



**Project Title:** ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS

**Project number:** 641762

**Project Acronym:** ECOPOTENTIAL

**Proposal full title:** IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS

**Type:** Research and innovation actions

**Work program topics addressed:** SC5-16-2014: “Making Earth Observation and Monitoring Data usable for ecosystem modelling and services”

## Deliverable No: D1.3

### ECOPOTENTIAL Data Management Plan



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762



## Document Metadata

<b>Project ref. number</b>	<b>641762</b>
<b>Project title</b>	<b>ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS</b>

<b>Document title</b>	ECOPOTENTIAL Data Management Plan
<b>Document status</b>	Final Draft
<b>Document version</b>	3.4
<b>Online access</b>	<a href="http://www.ecopotential-project.eu/products">http://www.ecopotential-project.eu/products</a>
<b>Diffusion</b>	Public
<b>Author(s)</b>	Alberto Basset (UNILE), Palma Blonda (CNR), Ghada El Serafy (DELTA RES), Francesco Lovergine (CNR), Ioannis Manakos (CERTH), Carmela Marangi (CNR), Joan Masò (CREAF), Stefano Nativi (CNR), Johannes Peterseil (EAA), Antonio Reppucci (STARLAB) and the EDSI (ECOPOTENTIAL Data Sharing and Interoperability) Task Force
<b>Editor(s)</b>	Stefano Nativi (CNR) and Carmela Marangi (CNR)
<b>Keywords</b>	Data Management Plan, Open Data Access, Open Science, Data Preservation, Interoperability, System Architecture.
<b>Abstract</b>	This document is the first version of the Data Management Plan of the H2020 ECOPOTENTIAL project (project number 641762).



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762





## Outline

<b>Document Metadata</b>	<b>2</b>
<b>Outline</b>	<b>3</b>
<b>Abstract</b>	<b>12</b>
<b>Scope</b>	<b>12</b>
<b>Motivations</b>	<b>12</b>
<b>Document History</b>	<b>12</b>
<b>ECOPOTENTIAL principles for Data Access and Services Interoperability</b>	<b>15</b>
<b>DMP Methodology</b>	<b>15</b>
<b>ECOPOTENTIAL Data Types</b>	<b>16</b>
<b>Data Types covered by the present DMP version</b>	<b>16</b>
<b>ECOPOTENTIAL dataset category #1: RS Raw Data (WP3 and WP4)</b>	<b>17</b>
<b>Data set reference and name: Landsat Raw Data Products</b>	<b>17</b>
Identifier for the data set to be produced	17
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>19</b>
Origin (in case it is collected)	19
Data model and scale	19
To whom it could be useful	20
Whether it underpins a scientific publication	20
Possibilities for integration and reuse	20
Reference to existing suitable standards of the discipline	20
<b>Data sharing (Description of how data will be shared)</b>	<b>21</b>
Access procedures	21
Embargo periods (if any)	21
Outlines of technical mechanisms for dissemination and necessary software	21
Definition of whether access will be widely open or restricted to specific groups	21
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)	21
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)	21
<b>Archiving and preservation (including storage and backup)</b>	<b>21</b>
Description of the procedures that will be put in place for long-term preservation	21
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered	21
<b>Data set reference and name: MODIS AQUA/TERRA Data Products</b>	<b>22</b>
Identifier for the data set to be produced	22
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>22</b>



Origin (in case it is collected) _____	23
Data model and scale _____	23
To whom it could be useful _____	23
Whether it underpins a scientific publication _____	23
Possibilities for integration and reuse _____	23
Reference to existing suitable standards of the discipline _____	23
<b>Data sharing (Description of how data will be shared) _____</b>	<b>24</b>
Access procedures _____	24
Embargo periods (if any) _____	24
Outlines of technical mechanisms for dissemination and necessary software _____	24
Definition of whether access will be widely open or restricted to specific groups _____	24
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	24
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related) _____	25
<b>Archiving and preservation (including storage and backup) _____</b>	<b>25</b>
Description of the procedures that will be put in place for long-term preservation _____	25
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	25
<b>Data set reference and name: Sentinel Raw Data Products _____</b>	<b>25</b>
Identifier for the data set to be produced _____	25
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>26</b>
Origin (in case it is collected) _____	26
Data model and scale _____	26
To whom it could be useful _____	26
Whether it underpins a scientific publication _____	26
Possibilities for integration and reuse _____	26
<b>Standards and metadata _____</b>	<b>27</b>
Reference to existing suitable standards of the discipline _____	27
<b>Data sharing (Description of how data will be shared) _____</b>	<b>27</b>
Access procedures _____	27
Embargo periods (if any) _____	27
Outlines of technical mechanisms for dissemination and necessary software _____	27
Definition of whether access will be widely open or restricted to specific groups _____	27
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	27
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related) _____	27
<b>Archiving and preservation (including storage and backup) _____</b>	<b>28</b>
Description of the procedures that will be put in place for long-term preservation _____	28
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	28
<b><i>ECOPOTENTIAL dataset category #2: RS High-level Data (WP4) _____</i></b>	<b>29</b>
<b>Data set reference and name: Water turbidity (inland waters)(WT_iw) _____</b>	<b>29</b>
Identifier for the data set to be produced _____	29



<b>Data set description (Description of the data that will be generated or collected)</b>	<b>29</b>
Origin (in case it is collected)	29
Data model and scale	29
To whom it could be useful	29
Whether it underpins a scientific publication	29
Information on the existence (or not) of similar data	29
Possibilities for integration and reuse	29
Reference to existing suitable standards of the discipline	29
<b>Data sharing (Description of how data will be shared)</b>	<b>30</b>
Access procedures	30
Embargo periods (if any)	30
Outlines of technical mechanisms for dissemination and necessary software	30
Definition of whether access will be widely open or restricted to specific groups	30
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)	30
<b>Archiving and preservation (including storage and backup)</b>	<b>30</b>
Description of the procedures that will be put in place for long-term preservation	30
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered	30
<b>Data set reference and name: Land Cover/use (LCLU) and LCLU change maps, multiple sites</b>	<b>31</b>
Identifier for the data set to be produced	31
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>31</b>
Origin (in case it is collected)	31
Nature and scale	31
To whom it could be useful	31
Whether it underpins a scientific publication	31
Information on the existence (or not) of similar data	31
Possibilities for integration and reuse	31
<b>Standards and metadata</b>	<b>32</b>
Reference to existing suitable standards of the discipline	32
<b>Data sharing (Description of how data will be shared)</b>	<b>32</b>
Access procedures	32
Embargo periods (if any)	32
Outlines of technical mechanisms for dissemination and necessary software	32
Definition of whether access will be widely open or restricted to specific groups	32
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)	32
<b>Archiving and preservation (including storage and backup)</b>	<b>32</b>
Description of the procedures that will be put in place for long-term preservation	32
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered	32
<b>Data set reference and name: Habitat and habitat change maps, multiple sites</b>	<b>33</b>
Identifier for the data set to be produced	33
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>33</b>
Origin (in case it is collected)	33



Nature and scale _____	33
To whom it could be useful _____	33
Whether it underpins a scientific publication _____	33
Information on the existence (or not) of similar data _____	33
Possibilities for integration and reuse _____	33
Reference to existing suitable standards of the discipline _____	33
<b>Data sharing (Description of how data will be shared) _____</b>	<b>34</b>
Access procedures _____	34
Embargo periods (if any) _____	34
Outlines of technical mechanisms for dissemination and necessary software _____	34
Definition of whether access will be widely open or restricted to specific groups _____	34
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	34
<b>Archiving and preservation (including storage and backup) _____</b>	<b>34</b>
Description of the procedures that will be put in place for long-term preservation _____	34
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	34
<b>Data set reference and name: Soil Moisture (SM) _____</b>	<b>35</b>
Identifier for the data set to be produced _____	35
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>35</b>
Origin (in case it is collected) _____	35
Nature and scale _____	35
To whom it could be useful _____	35
Whether it underpins a scientific publication _____	35
Information on the existence (or not) of similar data _____	35
Possibilities for integration and reuse _____	35
<b>Standards and metadata _____</b>	<b>35</b>
Reference to existing suitable standards of the discipline _____	35
<b>Data sharing (Description of how data will be shared) _____</b>	<b>35</b>
Access procedures _____	35
Embargo periods (if any) _____	35
Outlines of technical mechanisms for dissemination and necessary software _____	35
Definition of whether access will be widely open or restricted to specific groups _____	36
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	36
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related) _____	36
<b>Archiving and preservation (including storage and backup) _____</b>	<b>36</b>
Description of the procedures that will be put in place for long-term preservation _____	36
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	36
<b>Data set reference and name: Seasonal Water Bodies (SWB) _____</b>	<b>36</b>
Identifier for the data set to be produced _____	36
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>36</b>
Origin (in case it is collected) _____	36



Nature and scale _____	36
To whom it could be useful _____	37
Whether it underpins a scientific publication _____	37
Information on the existence (or not) of similar data _____	37
Possibilities for integration and reuse _____	37
<b>Standards and metadata _____</b>	<b>37</b>
Reference to existing suitable standards of the discipline. _____	37
<b>Data sharing (Description of how data will be shared) _____</b>	<b>37</b>
Access procedures _____	37
Embargo periods (if any) _____	37
Outlines of technical mechanisms for dissemination and necessary software _____	37
Definition of whether access will be widely open or restricted to specific groups _____	37
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	37
<b>Archiving and preservation (including storage and backup) _____</b>	<b>37</b>
Description of the procedures that will be put in place for long-term preservation _____	37
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	37
<b>Data set reference and name: Snow Cover Area (SCA) _____</b>	<b>38</b>
Identifier for the data set to be produced _____	38
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>38</b>
Origin (in case it is collected) _____	38
Nature and scale _____	38
To whom it could be useful _____	38
Whether it underpins a scientific publication _____	38
Information on the existence (or not) of similar data _____	38
Possibilities for integration and reuse _____	38
<b>Standards and metadata _____</b>	<b>38</b>
Reference to existing suitable standards of the discipline _____	38
Data sharing (Description of how data will be shared) _____	38
Access procedures _____	38
Embargo periods (if any) _____	38
Outlines of technical mechanisms for dissemination and necessary software _____	39
Definition of whether access will be widely open or restricted to specific groups _____	39
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	39
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related) _____	39
<b>Archiving and preservation (including storage and backup) _____</b>	<b>39</b>
Description of the procedures that will be put in place for long-term preservation _____	39
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	39
<b>Data set reference and name: Snow Cover Maps (SCM) _____</b>	<b>39</b>
Identifier for the data set to be produced _____	39
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>39</b>



