

Dealing with practical aspects of multilinguality across national and disciplinary boundaries

Session C1: Toward practices for FAIR data management and sharing by a researcher, in a community, and beyond a community

Alison Specht

Research Analyst, TERN, University of Queensland
ORCID: 0000-0001-1512-0743

@AlisonSpecht | a.specht@uq.edu.au

With contributions gratefully received from:

Shelley Stall¹, Jeaneth Machicao², Thibault Catry³, Marc Chaumont⁴, Romain David⁵, Rodolphe Devillers³, Rorie Edmunds⁶, Robin Jarry⁷, Laurence Mabile⁸, Nobuko Miyairi⁹, Margaret O'Brien¹⁰, Pedro Pizzigatti Correa², Solange Santos¹¹, Gérard Subsol⁷, Lesley Wyborn¹²

¹American Geophysical Union (AGU), Washington DC, United States of America

²University of São Paulo, Brazil

³Espace-Dev (IRD, University of Montpellier, University of Guyana, University of La Réunion, University of Antilles, University of New Caledonia), Montpellier, France

⁴Montpellier Laboratory of Computer Science, Robotics, and Microelectronics (LIRMM), University of Montpellier, Centre National de la Recherche Scientifique (CNRS), Montpellier, University of Nîmes, France

⁵European Research Infrastructure on Highly Pathogenic Agents (ERINHA), Brussels, Belgium

⁶DataCite, Hannover, Germany

⁷Montpellier Laboratory of Computer Science, Robotics, and Microelectronics (LIRMM), National Centre for Scientific Research (CNRS), University of Montpellier, Montpellier, France

⁸University of Toulouse III - Paul Sabatier, French National Institute of Health and Medical Research (Inserm), France

⁹National Institute of Information and Communications Technology, Japan

¹⁰University of California, Santa Barbara, United States of America

¹¹Scientific Electronic Library Online (SCIELO), São Paulo, Brazil

¹²National Computing Infrastructure, Australian National University, Canberra, Australia

*We acknowledge the Traditional Owners and Custodians of the land and sea in all nations.
We honour their profound connections to land, water, biodiversity and culture and pay our
respects to their Elders past, present and emerging.*

PARSEC

Today's
example

2019-2024



PARSEC: 自然保護区が社会経済に及ぼす影響の 多国融合研究を通じた新たなデータ共有・再利用手法の構築

国際コンソーシアムリーダー: ニコラ・ムケ、ダビド・ムイヨー、アリソン・スペクト、シェリー・ストール



助成総額: 1258千ユーロ

助成期間: 48か月

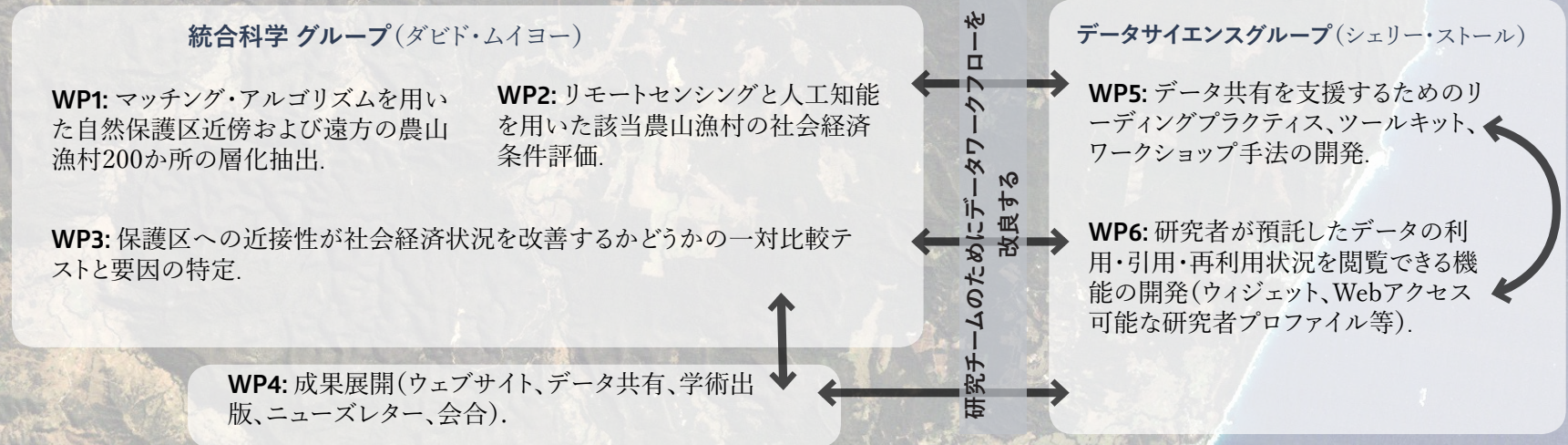
研究の目的

- (a) 衛星画像と人工知能の融合により、地域社会の社会経済に対する自然保護区(PA)の影響を予測する。
- (b) 保護区が地域社会の消費支出と地域資源の健全性にもたらす影響を特定する。

