



MOZART

WP 8: Deliverable D8.2 “Data Management Plan”

Responsible Partner	University of Birmingham (UoB)
Author(s)/Editor(s)	Cris Rocca (UoB), Dimosthenis Toliopoulos (PoliMi), Eva Valsami-Jones (UoB)
Dissemination Level	Public
Type	R — Document, report
Due Date	30/11/2022 (M06)
Submission Date	30/11/2022



Project Profile

Programme	Horizon Europe
Call	"A Digitized, Resource-Efficient and Resilient Industry 2021"
Topic	HORIZON-CL4-2021-RESILIENCE-01-12
Number	101058450
Acronym	MOZART
Title	Metal matrix nano-composite coatings utilization as alternative to hard chromium
Start Date	01.06.2022
Duration	42 months

Document History

Version	Date	Author	Remarks
1.0	17/09/2022	UoB (Cris Rocca)	1 st draft
1.4	21/10/2022	MOZART WPLs	Comments provided
1.5	25/10/2022	UoB (Cris)	Comments included
1.6	16/11/2022	UoB (Cris, Eva)	Harmonisation
1.7	24/11/2022	UoB (Eva)	Review
1.8	28/11/2022	PoliMi (Dimosthenis)	Final review by PM



Executive Summary

The Data Management Plan (DMP) of project MOZART aspires to ensure that project data are made findable, accessible, interoperable, reusable (FAIR) and secure. An important priority for the research data management will be to reduce the replication of research activities within the EU and the possibility of data loss as well as to achieve less time-consuming and more cost-effective research and development work in the areas of hard chromium (HC) replacement.

The DMP is aligned with the dissemination and exploitation activities while clarifying sensitive and important aspects such as IPR handling, backup storage, and archiving. The data produced within the MOZART project as well as the processes of their management and collection will be described. The access rights to the data, the level of confidentiality, and their management during the implementation and post-project period as well as the preservation, the identification of repositories and possible IPR issues will be analyzed in the DMP. Data that will be considered confidential for the commercialization and success of the project will be preserved in restricted access. Regarding the management of research data in post-project activities, the consortium will define the data that will be maintained, how they may be reused by third parties and the timeline plan for these data. Moreover, the findability and interoperability of the data will be ensured with the utilization of the appropriate metadata described by the partners, the naming convention for each dataset, the type of format (.docx, .xls, measurements etc.) with simultaneous respect to the GDPR principles. Such activities aim to make project data easily findable and their interoperability within the scientific community will be facilitated. The accessibility and reusability of the data will be facilitated through offer of appropriate types of licenses, the release of specific software to access the data, the access routes for the data (open or restricted) as well as active data filing in repositories such as Zenodo. The DMP will be continuously updated during the project by defining the type of research data that will be produced and the policies that apply to the data.



Table of Contents

1	Introduction	7
2	Data Summary	8
3	FAIR data	13
3.1	The FAIR principles	13
3.1.1	Findable	13
3.1.2	Accessible	13
3.1.3	Interoperable	13
3.1.4	Reusable	13
3.2	Making data findable, including provisions for metadata	14
3.2.1	Making data openly accessible	14
3.2.2	Making data interoperable	16
3.2.3	Making data re-use (through clarifying licenses)	17
4	Other research outputs	18
5	Allocation of resources	18
6	Data security	19
7	Ethics	20
8	Other issues	20



List of Figures

Figure 1 *Project structure* 8

List of Tables

Table 1 Data source within WPs..... 9



Table of Abbreviations

Abbreviation	Explanation
HC	Hard Chromium
EC	European Commission
EU	Europe
GDPR	General Data Protection Regulation
SSbD	Safe and Sustainable by Design
DMP	Data Management Plan
DoA	Description of Action
FAIR	Findable, Accessible, Interoperable, and Reusable
EC	European Commission
GA	Grant Agreement
PQP	Project Quality Plan
Dn.m	Deliverable
WP	Work Package
WPL	Work Package Leader
SEM-EDX	Scanning Electron Microscopy - Energy Dispersive X-Ray Analysis
XRF	X-ray fluorescence
DLS	Dynamic light scattering
ICP-MS	Inductively coupled plasma mass spectrometry



1 Introduction

The MOZART Data Management Plan (DMP) is a key deliverable (D8.2) for work package 8 (WP8), Dissemination, Exploitation & Communication, and the whole project.