Origin (in case it is collected) _____	39
Nature and scale _____	39
To whom it could be useful _____	39
Whether it underpins a scientific publication _____	40
Information on the existence (or not) of similar data _____	40
Possibilities for integration and reuse _____	40
<b>Standards and metadata _____</b>	<b>40</b>
Reference to existing suitable standards of the discipline _____	40
<b>Data sharing (Description of how data will be shared) _____</b>	<b>40</b>
Access procedure _____	40
Embargo periods (if any) _____	40
Outlines of technical mechanisms for dissemination and necessary software _____	41
Other tools for enabling re-use _____	41
Definition of whether access will be widely open or restricted to specific groups _____	41
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	41
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related) _____	41
<b>Archiving and preservation (including storage and backup) _____</b>	<b>41</b>
Description of the procedures that will be put in place for long-term preservation _____	41
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	41
<b>Useful References _____</b>	<b>41</b>
<b>Data set reference and name: Water Bodies Delineation (WBD) _____</b>	<b>41</b>
Identifier for the data set to be produced _____	41
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>42</b>
Origin (in case it is collected) _____	42
Nature and scale _____	42
To whom it could be useful _____	42
Whether it underpins a scientific publication _____	42
Information on the existence (or not) of similar data _____	42
Possibilities for integration and reuse _____	42
<b>Standards and metadata _____</b>	<b>42</b>
Reference to existing suitable standards of the discipline _____	42
Data sharing (Description of how data will be shared) _____	42
Access procedures _____	42
Outlines of technical mechanisms for dissemination and necessary software _____	42
Definition of whether access will be widely open or restricted to specific groups _____	42
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	42
<b>Archiving and preservation (including storage and backup) _____</b>	<b>43</b>
Description of the procedures that will be put in place for long-term preservation _____	43
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	43





<b>Data set reference and name: Maps of landscape measures indicating fragmentation or connectivity, multiple sites</b>	<b>43</b>
Identifier for the data set to be produced	43
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>43</b>
Origin (in case it is collected)	43
Nature and scale	43
To whom it could be useful	43
Whether it underpins a scientific publication	44
Information on the existence (or not) of similar data	44
Possibilities for integration and reuse	44
<b>Standards and metadata</b>	<b>44</b>
Reference to existing suitable standards of the discipline	44
<b>Data sharing (Description of how data will be shared)</b>	<b>44</b>
Access procedures	44
Embargo periods (if any)	44
Outlines of technical mechanisms for dissemination and necessary software	44
Other tools for enabling re-use	44
Definition of whether access will be widely open or restricted to specific groups	44
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)	44
In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)	45
<b>Archiving and preservation (including storage and backup)</b>	<b>45</b>
Description of the procedures that will be put in place for long-term preservation	45
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered	45
<b>Useful References</b>	<b>45</b>
<b><i>ECOPOTENTIAL dataset category #3: in-situ Data (WP5)</i></b>	<b>46</b>
<b>Data set reference and name: ECOPOTENTIAL in-situ Data</b>	<b>46</b>
Identifier for the data set to be produced	46
<b>Data set description (Description of the data that will be generated or collected)</b>	<b>46</b>
Origin (in case it is collected)	46
Nature and scale	46
To whom it could be useful	49
Whether it underpins a scientific publication	49
Information on the existence (or not) of similar data	49
Possibilities for integration and reuse	49
<b>Standards and metadata</b>	<b>49</b>
Reference to existing suitable standards of the discipline.	49
Specification of data format	50
If these do not exist, an outline on how and what metadata will be created	50
<b>Data sharing (Description of how data will be shared)</b>	<b>50</b>
Access procedures	51
Embargo periods (if any)	51



Outlines of technical mechanisms for dissemination and necessary software _____	51
Other tools for enabling re-use _____	51
Definition of whether access will be widely open or restricted to specific groups _____	51
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	52
<b>Archiving and preservation (including storage and backup) _____</b>	<b>52</b>
Description of the procedures that will be put in place for long-term preservation _____	52
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	52
<b>Data set reference and name: LifeWatch-ITA Data Products _____</b>	<b>52</b>
Identifier for the data set to be produced _____	52
<b>Data set description (Description of the data that will be generated or collected) _____</b>	<b>52</b>
Origin (in case it is collected) _____	52
Nature and scale _____	52
To whom it could be useful _____	53
Whether it underpins a scientific publication _____	53
Possibilities for integration and reuse _____	53
<b>Standards and metadata _____</b>	<b>53</b>
Reference to existing suitable standards of the discipline. _____	53
<b>Data sharing (Description of how data will be shared) _____</b>	<b>54</b>
Access procedures _____	54
Embargo periods (if any) _____	54
Outlines of technical mechanisms for dissemination and necessary software _____	55
Other tools for enabling re-use _____	55
Definition of whether access will be widely open or restricted to specific groups _____	55
Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.) _____	55
Description of the procedures that will be put in place for long-term preservation _____	55
Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered _____	55
<b><i>ECOPOTENTIAL dataset category #4: Model Results (WP6, WP7, and WP8) _____</i></b>	<b><i>56</i></b>
<b><i>ECOPOTENTIAL dataset category #5: Indicators/indexes (WP7, WP8, and WP9) _____</i></b>	<b><i>57</i></b>
<b><i>Annex A: GEOSS Data Sharing Principles _____</i></b>	<b><i>58</i></b>
<b><i>Annex B: GEOSS Data Management Principles _____</i></b>	<b><i>59</i></b>
<b>Discoverability _____</b>	<b>59</b>
<b>Accessibility _____</b>	<b>59</b>
<b>Usability _____</b>	<b>59</b>
<b>Preservation _____</b>	<b>59</b>
<b>Curation _____</b>	<b>59</b>
<b><i>Annex C: GEOSS Architecture Principles _____</i></b>	<b><i>60</i></b>
<b><i>Annex D: Guidelines on Data Management in H2020 _____</i></b>	<b><i>61</i></b>



<b><i>Annex E: Glossary</i></b>	<b>62</b>
<b><i>References</i></b>	<b>63</b>



## Abstract

This document is the first version of the Data Management Plan of the H2020 ECO POTENTIAL project (project number 641762).

## Scope

**To describe the data management life cycle of all data sets that will be collected, processed or generated by the ECO POTENTIAL project.**

This document outlines **how research data will be handled** during the ECO POTENTIAL project, and even after the project is completed.

This document describes **what data will be collected, processed or generated** and **what methodology and standards will be applied**, whether and **how this data will be shared and/or made open**, and **how it will be curated and preserved**.

The ECO POTENTIAL DMP is not a fixed document; it evolves and gains more precision and substance during the lifespan of the project. **This is the first version.**

## Motivations

The **described policy should clearly reflect the current state of consortium agreements regarding data management** and be consistent with those referring to exploitation and protection of results.

The data management section can be considered as **a checklist for the future** and as **a reference for the resource and budget allocations related to data management**.

In H2020 a pilot action on open access to research data will be implemented. ECO POTENTIAL, being one of the participating projects, is required to develop a Data Management Plan (DMP), in which the Consortium will specify what data will be open and how they will be accessible, usable and preserved.

## Document History

Date	Action	Authors
05 Oct 2015	Draft the first version	Stefano Nativi (EDSI Task Force and WP10)
21 Dec 2015	Sections Update	Stefano Nativi
21 Dec 2015	Section content: Section "RS Raw Data" <ul style="list-style-type: none"> <li>• Landsat data</li> <li>• MODIS AQUA/TERRA Data.</li> <li>• Sentinel Raw Data Products</li> </ul> Section content: Section "RS High-level Data" <ul style="list-style-type: none"> <li>• Water turbidity (inland waters)</li> </ul>	Creator: WP4 Leader: CREAM (Joan Masò) Editing: Stefano Nativi



21 Dec 2015	<p>Section content: Section “RS High-level Data”</p> <ul style="list-style-type: none"> <li>• Land Cover/use (LCLU) and LCLU change maps, multiple sites</li> <li>• Habitat and habitat change maps, multiple sites</li> </ul>	<p>Creator: CNR-ISSIA (Palma Blonda and Francesco Lovergine)</p> <p>Editing: Stefano Nativi</p>
21 Dec 2015	<p>Section content: Section “RS High-level Data”</p> <ul style="list-style-type: none"> <li>• Seasonal Water Bodies (SWB)</li> <li>• Snow Cover Area (SCA)</li> <li>• Water Bodies Delineation (WBD)</li> </ul>	<p>Creator: StarLab (Antonio Reppucci)</p> <p>Editing: Stefano Nativi</p>
30 Dec 2015	<p>Section content: Section “RS High-level Data”</p> <ul style="list-style-type: none"> <li>• Maps of landscape measures indicating fragmentation or connectivity, multiple sites</li> </ul>	<p>Creator: CERTH (Ioannis Manakos)</p> <p>Editing: Stefano Nativi</p>
30 Dec 2015	<p>Section: “Indicators/indexes”</p>	<p>Creator: WP7 Leader (Ghada El Serafy)</p> <p>Editing: Stefano Nativi</p>
26 Jan. 2016	<p>Section: “RS High-level Data”</p> <ul style="list-style-type: none"> <li>• Land Cover/use (LCLU) and LCLU change maps, multiple sites</li> <li>• Habitat and habitat change maps, multiple sites</li> <li>• Soil Moisture</li> </ul>	<p>Creator: CNR_ISSIA (Palma Blonda, Anna Balenzano)</p> <p>Editing: Stefano Nativi</p>
03 Feb. 2016	<p>Section: “ECOPOTENTIAL in-situ data”</p>	<p>Creator: WP5 leader (Johannes Peterseil, EAA)</p> <p>Editing: Carmela Marangi and Stefano Nativi</p>
11 Feb 2016	<p>Section: “LifeWatch-ITA Data Products”</p>	<p>Creator: UNILE (Alberto Basset)</p> <p>Editing: Carmela Marangi and Stefano Nativi</p>
27 Feb 2016	<p>Editorial activities</p>	<p>Carmela Marangi</p>
28 Feb 2016	<p>Final Revision and Editing</p>	<p>Stefano Nativi</p>





## ECOPOTENTIAL principles for Data Access and Services Interoperability

ECOPOTENTIAL introduced a set of architecture and interoperability principles to facilitate data (and the associated software) discovery, access, (re-)use, and preservation:

- AP1. To build the ECOPOTENTIAL data and services infrastructure on the existing and under development digital systems –noticeably, the digital systems managed by WP3, WP4, WP5, etc.
- AP2. Not to impose any “common solution/specification” but advocate the use of open (international and Community) standards and interoperability APIs.
- AP3. To provide a common, consistent, and “high-level” entry point ECOPOTENTIAL platform for discovering, accessing, and using ECOPOTENTIAL ecosystem services –for interoperability to GEOSS, Copernicus, and other EC-funded programmes.
- AP4. To comply with the GEOSS Architecture Principles (see Annex C).
- AP5. To comply with the GEOSS “resource sharing” and “resource management” principles –including quality and preservation (see Annexes A and B).
- AP6. To comply with the EC Open Data Access principles (see Annex D and the Guidelines on Data Management in Horizon 2020, Version 1.0, 11 December 2013).

## DMP Methodology

In keeping with the ECOPOTENTIAL principles, the ECOPOTENTIAL DMP describes how the different data types (generated and/or collected by the WP digital systems) are going to be managed in the project, in terms of:

- What standards will be used for these data discoverability, accessibility and (re-)use;
- How these data will be exploited and/or shared/made accessible for verification and reuse; if data cannot be made available, the reasons will be fully explained.
- How these data will be curated and preserved, even after the project duration.

For each data type recognized by the project (see the next section), the DMP considers a separate section consisting of the following 5 information strands to be filled:

1. Data set reference and name
  - a. Identifier for the data set to be produced.
2. Data set description (Description of the data that will be generated or collected)
  - a. Origin (in case it is collected);
  - b. Nature and scale;
  - c. To whom it could be useful;
  - d. Whether it underpins a scientific publication;
  - e. Information on the existence (or not) of similar data;
  - f. Possibilities for integration and reuse.
3. Standards and metadata
  - a. Reference to existing suitable standards of the discipline;
  - b. If these do not exist, an outline on how and what metadata will be created;
4. Data sharing (Description of how data will be shared)
  - a. Access procedures;
  - b. Embargo periods (if any);
  - c. Outlines of technical mechanisms for dissemination and necessary software;
  - d. Other tools for enabling re-use;



- e. Definition of whether access will be widely open or restricted to specific groups;
  - f. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.);
  - g. In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).
5. Archiving and preservation (including storage and backup)
    - a. Description of the procedures that will be put in place for long-term preservation;
    - b. Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered.