- (c) 将来の環境に対する意思決定のあり方をより良くする。
- (d) 研究者・資金・成果物・データ間のデジタル連携をより良くする。
- (e) 研究データフローを改善し、チームによる研究のスキルを向上させる。
- (f) データ引用の数を増やし、データ作成者への帰属をより明確に示す。

- (g) データの再利用に向けてオープンでFAIRなデータの管理・保存を推進する。
- (h) 研究者が、自分が保存したデータがどのように引用・再利用されているか視覚的に把握できるツールを提供する。

<http://parsecproject.org>



参加国

ブラジル: サンパウロ大学 - ブラジル連邦共和国サンパウロ州学術研究支援財団 (FAPESP) - (P. ビツィガッティ・コレア) + ポスドク・技術支援員 (FAPESP)

フランス: 生物多様性研究財団、トゥールーズ第三大学 - フランス国立研究機構 (N. ムケ)

日本: 情報通信研究機構、総合地球環境学研究所 - 科学技術振興機構 (村山泰啓)

米国: アメリカ地球物理学連合 - 国立科学財団 (S. ストール)

協力機関

オーストラリア国立大学計算基盤センター (L. ワイボーン)、英国地質調査所 (H. グレープス)

後援

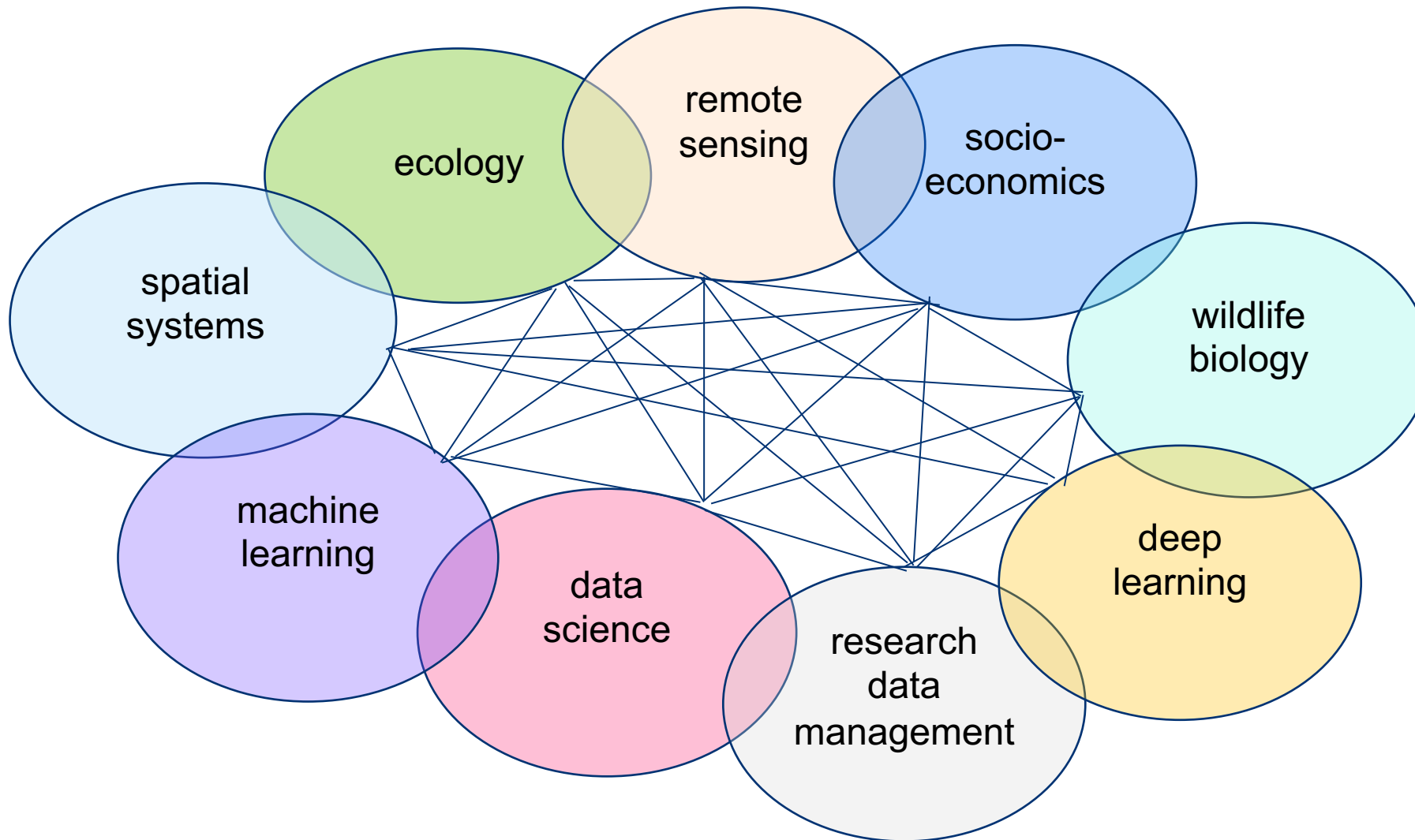
DataCite, ORCID, ESIP, RDA, EDI, WDS, AST, JWP, TNC