For the project members, the PQP will be a handbook with instructions on, for example, time plans and reviewing of deliverables. The reviewing process purpose is to ensure the quality of data, analysis and reports being submitted to the EC. This process will be improved continuously.

The MOZART consortium consists of 15 partners from 8 different countries. A broad range of data is going to be generated and/or collected (e.g., lists of stakeholders, reports, websites, flyers, oral presentations, laboratory and research outputs etc.) during the progress of the project.

In line with the FAIR principles, the project will assure that the data, collected and generated, is Findable, Accessible, Interoperable, and Reusable. The DMP will describe how the project will achieve this. The first version of the current report is due by M6 (30th November 2022) and will be continuously updated throughout the project with the collaboration of all the partners.

The current version of the DMP will be implemented whenever important changes to the project occur due to either inclusion of new datasets, changes in consortium policies or external factors. The DMP will also be important for the access of the data after the termination of the project.

This DMP was developed based on the EC template for DMP (version 1.0 – 5.5.2021) and covers the relevant aspects of data management in the project.



2 Data Summary

[Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

What types and formats of data will the project generate or re-use?

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

What is the expected size of the data that you intend to generate or re-use?

What is the origin/provenance of the data, either generated or re-used?

To whom might your data be useful ('data utility'), outside your project?]

The structure of the MOZART project is shown below in Figure 1, which can be also found in the DoA (Description of Action).

It shows how the work packages are connected to each other and how the management work package (WP1) constitutes a supporting foundation for the project as a whole. This illustration of the project organization is necessary for the understanding of the generation and flow of data within the project.

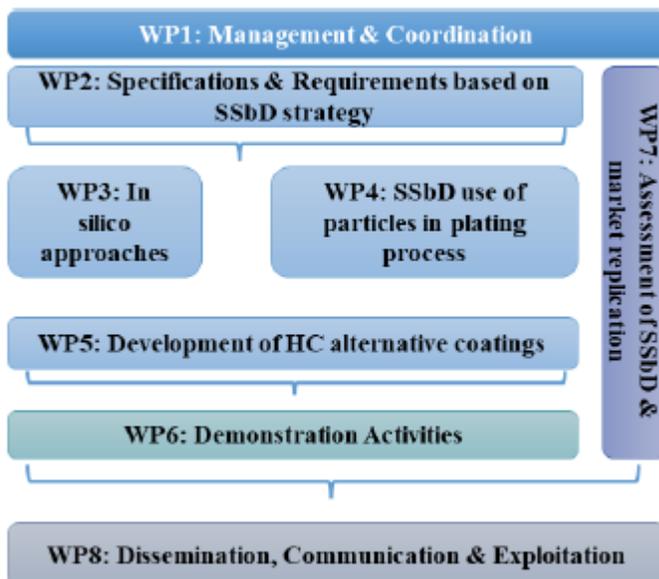


FIGURE 1 *Project structure*



The project consists of eight work packages (WPs), considering also the management and coordination WP (WP1); each WP may involve different input methods, analysis of and output data. Table 1 below shows an overview of the data sources, formats and where (ie in which WP) it was/will be generated/used. The MOZART project will collect and analyse the following type of data:

- Friction and wear data of different coatings in different tribology test instruments. Tests are still to be defined;
- Data regarding IPR protection and intentions of the partners to patent and bring their results to the market (Fernanda-Axia);
- Various simulation data for different particle types (OF2i Simulations as well as data obtained from CFD simulations within WP3).
- Experimental data from lab measurements and onsite at pilot lines (WP4, 5, 6). (Marko-Brave).

The project will also compile and analyse results and will collaborate with other past and currently running EU-funded projects. The inventory of data will be updated when and if new data types are identified.

Data obtained from previous EU projects will be used to compare the existing technologies developed within the MOZART project. Specifically, data related to the electroplating process for the production of Ni/SiC and Ni/GNP nanocomposite coatings (e.g., current density, plating time etc) will be compared with data already obtained from two H2020 running Projects: Sabydoma and Diagonal where different conditions were employed and may provide reference data for MOZART.