## ECOPOTENTIAL Data Types

ECOPOTENTIAL recognized five different Data types, in brackets the WPs mainly producing/using these data:

1. Remote Sensing (RS) Raw Data (WP3 and WP4);
2. Remote Sensing (RS) High-level Data (WP4);
3. In-situ data (WP5);
4. Model Results (WP6, WP7, and WP8);
5. Indicators/indexes (WP7, WP8, and WP9).

## Data Types covered by the present DMP version

**Time being, the Consortium reached an agreement on data management and sharing features only on the first three data types: RS Raw Data (WP3 and WP4), RS High-level Data (WP4), In-situ data (WP5).**

**The other two data types (i.e. Model Results and Indicators/indexes), managed by WPs 6, 7, 8 and 9, will be covered by the next version of the DMP.**





## ECOPOTENTIAL dataset category #1: RS Raw Data (WP3 and WP4)

For this kind of dataset (and associated software), ECOPOTENTIAL Consortium will adopt the following management strategy.

### Data set reference and name: Landsat Raw Data Products

Identifier for the data set to be produced

The original identifier will be kept for each Landsat product. The identifiers, according to their user's guides, are:

Product	ID	Elements	Values
Landsat MSS	LMSPPRRRRYYY DOYGSIVV_BN.XX X	<i>L = L → Landsat</i>	
		<i>M = Mission</i>	<i>M = Landsat MSS</i>
		<i>S = Satellite</i>	1 = Landsat 1 2 = Landsat 2 3 = Landsat 3 4 = Landsat 4 5 = Landsat 5
		<i>PPP = Three-digit starting path</i>	
		<i>RRR = Three-digit starting row</i>	
		<i>YYYY = Four-digit acquisition year</i>	
		<i>DOY = Three-digit acquisition day of year</i>	
		<i>GSI = Ground Station Identifier</i>	
		<i>VV = Two-digit version</i>	
		<i>BN = Product Component</i>	<i>B1 = Band 1</i> <i>B2 = Band 2</i> <i>B3 = Band 3</i> <i>B4 = Band 4</i> <i>B5 = Band 5</i> <i>B6 = Band 6</i> <i>B7 = Band 7</i> <i>GCP = GCP File</i> <i>VER = Verification File</i> <i>MTL = Metadata File</i>
		<i>XXX = File type</i>	<i>= TIF file extension for all image data</i> <i>= jpg file extension for the verification browse</i> <i>= txt file extension for GCP, VER, and Metadata (MTL) files</i>
		Landsat TM (NLAPS GeoTIFF)	LLNpppprrrOOYYD DDMM_AA.XXX
<i>N = Satellite number</i>	4 = Landsat 4 5 = Landsat 5		
<i>ppp = Starting path of the product</i>			
<i>rrr = Starting row of the product</i>			
<i>OO = WRS row offset (set to 00)</i>			



		<i>YY=Last two digits of the acquisition year</i>	
		<i>DDD=Julian date of acquisition</i>	
		<i>MM=Instrument mode: 50 = TM</i>	
		<i>AA=File Type</i>	<i>B1 = Band 1 B2 = Band 2 B3 = Band 3 B4 = Band 4 B5 = Band 5 B6 = Band 6 B7 = Band 7 WO = Processing history file</i>
		<i>XXX=File Extension</i>	<i>.tif = GeoTIFF file .txt = American Standard Code for Information Interchange (ASCII) text file</i>
<b>Landsat TM (LPGS GeoTIFF)</b>	LMSPPRRRRYYYY DOYGSIVV_BN.XX X	<i>L= Landsat</i>	
		<i>M=Mission</i>	
		<i>T = Landsat TM</i>	
		<i>S=Satellite</i>	<i>4 = Landsat 4 5 = Landsat 5</i>
		<i>PPP=Three-digit starting path</i>	
		<i>RRR=Three-digit starting row</i>	
		<i>YYYY=Four-digit acquisition year</i>	
		<i>DOY=Three-digit acquisition day of the year</i>	
		<i>GSI=Ground Station Identifier</i>	
		<i>VV=Two-digit version</i>	
		<i>BN=Product Component:</i>	<i>B1 = Band 1 B2 = Band 2 B3 = Band 3 B4 = Band 4 B5 = Band 5 B6 = Band 6 B7 = Band 7 GCP = GCP File VER = Verification File MTL = Metadata File</i>
		<i>XXX=File type</i>	<i>= TIF file extension for all image data = JPG file extension for the verification browse = .txt file extension for GCP, VER, and L1 Metadata (MTL) files</i>
		<b>Landsat 7</b>	



	L7CPFyyyyymmdd_ yyyyyymmdd.nn	<p><i>CPF = 3-letter CPF designator</i></p> <p><i>yyyy = 4-digit effectivity starting year</i></p> <p><i>mm = 2-digit effectivity starting month</i></p> <p><i>dd = 2-digit effectivity starting day</i></p> <p><i>_ = Effectivity starting/ending date separator</i></p> <p><i>yyyy = 4-digit effectivity ending year</i></p> <p><i>mm = 2-digit effectivity ending month</i></p> <p><i>dd = 2-digit effectivity ending day</i></p> <p><i>nn = Sequence number for this file</i></p>	
<b>Landsat 8</b>	Ls8pppprrrYYYYDD DGGGVV_FT.ext	<p><i>L= Landsat</i></p> <p><i>S= Sensor of: O = OLI, T = TIRS, C = Combined TIRS and OLI Indicates which sensor collected data for this product</i></p> <p><i>8= Landsat mission number</i></p> <p><i>ppp =Satellite orbit location in reference to the Worldwide Reference System-2 (WRS-2) path of the product</i></p> <p><i>rrr=Satellite orbit location in reference to the WRS-2 row of the product</i></p> <p><i>YYYY= Acquisition year of the image</i></p> <p><i>DDD= Acquisition day of year</i></p> <p><i>GGG= Ground station ID</i></p> <p><i>VV= Version</i></p> <p><i>_FT= File type, where FT equals one of the following: image band file number (B1–B11), MTL (metadata file), BQA (quality band file), MD5 (checksum file)</i></p> <p><i>.ext=File extension, where .TIF equals GeoTIFF file extension, and .txt equals text extension</i></p>	

### Data set description (Description of the data that will be generated or collected)

Landsat 1-5 MSS / Landsat 4-5 TM / Landsat 7 ETM / Landsat 8 OLI-TIRS

Origin (in case it is collected)

U.S. Geological Survey's Earth Resources Observation and Science (EROS, <http://eros.usgs.gov/>) Center or NASA's Land Processes Distributed Active Archive Center (LP DAAC)

Data model and scale

Raster at multiple scales (15 m, 30 m, 60 m, 100 m and 120 m)



### To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organizations

### Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

### Possibilities for integration and reuse

There are no restrictions on Landsat data downloaded from USGS EROS, and it can be used or redistributed as desired. However, a statement of the data source when citing, copying, or reprinting USGS Landsat data or images is requested. Standards and metadata

### Reference to existing suitable standards of the discipline

Level 1 products from all sensors contain the individual band files, as well and metadata files (some Landsat 4-5 scenes include a Work Order file, instead of the Metadata file, that contains the metadata information about scenes processed on the National Land Archive Production System (NLAPS).

#### *Landsat MSS / TM*

The MTL (metadata) file contains information specific to the product ordered. This file contains all applicable image description information. The MTL file complies with LSDS-524 Landsat Metadata Description Document (LMDD). (USGS/EROS. LSDS-524. Landsat Metadata Description Document (LMDD)).

#### *Landsat 7*

The MTL (metadata file) contains information specific to the product ordered. All calibration parameters are stored as American Standard Code for Information

Interchange (ASCII) text using the ODL syntax developed by JPL. ODL is a tagged keyword language developed to provide a human-readable data structure to encode data for simplified interchange.

#### *Landsat 8*

The metadata (MTL) file it is a human-readable text file in Object Definition Language (ODL) format. The complete description of the metadata file is provided in the Landsat 8 Level 1 Data Format Control Book (DFCB). In general, the MTL file includes the following parameters:

- Unique Landsat scene identifier
- WRS path and row information
- Scene Center Time of the date the image was acquired
- Corner longitude and latitude in degrees and map projection values in meters
- Reflective, thermal, and panchromatic band lines and samples
- File names included
- Image attributes including cloud cover, sun azimuth & elevation, number of ground control points used
- Band minimum and maximum reflectance and radiance rescaling



## Data sharing (Description of how data will be shared)

Open access.

### Access procedures

The data will be available to every partner member in the project via the FPT-CNR. Outside the consortium the data are available and accessible through the data provider facilities.

### Embargo periods (if any)

Not within the Consortium;

- Three-four months for use in paper submission on peer reviewed journals (if required)

### Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via the internationally accepted standards for exchange of satellite data and therefore can be used in all kinds GIS software and web mapping software capable of using OGC services.

### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors as well as the WP10 virtual platform as link to GEO/GEOSS.

In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

There are no restrictions on the use of data received from the U.S. Geological Survey's Earth Resources Observation and Science (EROS) Center or NASA's Land Processes Distributed Active Archive Center (LP DAAC) e.g. Landsat data downloaded from USGS EROS, and it can be used or redistributed as desired unless expressly identified prior to or at the time of receipt. However, a statement of the data source when citing, copying, or reprinting USGS Landsat data or images is requested. They request that the following statements be used when citing, copying, or reprinting data:

- *USGS Products: Data available from the U.S. Geological Survey.*
- *NASA Land Processes Distributed Active Archive Center (LP DAAC) Products: These data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at USGS/EROS, Sioux Falls, SD. <http://lpdaac.usgs.gov>*

## Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References:

- [http://landsat.usgs.gov/tools\\_project\\_documents.php](http://landsat.usgs.gov/tools_project_documents.php)
- <http://eros.usgs.gov/about-us/data-citation>



[http://landsat.usgs.gov/tools\\_project\\_documents.php](http://landsat.usgs.gov/tools_project_documents.php)

## Data set reference and name: MODIS AQUA/TERRA Data Products

Identifier for the data set to be produced

The original identifier will be kept for each MODIS product. The identifiers, according to their user's guides (see *Data set description* section for data set short names and specific product MODIS User guide). MODIS filenames follow a naming convention, e.g:

**MOD09A1.A2006001.h08v05.005.2006012234657.hdf**

**MOD09A1** - Product Short Name

**.A2006001** - Julian Date of Acquisition (A-YYYYDDD)

**.h08v05** - Tile Identifier (horizontalXXverticalYY)

**.005** - Collection Version

**.2006012234657** - Julian Date of Production (YYYYDDHHMMSS)

**.hdf** - Data Format (HDF-EOS)

## Data set description (Description of the data that will be generated or collected)

MODIS standard data products needed in the context of ECOPOTENTIAL project. The list of the available products is (MOD= TERRA / MYD= AQUA):

