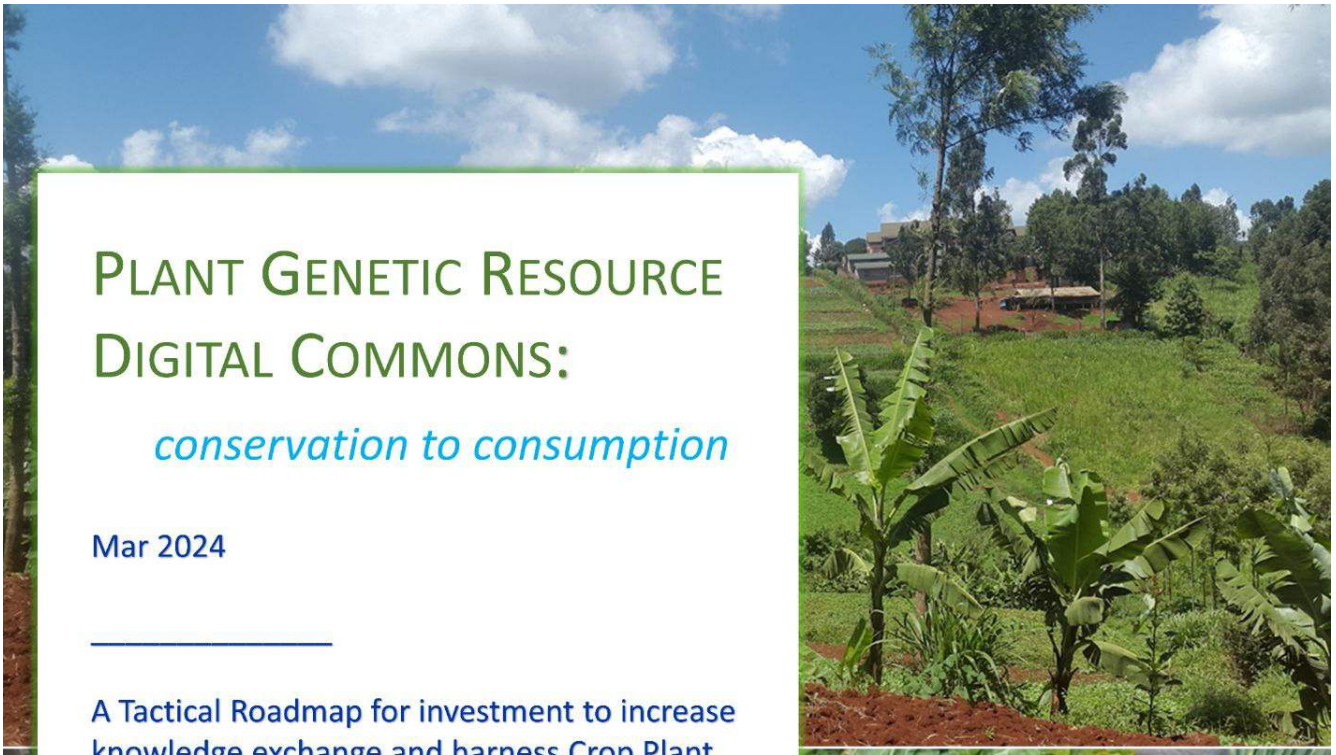
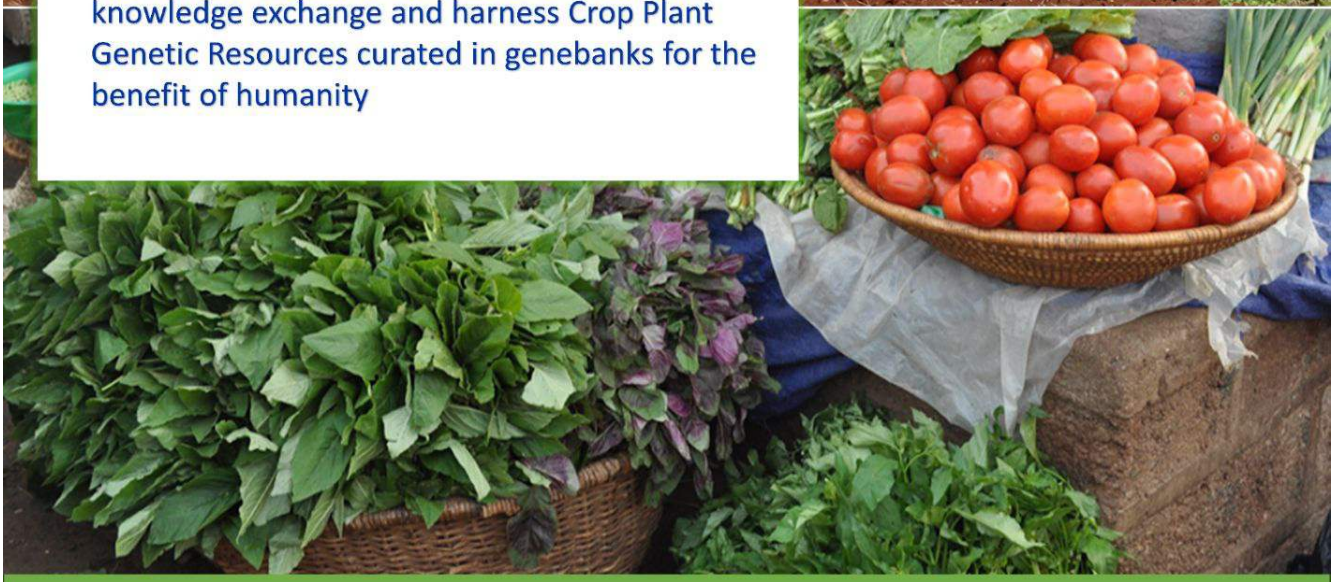