Table 1 **Data source within WPs**

Data source	WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8
Personal information	x							X
Re-use of other programmes data				X			X	X
Surveys	x						X	X
Interviews							X	X
Event/Workshop	x						X	X



Analysis outcome		X		X	X		X	
Presentations based on analysis outcome	X			X			X	X
Reports based on analysis outcome	X	X		X	X		X	X
Simulation outcome			X					
Specifications & Requirements based on SSbD		X					x	
Critical Parameters based on optimized computational experiments			X				x	
Practical SSbd parameters for particles in plating processes				X			x	
Performance and functionality of innovative coatings					X		x	
End Users requirements						X		
LCA, LCC, evaluation of SSbD metrics							X	

It is thus clear that the project will need to collect and share a significant amount of data – data which have to be stored, managed, and used according to standards that will ensure the safeguarding of the data and the quality of the project.

The data, both generated and re-used, will be stored in file formats such as .xls, .docx and .csv and will consist of tabular data, text documents, numerical data, figures. It is expected to generate 5-10 files of 3-20 MB size per experiment thus the size of the data will not be an issue.

The purpose of collecting and analysis of data within this project is to:



- evaluate the suitability of coatings for end users.
- To assess the technologies developed inside the project and identify the key exploitable results, as well as the ownership of results by each partner (Fernanda, Axia)
- New data might lead to unforeseen IP and novel analysis mechanisms (Marko, BRAVE)
- Develop a SSbD assessment tool supporting the different case studies that a stakeholder might need to run

The data is going to be generated from the Hull cell, lab scale, medium scale and pilot line electroplating with nickel-based baths using different acids in the bath to replace boric acid and electroplating parameters to improve coatings' properties. Also, from characterization techniques such as microhardness, optical microscopy, contact angle, SEM-EDX, XRF, DLS, ICP-MS and electrochemical techniques such as corrosion, linear polarization, cyclic voltammetry, etc. The new data and the re-used data will be combined to optimize the electroplating process of the Ni-matrix nanocomposite coatings using 2D or ceramic nanomaterials to replace HC.

The collected data within the project are going to be useful for:

- Industry end users who consider replacing HC coatings by alternatives;
- relevant stakeholders that might want to invest in the technologies developed
- Science community in the field of electroplating processes and particle characterization and dynamics in general
- Relevant SMEs of the plating industry who need a supportive environment to navigate themselves in the challenging action of implementing SSbD criteria within the area of the plating industry.

The origin of re-used data is anticipated from previous EU funded projects, such as Sabydoma & Diagonal. The new data will be generated from MOZART experimental procedures to demonstrate HC replacement by Ni-based nanocomposite coatings.



The project's generated data will be of primary relevance to the electroplating industry in general and specifically to plating companies and shops that use electrodeposition with Cr(VI). Also, to companies who produce Ni-based coatings and use boric acid in their electrolyte bath formulation.

Life cycle and Life Cost Analysis:

The data, both generated and re-used, will consist of either tabular data, text documents (digital or hard copy), google forms, audiovisual (e.g. images, sound recordings, videos), models or personal communications (email, phone), thus the size of the data will not be an issue, with the exception, perhaps, of the audiovisual datasets. All data will be stored in the internal project repository, hosted in partner NTUA's servers.

All results produced will not be made public by default, moreover data provided by some partners may be confidential and this has been stipulated in the Grant Agreement and Consortium Agreement. The contact information used for the registry, transparency of data sources, data requests, and information for the Life Cycle Assessment (LCA) and Life Cycle Cost (LCC) activities will be regulated according to the General Data Protection Regulation (GDPR), which further underlines the importance of ensuring data management high-quality is respected.

The purpose of collecting and analysing data within this project is to implement the LCA analysis and LCC analysis according to ISO 14040/44 and other applicable standards, for investigating the environmental impact of project MOZART innovations and solutions. Life Cycle Inventory (LCI) is the phase of an LCA that requires the most effort and resources. In this phase, data collection and modeling are to be done. The necessary data related to the operations, the processes, and the products will be supplied by data providers, including consortium partners and related stakeholders.

The collected data within the project is likely to be useful to the consortium for information, follow-up, improvements of the MOZART solutions based on the results and recommendations as well as for the development of business models.



3 FAIR data

FAIR is a set of guiding principles to make data Findable, Accessible, Interoperable, and Reusable. MOZART project follows these principles, and this chapter describes the process.

3.1 The FAIR principles

The FAIR principles define a set of characteristics that data, tools, vocabularies and infrastructures should have in order to be findable, accessible, interoperable and reusable.

3.1.1 Findable

Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

3.1.2 Accessible

Once the user finds the required data, they need to know how the data can be accessed, possibly including authentication and authorisation.