Data Type	Data Products
<b>Level 1</b>	MODIS Raw Radiances (M*D01) MODIS Calibrated Radiances (M*D021KM) MODIS Geolocation Fields (M*D03)
<b>MODIS Atmosphere Products</b>	MODIS Aerosol Product (M*D04) MODIS Total Precipitable Water (M*D05) MODIS Cloud Product (M*D06) MODIS Atmospheric Profiles (M*D07) MODIS Atmosphere Joint Product (M*DATML2) MODIS Atmosphere Gridded Product (M*D08) MODIS Cloud Mask (M*D35)
<b>MODIS Land Products</b>	MODIS Surface Reflectance (M*D09) MODIS Land Surface Temperature and Emissivity (M*D11) MODIS Land Cover Products (MCD12) MODIS Vegetation Index Products (NDVI and EVI) (M*D13) MODIS Thermal Anomalies - Active Fires (M*D14) MODIS Fraction of Photosynthetically Active Radiation (FPAR) / Leaf Area Index (LAI) (M*D15) MODIS Evapotranspiration (M*D16) MODIS Gross Primary Productivity (GPP) / Net Primary Productivity (NPP) (M*D17)



	MODIS Bidirectional Reflectance Distribution Function (BRDF) / Albedo Parameter (MCD43) MODIS Vegetation Continuous Fields (MOD44) MODIS Water Mask (MOD44W) MODIS Burned Area Product (MCD45)
<b>MODIS Cryosphere Products</b>	MODIS Snow Cover (M*D10) MODIS Sea Ice and Ice Surface Temperature (M*D29)
<b>MODIS Ocean Products</b>	MODIS Sea Surface Temperature MODIS Remote Sensing Reflectance (RRS_XXX) MODIS Sub-surface Chlorophyll-a Concentration (CHLOR_A) MODIS Diffuse Attenuation at 490 nm (KD490) MODIS Particulate Organic Carbon (POC) MODIS Particulate Inorganic Carbon (PIC) MODIS Fluorescence Line Height (FLH) (NFLH) MODIS Instantaneous Photosynthetically Available Radiation (IPAR) MODIS Daily Mean Photosynthetically Available Radiation (PAR)

Only the products needed in the context of ECOPOTENTIAL will be downloaded.

Origin (in case it is collected)

MODIS - NASA's GSFC. (LAADS, LP DAAC, NSIDC, OB DAAC) (<https://lpdaac.usgs.gov>)

Data model and scale

Raster at multiple scales (250 m, 500 m and 1000 m nominal spatial resolution) and at multiple temporal resolutions: Daily, 8-Day, 16-Day, Monthly, Quarterly, Yearly.

To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organizations

Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

Possibilities for integration and reuse

MODIS data and products are available free of charge through the LP DAAC (<https://lpdaac.usgs.gov>) and have no restrictions on subsequent use, sale, or redistribution. Standards and metadata

Reference to existing suitable standards of the discipline

*LAADS*

Metadata imbedded within each HDF product.

*LP DAAC*

MODIS products have two sources of metadata: the embedded HDF metadata, and the external ECS metadata. The HDF metadata contains valuable information including global attributes and dataset specific attributes pertaining to the granule. The ECS (generated by the EOSDIS Core System) .met file is the external metadata file



in XML format, which is delivered to the user along with the MODIS product. It provides a subset of the HDF metadata. Some key features of certain MODIS metadata attributes include the following:

- The Xdim and Ydim represent the rows and columns of the data, respectively
- The Projection and ProjParams identify the projection and its corresponding projection parameters
- The Sinusoidal Projection is used for most of the gridded MODIS land products, and has a unique sphere measuring 6371007.181 meters.
- The UpperLeftPointMtrs is in projection coordinates, and identifies the very upper left corner of the upper left pixel of the image data
- The LowerRightMtrs identifies the very lower right corner of the lower right pixel of the image data. These projection coordinates are the only metadata that accurately reflect the extreme corners of the gridded image
- There are additional BOUNDINGRECTANGLE and GRINGPOINT fields within the metadata, which represent the latitude and longitude coordinates of the geographic tile corresponding to the data

The Data Set attributes contain specific SDS information such as the data range and applicable scaling factors for the data. The LP DAAC data products page provides these details within a concise document for each of the products. An HDF-EOS file also contains EOS core metadata essential for EOS search services. Any tool that processes standard HDF files can read an HDF-EOS file. However, it is difficult for a standard HDF call to interpret HDF-EOS geolocation or temporal information without further knowledge of the file structure.

#### *NSDIC*

External metadata file in XML format.

#### *OB DAAC*

MODIS products have two sources of metadata: the embedded HDF metadata.

### Data sharing (Description of how data will be shared)

Open access

#### Access procedures

The data will be available to every partner member in the project via the FPT-CNR. Outside the Consortium the data are available and accessible through the data provider facilities.

#### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

#### Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via the internally accepted standards for exchange of data and therefore can be used in all kinds GIS software and web mapping software capable of using OGC services.

#### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors as well as WP10 virtual platform as link to GEO/GEOSS





In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

#### *LAADS*

All NASA Terra and Aqua MODIS comply with NASA's Data and Information policy, which promotes the full and open sharing of all data with the research and applications communities, private industry, academia, and the general public (the term data includes observation data, images, metadata, products, and documentation). All data products acquired through the Level-1 & Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC) are provided without any monetary charge, and have no restrictions on subsequent use or redistribution. However, they request to include an acknowledgment and citation.

#### *LP DAAC*

MODIS data and products acquired through the LP DAAC have no restrictions on subsequent use, sale, or redistribution. MODIS data and products are available at no charge from the LP DAAC.

#### *NSDIC*

NASA data are not copyrighted; however, as a condition of use, when you publish NSIDC NASA DAAC data or results derived from the data, they require citing the data sets in your work with a formal citation.

#### *OB DAAC*

They require citing the data sets in your work with a formal citation. Please provide acknowledgement of the use of OB.DAAC (Ocean Biology Distributed Active Archive Center) data products, images, and services in publications or presentations.

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP archive and GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References

- <https://ladsweb.nascom.nasa.gov/>
- <https://lpdaac.usgs.gov/>
- <http://nsidc.org/>
- <http://oceancolor.gsfc.nasa.gov/cms/>

---

### Data set reference and name: Sentinel Raw Data Products

Identifier for the data set to be produced

The original identifier will be kept for each Sentinel product. For specific details on the Sentinel naming convention, please refer to their user's guides:

*Sentinel-1(SAR)*

- Sentinel-1 Level-0: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-0-products/formatting>
- Sentinel-1 Level-1: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-product-formatting>
- Sentinel-1 Level-2: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-2-algorithms/formatting>

*Sentinel-2 (MSI)*

- <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/naming-convention>

*Sentinel-3 (OLCI)*

- <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-olci/naming-convention>

*Sentinel-3 (SLSTR)*

- <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-slstr/naming-convention>

*Sentinel-3 (Synergy)*

- <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-synergy/naming-conventions>

*Sentinel-3 (Altimetry-SRAL):*

- <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-altimetry/naming-conventions>

*Sentinel 4 / 5 / 5P (TROPOMI)*

TBD

## Data set description (Description of the data that will be generated or collected)

Sentinel-1 SAR / Sentinel-2 MSI / Sentinel-3 OLCI-SLSTR-Synergy-Altimetry\_SRAL / Sentinel-4/5/5P data and possible derivate products.

## Origin (in case it is collected)

ESA on behalf of the joint ESA/European Commission initiative GMES (Global Monitoring for Environment and Security).

## Data model and scale

Raster at multiple scales (10 m, 20 m, 60 m, 300 m, 500 m, 1000 m, 1200 m aprox)

## To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organizations

## Whether it underpins a scientific publication

Yes: a link to significant and/or related publications will be included in the dataset description, when they are available

## Possibilities for integration and reuse

Yes



## Standards and metadata

### Reference to existing suitable standards of the discipline

Sentinel-1 data is packaged as a file structure containing a manifest file in XML format listing general product metadata.

Sentinel-2 and 3 (OLCI- SLSTR-SYN) includes a *manifest.safe* file which holds the general product information in XML. The SAFE format wraps a folder containing image data in a binary data format and product metadata in XML.

Sentinel-3 (SRAL) is packaged as NetCDF and its attributes are used to store information about the data (ancillary data or metadata). The Common Data Language (CDL) will be used to describe the content of a data set. The CDL is textual notation that describes the netCDF object and it is readable by a human.

### Data sharing (Description of how data will be shared)

Open access.

### Access procedures

The data will be available to every partner member in the project via the FPT-CNR.

### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

### Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via the international standards for exchange of data and therefore can be used in all kinds GIS software and web-mapping software capable of using OGC services.

### Definition of whether access will be widely open or restricted to specific groups

Access will be defined according to the ESA policies (see <https://sentinel.esa.int/web/sentinel/home>)

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

The contents of ESA's Sentinel Online website are intended for non-commercial use. ESA grants permission to users to visit the site, and to download and copy information, images, documents and materials from the website for non-commercial use. ESA does not grant the right to resell or redistribute any information, documents, images or material from its website or to compile or create derivative works from material on its website. Use of material on the website is subject to the terms and conditions outlined below.

All material published on the ESA Sentinel Online website is protected by copyright and owned or controlled by ESA or the party credited as the provider of the content, software or other material. Copyright in the material must be recognised by an appropriate on-screen credit in a form such as "European Space Agency - ESA". If a remote sensing product is being used, the credit "produced from ESA remote sensing data" must be added, plus, if applicable, "image processed by (name of the institution as indicated on the Internet Server)".



Users may not modify, publish, transmit, participate in the transfer or sale of, reproduce, create derivative works from, distribute, perform, display or in any way exploit any of the content, software, material or services, in whole or in part, without obtaining prior written authorisation. In order to obtain authorisation to display or use any content of the ESA Sentinel Online website, please make a request for authorization through <https://earth.esa.int/web/guest/contact-us>.

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEO/GEOS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References:

- <https://sentinel.esa.int/web/sentinel/home>



## ECOPOTENTIAL dataset category #2: RS High-level Data (WP4)

For this kind of dataset (and associated software), ECOPOTENTIAL Consortium will adopt the following management strategy.

### Data set reference and name: Water turbidity (inland waters)(WT\_iw)

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming convention:

*xxn*-EO\_WT\_iw (xx=ISO country code, n=site number)

### Data set description (Description of the data that will be generated or collected)

Set of water turbidity maps in nefelometric units (NTU) generated using a statistical model calibrated and validated with field data using Landsat TM, ETM+ and OLI images.

### Origin (in case it is collected)

ECOPOTENTIAL Consortium Partners: EBD-CSIC, UAB.

### Data model and scale

Raster at Medium Resolution (30 m).

### To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organisations

### Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

### Information on the existence (or not) of similar data

Example of this data is currently accessible as WMS at <http://mercurio.ebd.csic.es/>.

### Possibilities for integration and reuse

Yes, by using OGC compliant standards and metadata.

### Reference to existing suitable standards of the discipline

General information of dataset will be included in a metadata form. The resulting metadata file contains the following information:

- title
- description and technical aspects
- spatial reference system (projection, datum and ellipsoid) and extent
- statistical model to derive the product
- contact information
- resource identifier (unique location where data is stored)



- other aspects like, date, lineage, language, authorization, copyright, acknowledgments, etc.

This general information will be collected and included in the metadata following the INSPIRE directive guidelines, using the ISO19115 metadata standard.