PLANT GENETIC RESOURCE DIGITAL COMMONS:

conservation to consumption

Mar 2024

A Tactical Roadmap for investment to increase knowledge exchange and harness Crop Plant Genetic Resources curated in genebanks for the benefit of humanity



DIVSEEK INTERNATIONAL NETWORK INC.

Plant Genetic Resource Digital Commons*:*
conservation to consumption

A Tactical Roadmap[1] for investment to increase knowledge exchange and harness Crop Plant Genetic Resources curated in genebanks for the benefit of humanity.

EXECUTIVE SUMMARY

What? **This document presents an analysis of the value proposition of Plant Genetic Resources (PGR) and the current state and the future potential of information flow around PGR collections conserved in genebanks**. This leads to the proposition that there is a pressing need to democratize access to the world's PGR to ensure that humanity can realize the benefits from the tangible connections between food security and health and peace outcomes.

Plants underpin earth systems and human civilization. PGR collections contain the intrinsic capacity to meet global food supply, mitigate many preventable human health problems, and tackle regionally relevant challenges posed by climate change and environmental degradation.

We present a value proposition and a theory of change for **democratizing access to information** about PGR, facilitated by a multi-domain community of practice.

Barriers to communication are identified, along with opportunities to democratize access to information derived from PGR and increase access and use for breeders, farmers, researchers, and the wider bio-economy.

Economic, social and environmental benefits are highlighted that are expected to arise from coordinated development of robust data standards and interoperable working practice. A vision is presented for increased investment aligned to practical next steps.

This document builds on discussions within the DivSeek International Network, including a focused workshop in 2023 co-organized with the Global Crop Diversity Trust.

Why now? Realizing PGR information value builds on rapid advances in genomics, genome editing, information technologies, and artificial intelligence (AI).

Who cares? Crop diversity, particularly that **contained within genebanks,** functions as a driver of plant breeding advancement and is a critical component of current and future **global food security** and **climate adaptation**. **Information flow** between different domains in the crop value chain is key to realizing the potential of PGR.