3.1.3 Interoperable

The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

3.1.4 Reusable

Metadata and data should be well-described so that they can be replicated and/or combined in different settings.

Each principle is broken down further into a series of measurable requirements. The FAIR principles have a low barrier-to-entry: they are minimally defined so that data producers, publishers and stewards can easily adhere to them, and they can be adhered to incrementally. The FAIR principles are also modular: they can be adhered to in any combination, which supports a wide range of applications and special circumstances.



3.2 Making data findable, including provisions for metadata

[Will data be identified by a persistent identifier?

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Will metadata be offered in such a way that it can be harvested and indexed?]

The Project Consortium members will be granted access to a protected shared space, directly provided by the Coordinator (PoliMi), Microsoft ShareDrive.

This will be used for storing data which can be shared within the consortium to enable easier collaboration between the partners. The folder structure will be built according to each WPL needs. Keywords will possibly be applied for searching the required files. Deliverables, agreements, and other formal documents (reports, recordings, etc) which are provided with document histories and version numbers in order to make it easier to track changes made and ensure that the reader has the most recent version of the file.

After the end of the project, the data indicated as confidential in the GA, that will have been generated throughout the project will be kept confidential for commercialization reasons.

3.2.1 Making data openly accessible

[Repository: Will the data be deposited in a trusted repository?

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

Data: Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate

interests or other constraints as per the Grant Agreement.



If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Will the data be accessible through a free and standardized access protocol?

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project? How will the identity of the person accessing the data be ascertained?

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

Metadata: Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not,

please clarify why. Will metadata contain information to enable the user to access the data?

During the MOZART project data will be generate but according to the significance of the results, the output data will be characterized as opened or restricted. If the data is open, then data will be made available (including meta-data) following FAIR principles and use open access repositories, such as Zenodo.

MOZART project **needs to keep some data confidential** as they are of sensitive nature to external industrial parties. These include:

- End users process and product information.
- Data regarding the KERs to be patented before IPR protection been applied (Fernanda-Axia)
- Internal measuring reports, internal outcome-presentation and technical approaches (partner related internal know-how) (Christian-BRAVE)
- Potentially, experimental data from technical partners. However, the DST should be downloadable and we plan to hide different sensitive data, and/or avoid them (Marinella, AXIA).
- Data related to processes, energy and material measurements required for life cycle environmental, cost analyses (NTUA)

The consortium will make, when required, aggregated data and results in presentations, and reports available and still protect the business confidentiality. No collected data will be published in an open-source database if not already agreed upon within the GA or subsequently directly with the Coordinator or the involved partners. Thus, there will be no arrangement within the project to enable external access to the internally collected data until its open publication, if possible. For project



partners, the issue of knowing what data and information are to be kept confidential is regulated in the Consortium Agreement or can be directly checked with the authors or relevant parties.

MOZART developed a public homepage where all visitors can access the full version of all PUBLIC and training materials for download. There will be no embargo applied to give time to publish or seek protection of patent within the MOZART project thus no related delay in access will apply in this project.

3.2.2 Making data interoperable

[What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Will your data include qualified references¹ to other data (e.g. other data from your project, or datasets from previous research)?

Interoperability is the real-time data exchange between different systems which speak directly to one another in the same language, instantly interpreting incoming data and presenting it as it was received while preserving its original context.

Interoperability of some raw data between WPs is not a critical issue for the progress of this project. When preparing deliverables, seminars, external communication activities, etc. results and data will have to be presented in a way that is comprehensive for the audience in question.

Data generated throughout the project will use standardised formats with the objective of facilitating the exchange and later use among the partners and different stakeholders. If information exchange is needed, the data will also be given standard and conventional naming, both inside and outside of the consortium.

The metadata is stored at first on the MOZART ShareDrive. The processed and published data will be available on the project homepage or directly via ZENODO or other platforms. Whenever possible interoperable file formats will be used, such as .CSV, .txt etc.



The project will re-use reports from other projects and refer to these and project internal reports through standard referencing.

For Interoperability, the dataset will follow standards, vocabularies, and methodologies, whenever possible, allowing data exchange between partners. Community-endorsed interoperable best practices will be explored and included in the revised version of the DMP if needed.

3.2.3 Making data re-useable (through clarifying licenses)

[How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Will the provenance of the data be thoroughly documented using the appropriate standards?

Describe all relevant data quality assurance processes.

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects

related to the allocation of resources, data security and ethical aspects.]