- INSPIRE directive guidelines, using the ISO19115 metadata standard
- Copernicus land monitoring services

### Data sharing (Description of how data will be shared)

Open access

#### Access procedures

The data will be available to every partner member in the project via the following OGC web services:

- WPS

#### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

#### Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via international standards for exchange of data and therefore can be used in all kinds GIS software and web mapping software capable of using OGC services.

#### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS. Also WMS at CSIC-Doñana.

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References:

- "Bustamante, J., Pacios, F., Díaz-Delgado, R. & Aragonés, D. (2009) Predictive models of turbidity and water depth in the Doñana marshes using Landsat TM and ETM+ images. *Journal of Environmental Management*, 90, 2219–2225.
- Díaz-Delgado, R., Amezttoy, I., Cristobal, J. & Bustamante, J. (2010) Long time series of Landsat images to reconstruct river surface temperature and turbidity regimes of Guadalquivir estuary. *Proceedings of the*



2010 IEEE International Geoscience & Remote Sensing Symposium on Remote Sensing (IGARSS2010)  
Geosciences and Remote Sensing Society."

---

### Data set reference and name: Land Cover/use (LCLU) and LCLU change maps, multiple sites

#### Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

#### Naming conventions:

*xxn*-EO\_LCLU (xx=ISO country code, n=site number)

and

*xxn*-EO LCLU Change maps

### Data set description (Description of the data that will be generated or collected)

Set of temporal LCLU maps provided by knowledge driven (Lucas et al., 2015) and data driven techniques, depending on the site characteristics (ground truth reference samples, site extension, site accessibility). The maps will be produced according to the Food and Agricultural Organization Land Cover Classification Taxonomy (FAO\_LCCS) or converted into FAP-LCCS from CORINE(in the case of pre-existing maps) .

#### Origin (in case it is collected)

ECOPOENTIAL Consortium Partners: CNR, UNSW, CERTH, ISPRA, CREAM, EURAC, UAB, EAA)

#### Nature and scale

Raster @ multiple scales, ranging from Medium Resolution to Very High Resolution (VHR) EO data integrated with ancillary data)

#### To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organisations

#### Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

#### Information on the existence (or not) of similar data

- CORE Copernicus products at Medium to High Resolution maps.
- Concerning VHR LCLU maps, these are not systematically produced from EO but are generally produced from costly in-field campaigns or visual interpretation of orthophotos for limited extents.

#### Possibilities for integration and reuse

Yes, adopting international standards like OGC/ISO.



## Standards and metadata

International standards like OGC/ISO

Reference to existing suitable standards of the discipline

- KEA (Clewley et al, 2015; Lucas et al. 2014; Bunting et al., 2013) and GeoTIFF
- GeoTIFF
- Copernicus land monitoring services

## Data sharing (Description of how data will be shared)

Open access

Access procedures

WPS

Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

Outlines of technical mechanisms for dissemination and necessary software

FTP and WCS

Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

## Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References

- **Lucas, R.M.,** Blonda, P., Bunting, P., Jones, G., Inglada, J., Aria, M., Kosmidou, V., Petrou, Z.I., Manakos, I., Adamo, M., Charnock, R., Tarantino, C., Mùcher, C.A., Jongman, R., Kramer, H., Arvor, D., Honrado, J.P. and Mairota, P. (2015). The Earth Observation Data for Habitat Monitoring (EODHAM) System. *International Journal of Applied Earth Observation and Geoinformation*. 37, 17-28.





### Data set reference and name: Habitat and habitat change maps, multiple sites

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_ANNEX I\_Ha of the Council Directive 92/43/EEC (xx=ISO country code, n=site number)

and

*xxn*-EO\_EUNIS\_Ha

*xxn*-EO\_GHC\_Ha (if needed)

*xxn*-EO\_Ha\_CHANGE

### Data set description (Description of the data that will be generated or collected)

- Set of temporal habitat maps expressed as ANNEX I and Eunis taxonomies. General Habitat Category (GHC) can also be produced (Adamo et al. 2014) if required by stakeholders.
- Set of habitat change maps

Origin (in case it is collected)

ECOPOTENTIAL Consortium Partners: CNR, UNSW, CERTH, ISPRA

Nature and scale

Raster @ multiple scales (from High Resolution and Very High Resolution EO data integrated with ancillary data)

To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments
- Non-governmental organisations

Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

Information on the existence (or not) of similar data

Generally produced from costly in-field campaigns or visual interpretation of orthophotos, but for limited extents.

Possibilities for integration and reuse

Yes, by adopting international and/or Community standards and metadata

Reference to existing suitable standards of the discipline

- KEA (Clewley et al, 2015; Lucas et al. 2014; Bunting et al., 2013) and GeoTIFF



- NATURA 2000 standard data form at

[http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/standarddataforms/notes\\_en.pdf](http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/standarddataforms/notes_en.pdf)

### Data sharing (Description of how data will be shared)

Open access

Access procedures

WPS

Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

Outlines of technical mechanisms for dissemination and necessary software

FTP and WCS

Definition of whether access will be widely open or restricted to specific groups

Open

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References

- Adamo M., Tarantino C., Tomaselli V., Kosmidou V., Petrou Z., Manakos I., Lucas R.M., Mucher C.A., Veronico G., Marangi C., De Pasquale V. and Blonda P., "Expert knowledge for translating land cover/use maps to General Habitat Categories (GHCs)" accepted for publication by Landscape Ecology. DOI: [10.1007/s10980-014-0028-9](https://doi.org/10.1007/s10980-014-0028-9).
- Bunting, P.; Gillingham, S. 2013. The KEA image file format. *Comput. Geosci.* 57, 54–58.
- Clewley D., Bunting P., Shepherd J., Gillingham S., Flood N., Dymond J., Lucas R., Armston J., Moghaddam M., 2014. A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA) Utilizing Raster Attribute Tables. *Remote Sensing*, 6, 6111-6135.
- **Lucas**, R.M., Blonda, P., Bunting, P., Jones, G., Inglada, J., Aria, M., Kosmidou, V., Petrou, Z.I., Manakos, I., Adamo, M., Charnock, R., Tarantino, C., Múcher, C.A., Jongman, R., Kramer, H., Arvor, D., Honrado, J.P. and Mairota, P. (2014). The Earth Observation Data for Habitat Monitoring (EODHAM) System. *International Journal of Applied Earth Observation and Geoinformation*. 37, 17-28. DOI: [10.1016/j.jag.2014.10.011](https://doi.org/10.1016/j.jag.2014.10.011)



### Data set reference and name: Soil Moisture (SM)

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_SM (xx=ISO country code, n=site number)

### Data set description (Description of the data that will be generated or collected)

Set of temporal SM maps provided by Synthetic Aperture Radar (e.g. Sentinel-1)

Origin (in case it is collected)

ECOPOTENTIAL Partner: CNR\_ISSIA, EURAC

Nature and scale

Raster @ 250-500m of resolution

To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers
- National and International governments

Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

Information on the existence (or not) of similar data

Generally produced from costly in-field campaigns, but for limited extents

Possibilities for integration and reuse

Yes, by adopting international and/or Community standards.

### Standards and metadata

Reference to existing suitable standards of the discipline

- GeoTIFF Int16, UTM projection

### Data sharing (Description of how data will be shared)

Access procedures

WCS and WPS

Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

Outlines of technical mechanisms for dissemination and necessary software

FTP, WCS and WPS



Definition of whether access will be widely open or restricted to specific groups

Open

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

Data will be shared

Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEOSS/GCI

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan Useful References

- Balenzano, Anna, et al. "Dense temporal series of C-and L-band SAR data for soil moisture retrieval over agricultural crops." Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 4.2 (2011): 439-450.
- Balenzano, Anna, et al. "SMOSAR algorithm for soil moisture retrieval using Sentinel-1 data." Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International. IEEE, 2012.
- Balenzano, Anna, et al. "On the use of temporal series of L-and X-band SAR data for soil moisture retrieval. Capitanata plain case study." European Journal of Remote Sensing 46.1 (2013): 721-737.

Data set reference and name: Seasonal Water Bodies (SWB)

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_SWB (xx=ISO country code, n=site number)

Data set description (Description of the data that will be generated or collected)

Set of temporal SWB maps provided by Synthetic Aperture Radar (sentinel-1) Data processing. The maps will be produced using a binary format (wetland/others)

Origin (in case it is collected)

ECOPOTENTIAL Partner: StarLab

Nature and scale

Raster, 20 m resolution



#### To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers

#### Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

#### Information on the existence (or not) of similar data

Copernicus global Land service products at High Resolution (20) maps.

#### Possibilities for integration and reuse

Yes, by using international and/or Community standards.

#### Standards and metadata

Reference to existing suitable standards of the discipline.

- Copernicus land monitoring services and Geo-TIFF

#### Data sharing (Description of how data will be shared)

Open access

#### Access procedures

- FTP

#### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

#### Outlines of technical mechanisms for dissemination and necessary software

- FTP

#### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

WP10 virtual platform as linkage to GEO/GEOSS

#### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan



---

### Data set reference and name: Snow Cover Area (SCA)

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_SCA (xx=ISO country code, n=site number)

### Data set description (Description of the data that will be generated or collected)

Set of temporal SCA maps provided by Synthetic Aperture Radar (sentinel-1) and Multispectral (MODIS, Sentinel-3) Data processing. The maps will contain values of snow cover, in percentage of covered area per pixel, with values between 0 (no snow) and 1 (100% snow covered).

Origin (in case it is collected)

ECOPOTENTIAL Partner: StarLab

Nature and scale

Raster , 500 m resolution

To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers

Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

Information on the existence (or not) of similar data

TBD in the next version of the DMP

Possibilities for integration and reuse

Yes, by using international and/or Community standards.

### Standards and metadata

Reference to existing suitable standards of the discipline

- NOAA standards (OGC, netCDF-CF, THREDDS, and OPeNDAP) and GeoTIFF

Data sharing (Description of how data will be shared)

Open access

Access procedures

- FTP

Embargo periods (if any)

Time being, NO embargo period is reported.



Outlines of technical mechanisms for dissemination and necessary software

- FTP

Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

WP10 virtual platform as linkage to GEO/GEOSS

In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

NA

Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOENTIAL CNR\_FTP and mirrors as well as WP10 virtual platform as link to GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Data will be preserved in the ECOPOENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and associated costs will be given in a successive update of the current data management plan.

Data set reference and name: Snow Cover Maps (SCM)

Identifier for the data set to be produced

A unique identifier will be given to each data object. Identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_SCM (xx=ISO country code, n=site number)

Data set description (Description of the data that will be generated or collected)

The data set is a classified raster which contains information on snow cover (class 1), snow free area (class 2), clouds (class 3), open water (4) and inland water bodies (5).

Origin (in case it is collected)

NA

Nature and scale

Raster at Medium Resolution (MR)

Spatial resolution: 250 m Nature and scale

To whom it could be useful

- Decision makers
- National and International governments



- Non-governmental organizations
- Environmental Agencies
- Research institute
- Civil protection
- Regional authorities

#### Whether it underpins a scientific publication

The dataset can be used freely for research purposes and publications always acknowledging the source and mentioning the disclaimer policy if any.

#### Information on the existence (or not) of similar data

MODIS standard products (MOD10 and MYD10) but with coarser spatial resolution (500 m).

#### Possibilities for integration and reuse

Yes, by using international and/or Community standards.

#### Standards and metadata

The data format provided is GeoTIFF.