[1] *This document was prepared by a community of practice led by DivSeek International, as guardians and mediators of an open content process that includes workshops, publications, open-data resources and wider consultation. Readers may also wish to consult the DivSeek International Network Strategic Plan (https://divseekintl.org/strategic-plan/), including the Glossary of Terms (p14-15).*

Recommended Actions:

This vision outlines the urgent need for a **global reassessment of the value of PGR-derived *information flow*** and engagement to establish a **PGR *Digital Commons.***

An associated **set of actions** are proposed, requiring:

1. Formation of federated[2] Communities of Practice with a wider view of PGR value.
2. Broadening the remit of a dedicated crop PGR *Digital Commons*.
3. Increasing connectivity between national/international informatics infrastructure providers.
4. Development of scalable sets of best practices and standard operating procedures.
5. Creation of new R&D and industry outreach centers interconnected with genebanks.
6. Economic analyses to quantify the PGR value chain.
7. Extension of available examples for traceability of germplasm from source to consumer.
8. Increase knowledge and awareness of the value of PGR and derived information for the wider public, including harnessing participatory citizen science.

A further five **specific actions** are proposed to address data standards for PGR conservation and utilization.

9. Demonstration projects to highlight the value of adopting relevant, generic, and scalable standards.
10. Identify incentives and innovative approaches to stimulate adoption, completion and sustaining of standards.
11. Establish a PGR Standing Group for standard authorisation and deployment.
12. Tailor controlled vocabulary (ontology) use for specific production, supply chain and end-uses.
13. Apply generic ontologies to describe traits of minor and under-utilized crops.

[2] eg. a union of independent regional hubs that share best practices and outcomes

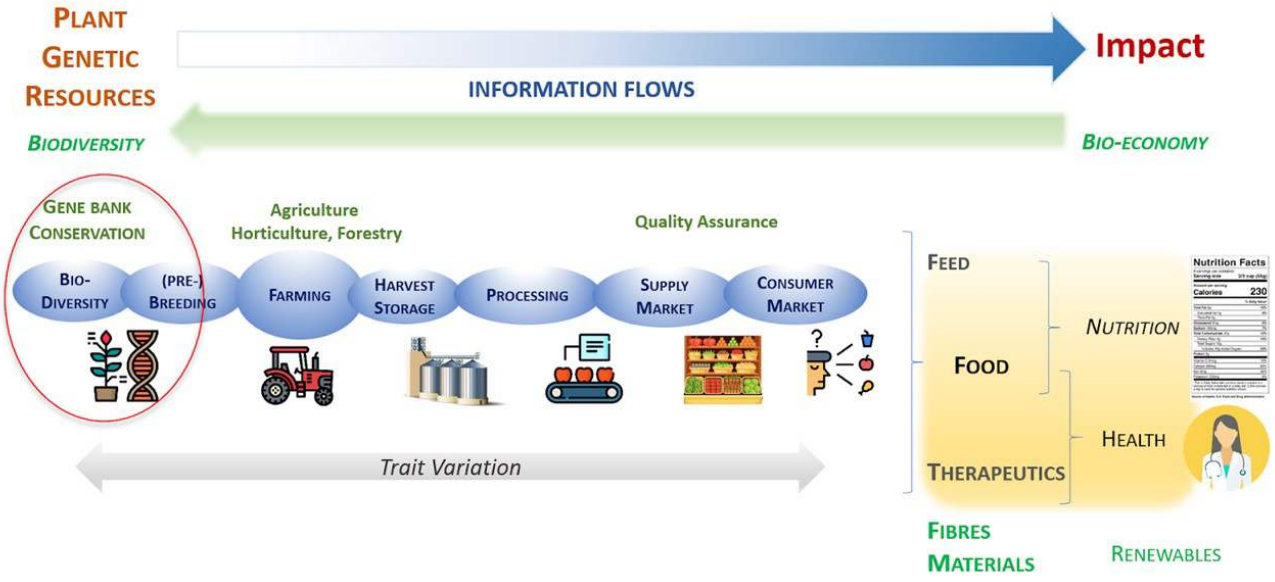
The Value Proposition

Plants underpin earth systems[3] and human civilization. **Biodiversity curated** within *ex situ* plant genetic resource (PGR) collections (genebanks) represents an irreplaceable source of variation to meet challenges of climate adaptation and mitigation, food security and human health.