The aggregated data and reports generated within the MOZART project, to succeed in the objectives stated in the DoA, will be openly accessible either on the project homepage (<https://www.mozart-project.eu/>) or on other open platforms throughout the duration of the project. All PUBLIC reports, publications and exploitation material generated in the project will be unlimited to share, thus a CC BY 4.0 license is sufficient. This license lets others distribute, remix, adapt, and build upon all materials, even commercially, as long as they credit the project for the original creation. This is recommended for maximum dissemination and use of licensed materials and will be applied to MOZART publications and reports.

The project management office will, through project internal meetings, presentations and newsletters to the consortium members, endorse the reuse of data within the project. To enable the reuse of the data, definitions, methodologies, assumptions etc. will be described in conjunction with the data presentation such as in file headers, comments in data sheets or in annexes in reports.



Challenging in this work will be coping with the wide, not limited, range of members, being global industrial organisations, authorities, universities, researchers etc. who can have different standards and vocabulary.

The significance of the results will determine the actions afterwards. In order to re-use the data, an agreement, for example, between partner Cnano and third parties needs to be put in place.

4 Other research outputs

[In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.]

Software involved in the data processing for the LCA and LCC implementation is GaBi 8.5 Sphera™ and Microsoft Office (Excel, Teams, OneDrive, and Outlook). All data collected and processed are in accordance with Data Protection Legislation and comply with Grant Agreement regulations.

5 Allocation of resources

[What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

Who will be responsible for data management in your project?

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?]

Data management in the MOZART project will be handled within all individual WPs. At project level, Prof. Luca Magagnin, representing Polimi as project coordinator, will be responsible for data management as part of WP1. Polimi has allocated a part of the overall WP1 budget and person months to these activities. Beyond this, there are no specific costs, directly related to data management within the project and indirect cost, such as Microsoft tools licenses etc, are included in



the overhead of all partners. The overhead will be covered by the partner organisations or authorities within each country.

No license fees for data management software have been included in the budget. Long-term preservation of data might be a potential issue, once the project reaches its natural end (42M duration). The data will be preserved for the time stated in the GA.

Responsibilities lie with the WP leader in which WP the data is collected. There will be a regular meeting between Project Management Office and all WP leaders with data management as a standing point on the agenda.

The responsible person for the data management for Cnano's part is Dr. Alexandros Zoikis-Karathanasis.

6 Data security

[What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

Will the data be safely stored in trusted repositories for long term preservation and curation?]

The data will be kept safely within the internal data management system and will only be transferred amongst the partners during the project duration.

MOZART findings and data will be stored in online cloud systems provided by the IT department of PoliMi.

Data security will be assured using the security systems at each partner's organisation and also the degree of security provided by Microsoft for data stored on partners' individual cloud storage and made available to the project. Virus security systems is also provided by Microsoft and all files uploaded will go through a process of virus detection. The access to the project SharePoint is administrated by the Project Management Office. For new project members to have access, it will be needed for them to request and explain in what role they need access. The WP leaders have to approve any new person within their WPs. All new project members will be given an introduction by the WP leader to this DMP to ensure the requirements are applied.



SharePoint is a cloud-based tool for cooperation in projects. It has a two-step authorisation access for all users. This is the standard protocol globally and is sufficient for this project's purposes. The coordinating organisation has a recovery solution with backup put on a separate server.

SharePoint is the branch standard for digital office environments, and this is why the coordinating organisation encourages, supports, and offers education in its usage for this project.

The generated data will be shared between the partners as long as the project runs.

7 Ethics

[Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA)]

Personal information can be an ethical aspect and the MOZART project will handle a significant amount of this type of data.

All personal data (stakeholder input and data collected) will be anonymized or permission to publish will be requested from the data sources. All data management will be compliant with EU Regulation 2016/679, General Data Protection Regulation. The partner in charge of data management, as stated above is Prof. Luca Magagnin representing PoliMi, and will deploy a GDPR-compliant tool and procedures; This includes a project internal GDPR content for project members as well as a general consent policy sent to any external stakeholders contributing, through workshops, surveys etc.

8 Other issues

[Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?]

MOZART has not identified any other ethical aspects related to the project. Nor have other ethical issues been identified with regard to data management for the moment.

The DMP will be monitored throughout the project's lifetime and any new data related issues identified will be resolved internally, through partners' individual processes and actions. In such a case the DMP will be updated appropriately.