Data are stored on EURAC servers and the information as well as metadata are available via the following OGC compliant services:

- CSW (Catalogue Service for the Web)
- Geoportal (<http://webgis.eurac.edu/geoviewer>)
- WMS (Web Mapping Service)

General information of dataset are included in a metadata form. The resulting metadata file contains the following information:

- title
- description
- contact information
- resource identifier (unique location where data is stored)
- other aspects like, date, lineage, language, authorization, copyright, acknowledgments, etc.

This general information will be collected and included in the metadata following the INSPIRE directive guidelines, using the ISO19115 metadata standard.

#### Reference to existing suitable standards of the discipline

Metadata follows the INSPIRE directive guidelines, using the ISO19115 metadata standard.

#### Data sharing (Description of how data will be shared)

Open access, under CC-BY license.

#### Access procedure

The data will be available to every partner member in the project via the following OGC web services:

- Metadata at EURAC Geonetwork
- Data on WMS (Web Mapping Service) and on EURAC Geoviewer

#### Embargo periods (if any)

No embargo period foreseen for present and future datasets





Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via the internally accepted standards for exchange of data and therefore can be used in all kinds GIS software and web-mapping software capable of using OGC services.

Other tools for enabling re-use

NA

Definition of whether access will be widely open or restricted to specific groups

Full, open and free access, under CC-BY.

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

- Regional Geoportal (Institutional geoportal of the Province of Bolzano)
- WMS (Institutional Web Mapping Service)

In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

NA

Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

EURAC is archiving and mirroring all relevant data within the scientificnet.org, in cooperation with the Free University of Bolzano. A LTDP policy is going to be defined during the 2016.

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Permanently: EURAC

Useful References

- Notarnicola, C.; Duguay, M.; Moelg, N.; Schellenberger, T.; Tetzlaff, A.; Monsorno, R.; Costa, A.; Steurer, C.; Zebisch, M. Snow Cover Maps from MODIS Images at 250 m Resolution, Part 1: Algorithm Description. *Remote Sens.* 2013, 5, 110-126.
- Notarnicola, C.; Duguay, M.; Moelg, N.; Schellenberger, T.; Tetzlaff, A.; Monsorno, R.; Costa, A.; Steurer, C.; Zebisch, M. Snow Cover Maps from MODIS Images at 250 m Resolution, Part 2: Validation. *Remote Sens.* 2013, 5, 1568-1587.
- G.Thirel, C. Notarnicola, M. Kalas, M. Zebisc, T.Schellenberger, A. Tetzlaff, M. Duguay, N.Mölg, P.Burek, A. de Roo, Assessing the quality of a real-time Snow Cover Area product for hydrological applications, *Remote Sensing of Environment*, Volume 127, December 2012, Pages 271–287

Data set reference and name: [Water Bodies Delineation \(WBD\)](#)

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:



*xxn*-EO\_WBD (*xx*=ISO country code, *n*=site number)

#### Data set description (Description of the data that will be generated or collected)

Set of temporal WBD maps provided by Synthetic Aperture Radar (sentinel-1) Data processing. The maps will be produced using a binary format (water/No-water)

#### Origin (in case it is collected)

ECOPOTENTIAL Partner: StarLab

#### Nature and scale

Raster, 20 m resolution

#### To whom it could be useful

- Researchers
- Management authorities of protected sites
- Decision makers

#### Whether it underpins a scientific publication

Yes. A link to significant and/or related publications will be included in the dataset description, when they are available

#### Information on the existence (or not) of similar data

Copernicus global Land service products at Medium Resolution (1 Km) maps.

#### Possibilities for integration and reuse

Yes, by adopting international and/or Community standards.

#### Standards and metadata

##### Reference to existing suitable standards of the discipline

- Copernicus land monitoring services and GeoTIFF

##### Data sharing (Description of how data will be shared)

Open access

##### Access procedures

FTP

##### Outlines of technical mechanisms for dissemination and necessary software

FTP

##### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

WP10 virtual platform as linkage to GEO/GEOSS



### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as linkage to GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and costs will be given in a successive update of the current data management plan.

---

### Data set reference and name: Maps of landscape measures indicating fragmentation or connectivity, multiple sites

Identifier for the data set to be produced

A unique identifier will be given to each data object; identifiers will be unique and do not change over time.

Naming conventions:

*xxn*-EO\_LANDSCAPE\_YYYYY (*xx*=ISO country code, *n*=site number, YYYYY=landscape measure code, e.g. PLAND, PD, SHAPE)

and

*xxn*-EO LCLU Change maps

### Data set description (Description of the data that will be generated or collected)

Set of landscape measures indicating the degree of habitat or land cover fragmentation or connectivity in a studied landscape. They will be calculated following a cell-based approach, dividing the scene in (overlapping) cells, calculating the measures therein, and aggregating the results for the whole study site, as followed in BIO\_SOS project to provide more localized results (Petrou *et al.*, 2013; Mairota *et al.*, 2013). Examples of calculated measures include (McGarigal *et al.*, 2002): percentage of landscape (PLAND), patch density (PD), shape index distribution (SHAPE), total class area (CA), mean patch size (MPS), effective mesh size (MESH), Area-Weighted Mean Patch Fractal Dimension (AWMPFD). Land cover and/or habitat maps, mainly provided through ECOPOTENTIAL mapping activities, will be used as primary input for the calculation of the measures.

Origin (in case it is collected)

ECOPOTENTIAL Consortium Partners: CERTH leads the subtask. ESL. FORTH will work on it for Samaria National Park

Nature and scale

Raster @ multiple scales, ranging from Medium Resolution to Very High Resolution (VHR) following the resolution of the input land cover and/or habitat maps

To whom it could be useful

- Researchers
- Management authorities of protected sites



- Decision makers
- National and International governments
- Non-governmental organisations

#### Whether it underpins a scientific publication

Yes. For landscape measures used in ECOPOTENTIAL publications, a link (preferably as DOI) to the publication will be part of the meta-data of the data record.

#### Information on the existence (or not) of similar data

Fragmentation and connectivity landscape measures are not systematically produced from EO data.

Relevant COPERNICUS data that may be used as input for their calculation, instead of ECOPOTENTIAL land cover and/or habitat maps, include: 1) High Resolution Layers: a) Tree Cover Density 2012 and b) Forest Type 2012; 2) CORINE Land Cover (2000, 2006, 2012).

#### Possibilities for integration and reuse

Yes, by using international and/or Community standards.

#### Standards and metadata

Reference to existing suitable standards of the discipline

- KEA (Clewley et al, 2015; Lucas et al. 2014; Bunting et al., 2013) and GeoTIFF
- Copernicus land monitoring services

#### Data sharing (Description of how data will be shared)

Open access

#### Access procedures

WPS: this standard interface was selected to be consistent with the (i) “Land Cover/use (LCLU) and LCLU change maps, multiple sites,” and (ii) “Habitat and habitat change maps, multiple sites” products.

#### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)

#### Outlines of technical mechanisms for dissemination and necessary software

FTP and WCS. “FTP and WCS” were selected to be consistent with the (i) “Land Cover/use (LCLU) and LCLU change maps, multiple sites,” and (ii) “Habitat and habitat change maps, multiple sites” products.

#### Other tools for enabling re-use

Any open source or commercial GIS software compatible with GeoTiff and KEA file formats and any custom-made or existing scripts or libraries in programming languages such as Matlab, R, Python, or Java processing such files, will be able to re-use the produced data.

#### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors as well as WP10 virtual platform as linkage to GEO/GEOSS



In case the dataset cannot be shared: the reasons for this should be mentioned (e.g. Ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

No reasons for restricting the open access to the data are identified, apart from any embargo period for scientific publications, as defined above.

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

ECOPOTENTIAL CNR\_FTP and GEO/GEOSS

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

The data will be preserved in the ECOPOTENTIAL FTP archive for at least 5 years after the end of the project. Details on the estimated end volume and costs will be given in a successive update of the current data management plan

### Useful References

- Petrou, Z., Manakos, I., Kosmidou, V., Lucas, R., Adamo, M., Tarantino, C., Blonda, P., 2013. Indicator extraction software: Software for indicator extraction including warning signal. FP7 BIO\_SOS project, Deliverable D6.12.
- Mairota, P., Cafarelli, B., Boccaccio, L., Leronni, V., Labadessa, R., Kosmidou, V., Nagendra, H., 2013. Using landscape structure to develop quantitative baselines for protected area monitoring. *Ecol. Ind.* 33, 82–95.
- McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: Spatial pattern analysis program for categorical maps. Comp. software prog. Univ. Mass., Amherst. Documentation. Available online at: [http://www.umass.edu/landeco/research/fragstats/documents/fragstats\\_documents.html](http://www.umass.edu/landeco/research/fragstats/documents/fragstats_documents.html) (Accessed on 06/11/2015).



## ECOPOTENTIAL dataset category #3: in-situ Data (WP5)

For this kind of dataset (and associated software), ECOPOTENTIAL Consortium will adopt the following management strategy:

### Data set reference and name: ECOPOTENTIAL in-situ Data

Identifier for the data set to be produced

Depending on the source and origin of the dataset we can differentiate between:

- Data created by the project consortium within the framework of the current ECOPOTENTIAL project, e.g. sampling campaigns funded by and related to the project work plan
- Data created by the project consortium outside the framework of the current ECOPOTENTIAL project, e.g. historic data

For the first group of datasets, a unique identifier will be given to each data object; identifiers will be unique and do not change over time. Depending on the data repository applied, the unique identifier can either be a handle (e.g. EPIC PID or compatible with the Digital Object Identifier (DOI) System) or a permanent URL (by the MD record). The MD record will contain the information on the PID.

For the second group, existing identifiers are maintained. So, if data are resulting from an existing data repository or storage (e.g. data management system of a protected area), the identifiers used in the local system will be maintained and used in the referencing.

All unique identifiers (e.g. PID) need to be linked to a metadata record describing the content and origin of the dataset or data series.

### Data set description (Description of the data that will be generated or collected)

Origin (in case it is collected)

ECOPOTENTIAL Consortium Partners and related protected areas. In addition, relevant data sources might be downloaded from existing open source data repositories.

### Nature and scale

Content, scale and resolution of the datasets depend on the specified scientific analysis as defined in the story lines. An overview on relevant in-situ datasets is given in Table 1. A detailed list of datasets needed will be developed within the runtime of the project together with other relevant workpackages.

A catalogue on existing datasets (on the basis of aggregated parameters) on the level of the protected or research areas will be developed within the project and published using DEIMS Site Catalogue (<https://data.iter-europe.net/deims/research-site>). This will allow a first level evaluation of data availability.