The crop value chain globally spans the PGR, trait discovery, R&D, breeding, production, post-harvest, processing, supply-chain and consumption sectors, with associated socio-economic and health outcomes. In this context, the capacity for annotating complex datasets and exchanging data related to PGR characterization is limited.

The tangible connections between health outcomes and PGR are largely overlooked in policy decision-making. This under-valuation of PGR in modern socio-economic systems presents an existential risk for human civilizations.

Food supply beyond primary production spans multiple processing and supply chain sectors. Despite consumer awareness and concern about plant varieties being limited, crop varietal attributes are fundamental to nutrition and health outcomes as well as to quality, provenance and regional identity.



PGR management and evaluation efforts are often not integrated with the PGR-derived value chain, with interactions and interfaces (human, live plant, digital) not visible to individual players or policy makers.

Plant breeding of new cultivars adapted to increasing demands and challenges of changeable growing environments, market diversification, and functional end-uses requires access to genetic variation alongside characterization and evaluation[4] data.

[3] Coates, Juliet C., et al. "Plants and the Earth System - Past Events and Future Challenges." The New Phytologist, vol. 189, no. 2, 2011, pp. 370–73. JSTOR, https://doi.org/10.1111/j.1469-8137.2010.03596.x.

[4] PGR *characterization* involves i) a description of the distinctive nature or features of an organism; ii) detection of variation as a result of differences in either DNA sequences or specific genes (genotype), or any aspect of morphology, physiology, biochemistry, interaction with biotic or abiotic environment, appearance or quality traits (phenotype). *Evaluation* includes assessment of germplasm for any agronomic (production) or end-use quality traits.

Governance and management of Plant Genetic Resource (PGR) collections

PGR collections are represented by a system of national and international crop-specific genebanks. A wider range of research organizations, breeders and farmers manage PGR collections around the world. These often support regionally relevant vertically integrated grower/producer collectives, added value enterprises and regional industries.

Access, exchange and downstream use of PGR are regulated by **international agreements**, including the Convention on Biological Diversity (CBD) and principles of Access and Benefit Sharing (ABS), the International Treaty on Plant Genetic Resources for Food and Agriculture (FAO ITPGRFA), and the Nagoya Protocol[5].