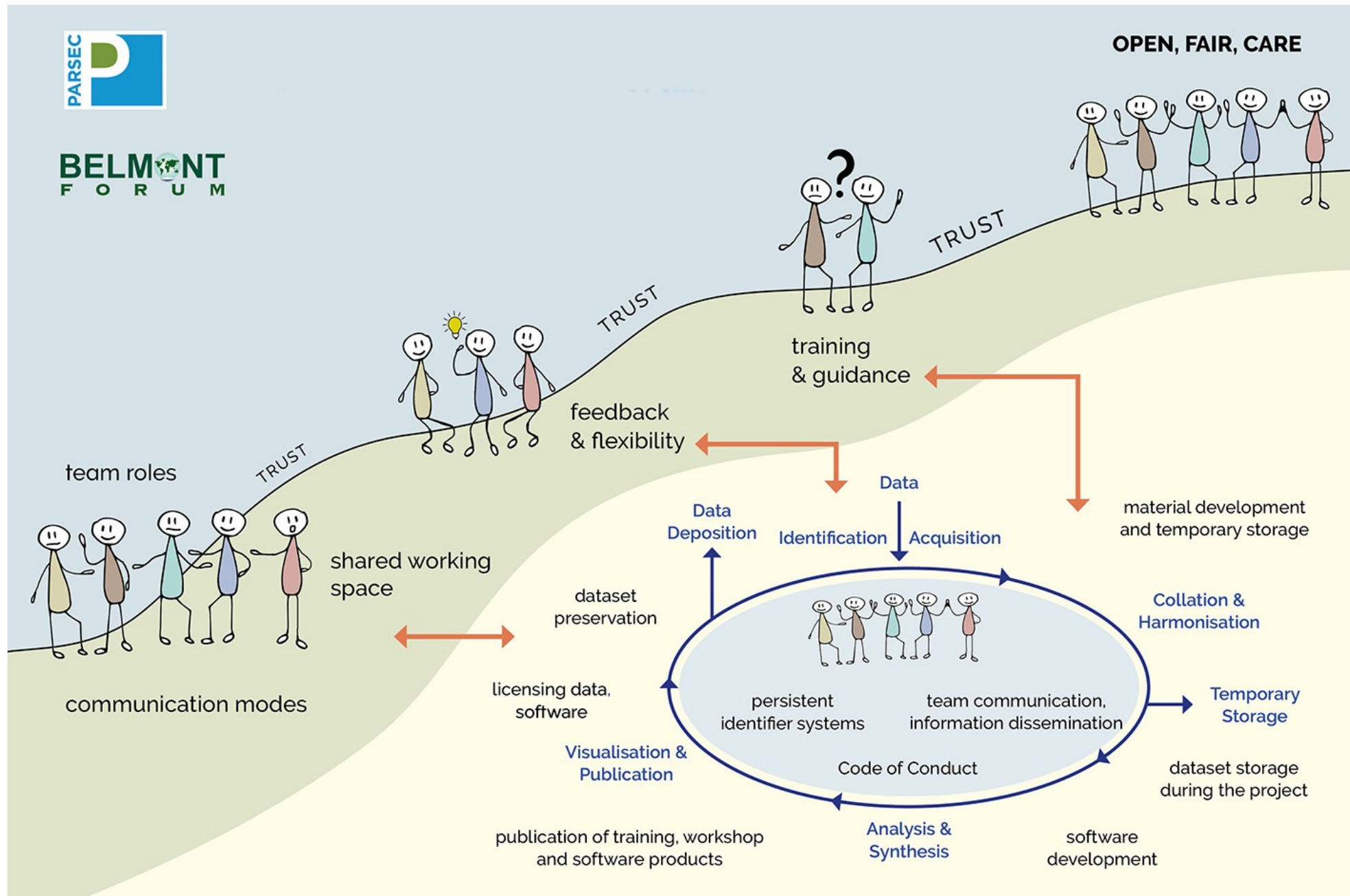
PARSEC : a transnational team



PARSEC : a multi-disciplinary team



- 33 researchers
 - 5 countries
- from
- universities,
 - research institutes,
 - other organisations



