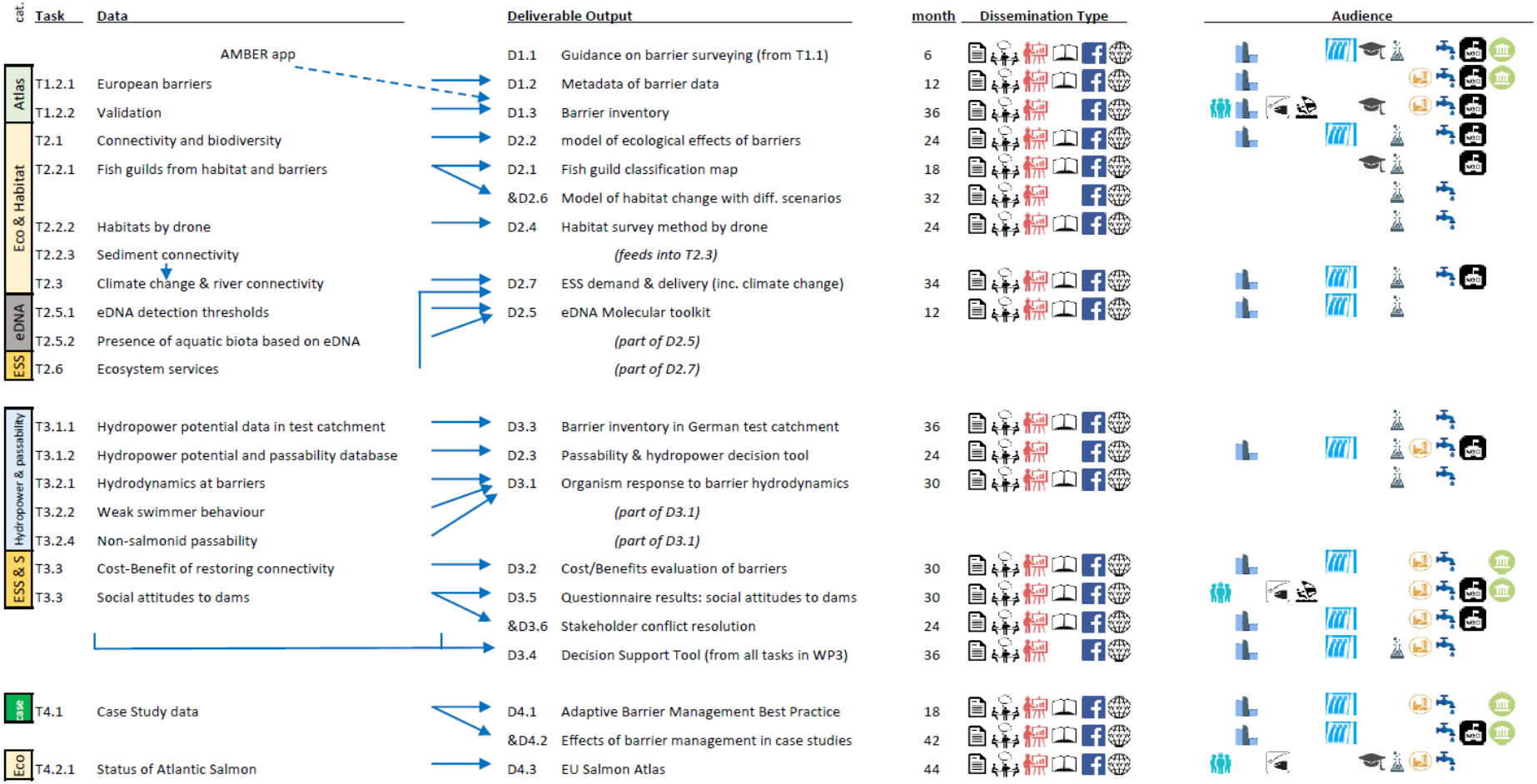

Table 7. Summary of outputs from non-specific data sources (WP5). NB. WP6 and WP7 outputs are internal management documents and not for external audiences.

Data	Deliverable Output		month	Dissemination Type	Audience
amalgamation	(D5.7)	AMBER project and call to action	4		
from	(D5.7)	Project setup video	12		
diverse	(D5.7)	Progress and preliminary results video	24		
data	(D5.7)	Project outcome and results video	48		
sources	D5.4	AMBER app	18		
throughout	D5.8	Citizen science material	18		
project	D5.10	Book: Adaptive Barrier Management	48		
		Publications:			
(D1.3)	D5.3	Development of the Barrier Inventory and online Atlas	48		
(D1.3)	D5.3	Extent of river fragmentation in Europe	48		
(D2.2)	D5.3	Conceptual framework for estimating barrier effects on fluvial processes	48		
(D2.3)	D5.3	Conceptual framework for estimating Barrier effects for a range of aquatic biota	48		
(D3.4)	D5.3	Application of the AMBER decision support tool for barrier mitigation and planning	48		
(D3.5/D3.6)	D5.3	Socio-economic drivers and impediments for successful reconnecting of rivers	48		
(D4.2)	D5.3	Case studies in restoration of stream connectivity	48		

Key

as Table 2 with addition of  = book publication