Table 1 Exemplary overview on relevant in-situ data for the project

Level: Atmosphere	
Domain	<b>Parameter</b>
Meteorological	Weather stations data (precipitation, temperature, humidity, wind) Pan evapotranspiration measurements Solar irradiance measurements
Climatic	Air temperature Wind speed Wind direction Water vapor



Level: Geosphere Domain Geological	Pressure Precipitation Surface radiation budget Backscattering coefficient
	<b>Parameter</b> Stratigraphic Map Petrographic Map
Level: Pedosphere Domain Pedological	<b>Parameter</b> Soil type Soil chemistry (e.g. C, N) Soil depth Soil carbon Decomposition rate Nutrient retention Soil characteristics (e.g. silt/sand/clay, permeability, erodibility, moisture)
	Topography Ground Control Points (accurately known locations for ortorectification) Elevation range (m) Min. elevation (m asl) Max. elevation (m asl) Slope Aspect
Digital elevation data	LIDAR data STRM
Level: Hydrosphere Domain Rivers/streams	<b>Parameter</b> Stream gauge/pressure sensor data (streamflow) Tracer experiments data Water chemistry (e.g. dissolved O <sub>2</sub> , DOC, nitrates, chemical substances) Water temperature Sediments (concentration, characterization) Groundwater River discharge Inundation
	Lakes Number of lakes Area of lakes Water chemistry (e.g. dissolved O <sub>2</sub> , DOC, nitrates, chemical substances) Water depth / bathymetry Water clarity (e.g. turbidity / Secci depth) Water temperature Sediments (concentration, characterization) Groundwater
Sea	Sea Surface Temperature Currents (velocity, direction, regime) Eddies, Upwellings (regime, location, temperature) Water chemistry (e.g. dissolved O <sub>2</sub> , DOC, nitrates, chemical substances) Sediments (concentration, characterization) Water depth / bathymetry Water clarity (e.g. turbidity / Secci depth)
Glaciers and ice caps	Ice sheets (regime, area, density, etc) Permafrost (thermal regime, density, etc) Snow cover Albedo Sea Surface Temperature
Level: Biosphere Domain Vegetation	<b>Parameter</b> Point species occurrence data / species lists / census plots



Fauna (terrestrial or aquatic)	Coverage Distribution Primary productivity Traits (e.g. phenology, phenologic variability, fraction of absorbed photosynthetically active radiation (FAPAR), above-ground biomass, below-ground biomass, litter (standing and fallen dead biomass), leaf nitrogen content, leaf area index (LAI), specific leaf area) Plant Communities traits Disease / pathogens occurrence / spread data NDVI (ground truth)
	Species occurrence data / presence / absences / species abundance Distribution Traits data (e.g. body mass, foraging behavior, life-history habitat requirements) Population dynamics (e.g. breeding success rate, age structure) Animal movement data (from GPS tracks/ observations) Home range
Biodiversity	Biodiversity status Biodiversity maps Ecosystem diversity Habitat diversity Functional diversity Taxonomic diversity Species diversity Allelic diversity Genetic diversity
General	Ecosystem type Fire occurrence Inundation Decomposition rate Disturbance Nutrient retention
Level: Anthroposphere	
Domain	<b>Parameter</b>
Information about the Protected Area	Year of creation Contact Person Total Area (Area [ha] that has been protected at a given time)
Management	Management plans Protection Status
Infrastructure	Maps of trails, roads, paths Access restrictions (e.g. fenced areas) Man-made structures (e.g. dykes, dams, buildings, bridges)
Ecosystem Services	Cultural ESS (e.g. visitors numbers / recreational activities surveys) Provisioning ESS (e.g. fish catch / herds size / crops and yields) Census of local population (e.g. population, age, demographics, socio-economic data) Use of other natural resources or environment (e.g. mining concessions, wind turbines)

The information is divided in the different spheres of the Earth: atmosphere, hydrosphere, geosphere, biosphere and anthroposphere. The parameters aggregated in corresponding domains can be applied to each protected area. However, this table is not restricted; more data not shown here can also be added.

In general, the temporal parameter resolution should be as high as possible, i.e. min, max and annual, but also down to minutes, days etc. In addition, the spatial resolution of the parameter should be as high as possible, i.e. km, meters and/or meaningful regions (ecosystems, administrative areas, etc).





### To whom it could be useful

- Researchers
- Management authorities of protected sites / areas
- Decision makers
- National and International governments
- Non-governmental organizations

### Whether it underpins a scientific publication

A link to significant and/or related publications will be included in the dataset description, when they are available

A link (preferably as DOI) to the publication will be part of the meta-data of the data record. The data can be used freely for research purposes and publications always acknowledging the source and mentioning the disclaimer policy if any.

### Information on the existence (or not) of similar data

If similar data exists, a reference or a link to such data will be provided in the metadata.

### Possibilities for integration and reuse

Yes. Further research can be conducted using this data. Standards adoption will facilitate this task.

Since the data will be free, it can be reused or integrated into tools, software, websites, infographics, policy briefs, decision-support systems etc. Proper description of the dataset by content related metadata will ensure the reusability of the data.

### Standards and metadata

Reference to existing suitable standards of the discipline.

Specification of metadata models

Research locations and datasets will be – if not otherwise existing in a proper format – described using the following standards:

Datasets<sup>1</sup>:

- INSPIRE MD regulation (based on ISO19115/19139)
- EML (Ecological Metadata Language)

Research locations<sup>2</sup>

- Site documentation developed by LTER Europe
- INSPIRE EF Data Model (Environmental Monitoring Facility) – work in progress

General information of dataset will be included in a metadata form. The resulting metadata file contains the following information:

- title
- description and technical aspects

---

<sup>1</sup> See <https://data.lter-europe.net/deims/dataset/documentation>

<sup>2</sup> See <https://data.lter-europe.net/deims/research-site/documentation>



- temporal and spatial extent (including aspects on spatial reference system (projection, datum and ellipsoid))
- method description (e.g. statistical model to derive the product)
- contact information
- resource identifier (unique location where data is stored)
- other aspects like, date, lineage, language, authorization, copyright, acknowledgments, etc.

A more detailed documentation of the MD standards and fields applied can be found at the DEIMS Metadata catalogue at <http://data.lter-europe.net/>. Within WP5 of ECOPOTENTIAL extensions of this MD models will be discussed based on the needs within the project.

MD model implementation and best practice examples will be published open to the community.

#### Specification of data format

A harmonization of newly collected data will be done, as far as possible, based on the OGC O&M standard. Where OGC O&M could not be applied, alternative data models for the data exchange will be developed within ECOPOTENTIAL. Specification of the data models will be shared openly with the datasets.

#### If these do not exist, an outline on how and what metadata will be created

For each final in-situ dataset created and deposited within the context of ECOPOTENTIAL a basic set of metadata need to be created and shared using an existing platform. If not otherwise specified the DEIMS Site and Dataset Registry<sup>3</sup> will be used. Both, the research location (e.g. a protected area) as well as the dataset generated (e.g. soil sample) can be described based on the MD standards based above.

The metadata will be generated by the user when depositing the data. Minimal metadata need to be provided, e.g.

- data provider (e.g. institute creating the data)
- metadata provider (e.g. contact information for person depositing the data)
- dataset title
- short description of the dataset
- etc.

A more detailed description can be found at <https://data.lter-europe.net/deims/dataset/documentation>. The MD record will point to the resource location (e.g. an internal or external data repository). Metadata will be added by user before depositing into Data Repository.

#### Data sharing (Description of how data will be shared)

In general, an open access policy to data directly generated within the realm of the ECOPOTENTIAL project will be applied. Existing data policies for data generated outside the realm and funding of the ECOPOTENTIAL project will not be affected.

Nevertheless, fostering an open access policy for science related data will be one of the activities within ECOPOTENTIAL.

---

<sup>3</sup> See <http://data.lter-europe.net/deims/>



### Access procedures

Depending on the origin of the data, different data access points and access procedures will be applied. Datasets generated and stored within the ECOPOTENTIAL domain will be accessible by the following options:

- File based datasets (e.g. NetCDF) will be available to every partner member in the project via the FPT-CNR or related data repository (e.g. EUDAT B2SHARE instance).
- Spatial data (e.g. habitat maps) will be available to every partner member in the project via OGC web services, e.g. WMS, WCS or WFS. . The final list on data services will be defined based on the data specification needed.
- Time series data (e.g. temperature sensors) will be available to every partner member in the project via OGC web services, e.g. SOS. The final list on data services will be defined based on the data specification needed.

In addition, linkage to existing local or institutional data repositories will be made.

- The data will be deposited either in institutional research data repositories or public data repositories

As a general rule, the use of web services for data sharing will be fostered. If applicable, the data will be available to every partner member in the project via web frontend and OGC web services (if implemented in repository).

A common metadata catalogue to discover and access data will be developed via the GEOSS data portal tools (see WP10).

### Embargo periods (if any)

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required). To be further clarified in the next version of the DMP.

### Outlines of technical mechanisms for dissemination and necessary software

Data is disseminated via the accepted standards for exchange of data in order to ensure interoperability. This includes e.g. the use of the data in all kinds statistical (e.g. R) and GIS software (e.g. ArcGIS, GoogleEarth or QuantumGIS).

For spatial data existing data specifications, e.g. INSPIRE data models, will be evaluated for their usability and applicability within the scientific workflows. For file-based data, common data models will be defined based on best practice examples.

Wherever possible web services will be used to share data, in order to avoid duplicate storage of data. Metadata records will point to the service endpoints for the user. Direct linking e.g. of SOS services to R Statistical software will be evaluated.

Metadata will be published providing a standard CSW (e.g. as provided by geonetwork) in order to allow sharing and harvesting of metadata.

### Other tools for enabling re-use

Statistical software and corresponding scripting languages like R, SPSS, Matlab, Python, Ruby, Java

### Definition of whether access will be widely open or restricted to specific groups

Full, open and free access. The data will be widely open, unless explicitly stated a restriction to specific groups.



Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

ECOPOTENTIAL CNR\_FTP and mirrors and WP10 virtual platform as link to GEO/GEOSS

In addition, the use of EUDAT B2SHARE repository (will be setup by FZ Jülich) will be evaluated for the use of ECOPOTENTIAL. EUDAT B2SHARE will support PID provision to uploaded datasets.

A database with in-situ data collected in ground or from sensors (e.g. air temperature) is proposed. For this the implementation of a virtual data node (e.g. <https://github.com/SP7-Ritmare/starterkit>) will be evaluated within the project context.

For datasets not generated by the ECOPOTENTIAL project the access and sharing rules of the providing organization needs to be applied. This is also true if these rules include rights on derived products.

### Archiving and preservation (including storage and backup)

Description of the procedures that will be put in place for long-term preservation

- GEO/GEOSS , LifeWatch
- EUDAT B2SHARE repository will be evaluated

Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

Details on the long term preservation of the data, the estimated end volume and costs will be given in a successive update of the current DMP.

### Data set reference and name: LifeWatch-ITA Data Products

The ECOPOTENTIAL partner UNILE can take advantage from the Data Management facilities of the LifeWatch-ITA Joint Research Unit, including the Data Portal and Service Centre facilities. The following management strategy of the LifeWatch-ITA Data Portal and Service Centre will be available for the in-situ Data produced by WP5.

### Identifier for the data set to be produced

LifeWatch-ITA Data Repository uses the original global unique identifier if it exists; alternatively, the identifier is generated according to the LifeWatch guidelines.

### Data set description (Description of the data that will be generated or collected)

#### Origin (in case it is collected)

ECOPOTENTIAL Consortium, LifeWatch Italy Network, other data providers as national and international networks, initiatives, organisations, administrations.

#### Nature and scale

LifeWatch-ITA Data Repository can receive biodiversity and ecosystem data, including species traits data with grain size from individual to species. LifeWatch-ITA Data Repository can import CSV, DarwinCore and ABCD formats and can export in CSV format or supply REST interface using JSON. Scale will be part of the metadata description.



## To whom it could be useful

### User Groups:

- Research
- Administration
- Education
- Citizen Science
- Others

## Whether it underpins a scientific publication

Links to useful publications will be included in the field “Associated publications” in the metadata description, preferably as DOI if it exists or in alternative providing the reference according to the BibTeX Style Examples (<https://verbosus.com/bibtex-style-examples.html>).