Sources of multilinguality



- country of origin
- disciplinary background

Source:

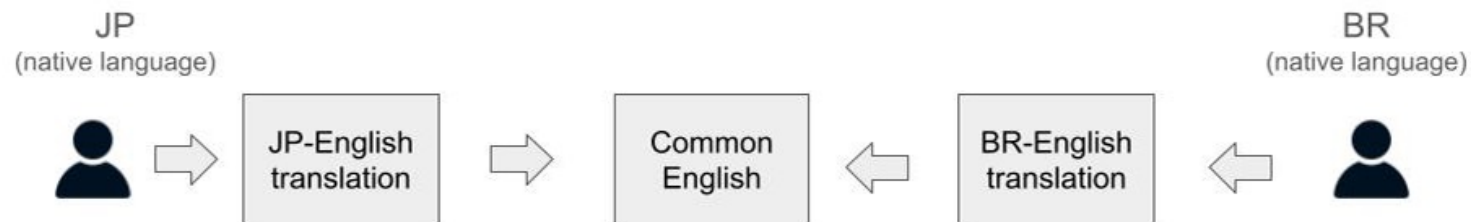
Managing linguistic obstacles in multidisciplinary, multinational, and multilingual research projects.
Submitted to PLOS ONE. Responding to reviewers comments now



Country of origin

Communication and cultural challenges can arise from:

- different languages
- ways of working
- time zone mis-matches
- different terminology used for in-country administrative units
- different terminology and algorithms to calculate country-based metrics



Examples of vocabulary translations for Open Science

English	Portuguese	Spanish	French	Japanese
Data Management Plan (DMP)	Plano de Gestão de Dados (PGD)	Plan de Gestión de Datos (PGD)	Plan de Gestion des Données (PGD)	データマネジメントプラン (DMP) [phonetic transcription]
Computational Notebooks (e.g., Jupyter Notebook, R Markdown)	Cadernos computacionais	Cuadernos computacionales	Notebook (in routine) / Carnets de notes pour logiciel / Calepin électronique	計算機上のノートブック環境 [English: Computational Notebook Environment]
Open Science Journey	Jornada pela Ciência Aberta	Viaje por la Ciencia Abierta	Parcours Science Ouverte	オープンサイエンスの旅

Disciplinary challenges

Disciplinary challenges can arise between disciplines or even within sub-disciplines due to:

- discipline-specific terminology,
- different kinds of data being seen as more important than others
- a perception that different kinds of analyses are more valid than others
- modelling in multidisciplinary space needing input from all relevant disciplines to ensure model validity, relevance, transparency and acceptability
- Different programming languages, programming styles, and frameworks

Disciplinary obstacles to the definition of the work (scientific question, hypothesis, methodology)

Obstacles encountered *

high (4: stops the work) to low (1: minor interruption)

Examples

a. Alignment of principal keywords in ecology across domains [4]

The term 'ecoregion' was misunderstood.

b. Understanding (alignment) of the DL model by all the team [1]

Confusion between the meaning of the terms 'prediction' and 'estimation'.

c. Regional/national key development questions were differently understood by team members [1]

The difference between the terms 'survey' and 'census' was not initially appreciated.

d. Setup of the terminology for the methodology presented challenges [2]

Use of various terms, e.g. 'control - reference'; 'treated - non-treated'; 'mirror - control'; 'experiment - run'

What will help bridge multilingual gaps?

1. Ensure all team members have a strong common understanding of the project purpose, good knowledge of the tasks to be done and their contribution.
2. At the start of the project define the terms you will be working with and record them in a shared space.
3. Never assume understanding across boundaries.
4. Identify the obstacles encountered and share promptly with the wider team.
5. Ensure relevant 'non-core' participants are involved and have opportunities for meaningful input.
6. If disciplinary-specific language and new concepts are being used, double check on-line translations.



Inter-University Research Institute Corporation
National Institutes for the Humanities
**Research Institute for
Humanity and Nature**



**TOKYO
METROPOLITAN
UNIVERSITY**





Inter-University Research Institute Corporation
National Institutes for the Humanities
**Research Institute for
Humanity and Nature**



**TOKYO
METROPOLITAN
UNIVERSITY**



ORCID
thankyou



European Research Infrastructure
on Highly Pathogenic Agents



**THE UNIVERSITY
OF QUEENSLAND**
terrestrial
A U S T R A L Ecosystem Research
Infrastructure