## Possibilities for integration and reuse

Each Dataset can be reachable via REST web services using JSON. It makes the dataset interoperable and reusable.

## Standards and metadata

Reference to existing suitable standards of the discipline.

The LifeWatch-Italy Data Management System would provide the infrastructure for the curation, protection, discovery, access, and analysis of data resources on Biodiversity and Ecosystems in support of multi-user needs.

The LifeWatch-Italy Data Management System is based on an Extensible Observation Ontology (OBOE; Madin et al., 2007 Ecological Informatics 2), which has been expanded to the spatial/geographic context and the thematic dominion and connected to the LifeWatch Thesauri Catalogue. The goal is to serve as a formal and generic conceptual framework for describing the semantics of observational data sets (i.e., data sets consisting of observations and measurements). The LifeWatch methodology will also prescribe a structured approach for organizing domain specific ontologies through the use of “extension points”. The LifeWatch extension points allow ontology classes, properties, and constraints to be easily defined for a particular domain specific terminology, and existing domain extensions to be interrelated. Thus, LifeWatch model can serve as a framework for defining new domain ontologies as well as interoperating and relating existing ones.

The LifeWatch Data Management System consists of two major components: LifeWatch Metadata Standard and LifeWatch Data Standard.

The LifeWatch Metadata Standard, related to the **dataset description**, follows the Dataset Metadata Model of DEIMS (Drupal Ecological Information Management System; LTER), which provides a unified framework for ecological information management for LTER sites or similar research groups, and it is compliant with the Ecological Metadata Language (EML) and the INSPIRE Directive. The LifeWatch Metadata Standard provides a list of the metadata elements necessary for the description of resources collected within the LifeWatch domain on the dataset level. Each metadata element is described in a tabular form and provides following details:

- The name of LifeWatch metadata element
- Metadata type
- Metadata description given by DEIMS specification based on EML (Ecological Metadata Language) specification and amended by LifeWatch expert community
- Obligation/condition for the metadata element (Required or Optional)



- Metadata Reference (e.g. <http://...>)
- Metadata category

The metadata will be created using the Dataset MetaData Editor which provides entry forms for authorised users to create metadata description in accordance with LifeWatch Metadata Specification ([www.servicecentrelifewatch.eu](http://www.servicecentrelifewatch.eu)).

The LifeWatch Data Standard, related directly to the **data**, provides the relevant Meta information about the dataset fields (e.g. Dataset Field Name, Description, Dataset Field DataType, Unit of measure, Dataset Field Standard etc.). At the state of the art it includes the Darwin Core standard and Controlled Vocabularies:

**Darwin Core**, includes a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biodiversity primary data by providing reference definitions, examples, and commentaries. The Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

**Controlled Vocabularies (Thesauri)** consists of lists of standardised terms that cover a broad spectrum of disciplines of relevance to the LifeWatch community. Using standardised sets of terms solves the problem of ambiguities associated with data markup and also enables records to be interpreted by computers. Use of common vocabularies in all metadatabases and data formats is an important prerequisite towards consistency and interoperability. LifeWatch uses existing thesauri such as EnvThes3, LusTRE, CAST, SeaDataNet Vocabulary, ecc..., and thesauri implemented by LifeWatch community concerning functional traits of aquatic organism and alien trait. According to the state of the art of SKOS/RDF thesauri, links among existing thesauri and LifeWatch thesauri are established (e.g. exact match, close match, narrower match, broader match, related match).

## Data sharing (Description of how data will be shared)

### Access procedures

After all data have been collected and metadata have been generated, the dataset will be shared using the LifeWatch-Italy Data Portal. There are two types of search engines to discover the dataset using Graphical User Interfaces: a geographical search using a map and a parameters search (Scientific name, Country, traits type, etc). Data will be also accessible via on-line services, including, at minimum, direct download but preferably user-customizable services for visualization and computation. Data access and use conditions, including licenses, will be clearly indicated in the metadata.

ECOPOTENTIAL Consortium data will be free and open access, according to the Consortium data policy

### Embargo periods (if any)

Before to be published the dataset is stored in a “staging area”. In this area a set of LifeWatch tools will support the researcher and the LifeWatch scientific team to check the quality of the data (tools for taxonomic, numeric and semantic checks). The researchers can verify consistency, accuracy, and precision of data. After the “staging period”

For ECOPOTENTIAL Consortium data, including those made available from the LifeWatch-ITA network

- Not within the Consortium;
- Three-four months for use in paper submission on peer reviewed journals (if required)



#### Outlines of technical mechanisms for dissemination and necessary software

The LifeWatch-Italy Data Portal is a Web application and it is a single point of access shared among all partners. There are two types of search engines to discover the dataset using Graphical User Interfaces: a geographical search using a map and a parameters search (Scientific name, Country, traits type, etc.).

#### Other tools for enabling re-use

Data will be also accessible via on-line services, including, at minimum, direct download but preferably user-customizable services for visualization and computation.

#### Definition of whether access will be widely open or restricted to specific groups

ECOPOTENTIAL Consortium data will be free and open access, according to the Consortium data policy

Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.)

The institutional repository of LifeWatch-Italy.

#### Archiving and preservation (including storage and backup)

##### Description of the procedures that will be put in place for long-term preservation

LifeWatch preserve in its infrastructure all the datasets (and the associated metadata) generated and it will be responsible for providing storage and backup resources. The dataset will be always available to the scientific community via the LifeWatch facilities (Data Portal, Virtual Lab, etc.). Specific storage, backup, replication and versioning policies will be described in more detail in the LifeWatch Disaster Recovery Plan. The database will be long-term preserved, hosted in the LifeWatch-ITA Data Repository infrastructure based in University of Salento (Italy). Being LifeWatch a European Research Infrastructure, in principle the data preservation shall be guaranteed for an indefinite period of time.

##### Indication of how long the data should be preserved, its approximated end volume, the associated costs and how these are planned to be covered

LifeWatch-ITA Data Repository, as member of a European Research Infrastructure, can commit to long-term sustainability of the ECOPOTENTIAL data, well beyond the completion of the project, independently on the data volume. Data preservation is free of charge for the data provider, i.e. the ECOPOTENTIAL Consortium.



## ECOPOTENTIAL dataset category #4: Model Results (WP6, WP7, and WP8)

THIS DATA CATEGORY WILL BE FIRST COVERED IN THE NEXT VERSION OF THE DMP.





## ECOPOTENTIAL dataset category #5: Indicators/indexes (WP7, WP8, and WP9)

THIS DATA CATEGORY WILL BE FIRST COVERED IN THE NEXT VERSION OF THE DMP.



## Annex A: GEOSS Data Sharing Principles

The following data sharing principles were adopted by GEOSS.

DSP1. There will be **full and open exchange** of data, metadata and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation;

DSP2. All shared data, metadata and products will be made available with **minimum time delay** and at **minimum cost**;

DSP3. All shared data, metadata and products being **free of charge or no more than cost of reproduction** will be encouraged for research and education.



## Annex B: GEOSS Data Management Principles

The following management principles were introduced by the GEOSS IIB (Infrastructure Interoperability Board) and will be adopted at the next GEO Plenary in Mexico City (November 2015).

### Discoverability

DMP-1. Data and all associated metadata will be discoverable through catalogues and search engines, and data access and use conditions, including licenses, will be clearly indicated.

### Accessibility

DMP-2. Data will be accessible via online services, including, at minimum, direct download but preferably user-customizable services for visualization and computation.

### Usability

DMP-3. Data will be structured using encodings that are widely accepted in the target user community and aligned with organizational needs and observing methods, with preference given to non-proprietary international standards.

DMP-4. Data will be comprehensively documented, including all elements necessary to access, use, understand, and process, preferably via formal structured metadata based on international or community-approved standards. To the extent possible, data will also be described in peer-reviewed publications referenced in the metadata record.

DMP-5. Data will include provenance metadata indicating the origin and processing history of raw observations and derived products, to ensure full traceability of the product chain.

DMP-6. Data will be quality-controlled and the results of quality control shall be indicated in metadata; data made available in advance of quality control will be flagged in metadata as unchecked.

### Preservation

DMP-7. Data will be protected from loss and preserved for future use; preservation planning will be for the long term and include guidelines for loss prevention, retention schedules, and disposal or transfer procedures.

DMP-8. Data and associated metadata held in data management systems will be periodically verified to ensure integrity, authenticity and readability.

### Curation

DMP-9. Data will be managed to perform corrections and updates in accordance with reviews, and to enable reprocessing as appropriate; where applicable this shall follow established and agreed procedures.

DMP-10. Data will be assigned appropriate persistent, resolvable identifiers to enable documents to cite the data on which they are based and to enable data providers to receive acknowledgement of use of their data.



## Annex C: GEOSS Architecture Principles

The following architectural principles were introduced by the GEOSS IIB (Infrastructure Interoperability Board) and are under discussion.

- Given the nature of a “system of Systems” it was recognized that the **success would depend on building interoperability among the different and autonomous systems.**
- As the basis for evolution and ensure interoperability with relevant research and policy-driven data infrastructures
  - **Openness**
  - **Effectiveness**
  - **Flexibility**
  - **Sustainability**
  - **Reliability**
- Support the implementation of the **Data Sharing and Management principles.**



## Annex D: Guidelines on Data Management in H2020

Any Research Data and the associated Software must be:

- **Discoverable**
  - readily **located**
  - standard **identification** mechanism
- **Accessible**
  - **modalities, scope, licenses**
- **Assessable and intelligible**
  - scientific **scrutiny**
  - peer **review**
- **Useable beyond the original purpose for which it was collected**
  - safely stored in **certified repositories**
  - **long term preservation** and **curation**
  - **minimum software, metadata** and **documentation** to make it useful
- **Interoperable to specific quality standards**
  - standards for **data annotation** and **data exchange**
  - **recombinations with different datasets** from different origins



## Annex E: Glossary

These terms are defined by the “Guidelines on Data Management in Horizon 2020”, version 1.0, 11 December 2013.

### **Data set reference and name**

Data set reference and name Identifier for the data set to be produced.

### **Data set description**

Description of the data that will be generated or collected, its origin (in case it is collected), nature and scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or not) of similar data and the possibilities for integration and reuse.

### **Standards and metadata**

Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created.

### **Data sharing**

Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.). In case the dataset cannot be shared, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

### **Archiving and preservation (including storage and backup)**

Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered.



## References

- Commission-European-DG-R&I. (2015, October 01). *Guidelines on Data Management in H2020*. Retrieved from [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)
- GEO. (2015, October 01). *GEOSS Data Management Principles*. Retrieved from [http://www.earthobservations.org/documents/dswg/201504\\_data\\_management\\_principles\\_long\\_final.pdf](http://www.earthobservations.org/documents/dswg/201504_data_management_principles_long_final.pdf)
- GEO. (2015, October 01). *The GEOSS Data Sharing Principles*. Retrieved from [https://www.earthobservations.org/geoss\\_dsp.shtml](https://www.earthobservations.org/geoss_dsp.shtml)
- GEOSS-IIB. (n.d.). *GEOSS requirements, functions, architecture*. Retrieved from GEO Web site: [ftp://ftp.earthobservations.org/201505\\_Work\\_Plan\\_Symposium/presentations/](ftp://ftp.earthobservations.org/201505_Work_Plan_Symposium/presentations/)