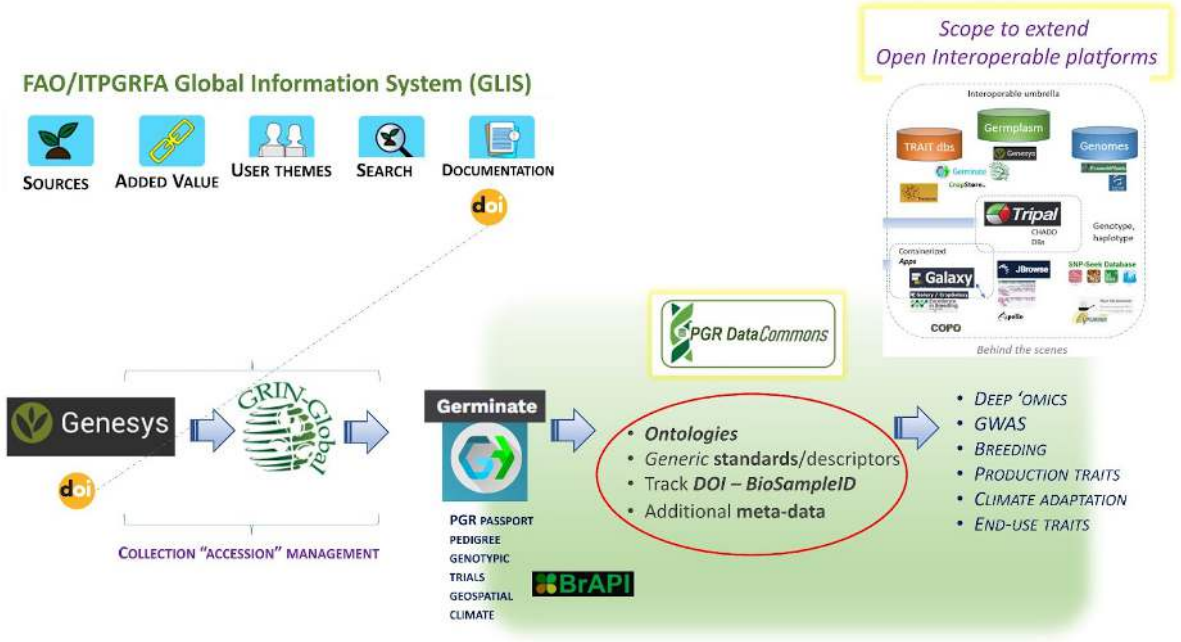
- While the stated goals and principles are well–intentioned, the real world impact has resulted in confusion, lack of a workable ABS framework and in particular, limited freedom to operate.

- Uncertainties and inconsistencies contribute to many commercial private breeders managing private diversity collections independently from the global PGR system, with unknown levels of overlap.

The role of information systems in the realm of PGR

PGR information sharing underpins the value proposition of PGR collections and needs to be findable, accessible, interoperable and re-usable (FAIR). Current data infrastructure and information sharing mechanisms are limited in their ability to link collections with assessment of useful PGR variation. A wide range of stakeholders could contribute and benefit from high quality characterization and evaluation data, using common vocabularies and ontologies to describe crop phenotypes, traits and downstream attributes.

A loosely connected data ecosystem has been promoted through combined actions of the FAO/Treaty GLIS[6] and the public sector R&D community across the world. This ecosystem is primarily focused on germplasm inventory and targeted characterization and evaluation programs, rather than downstream evaluation.



[5] https://doi.org/10.1073/pnas.2205773119
[6] https://glis.fao.org/glis/

No coordinated or comprehensive system or portal is available for users to explore the full landscape and impact of PGR data, or to link this with downstream data and information processing resources. The *DivSeek Commons* 'matrix'[7] is a first step in this direction.

The DivSeek International Network

DivSeek began in 2012 as an informal global community of practice for sharing information about plant genetic resources. Since publication of a 2014 White Paper '*Harnessing the power of crop diversity to feed the future*' [8] there has been considerable change, including rapid advances in informatics, low cost whole genomes, global warming, and a 706m (9.6%) increase in world population.

DivSeek International Network Inc. was established in 2018 as a not-for-profit with a Board elected by the ¬70 institutional members. Its mission is to enable those who conserve and benefit from the sustainable utilization of plant biodiversity to access and harness expertise, scientific evidence, technologies and global best practices that facilitate the characterization, utilization, exchange and traceability of genetic resources.

- A vision and strategic requirements are outlined in the *DivSeek Strategic Plan*[9] (2021-2026);
- DivSeek has established an initial global set of regional and thematic Hubs[10].
- A Memorandum of Understanding has been signed with the FAO International Treaty on Plant Genetic Resources for Food & Agriculture.
- The *DivSeek Commons*[11] concept was proposed in 2022 as an ecosystem of Standards, Open Data Resources, Analysis Tools, and Best Practices for Plant Genetic Resources.

Improved interoperable PGR information systems are expected to facilitate de-risking of global and regional system failures to:

- Contribute to AgroBiodiversity and land productivity
- Accelerate transition of food systems to climate change mitigation (climate smart targets)
- Improve food security through locally adapted crops
- Meet health/nutrition enhanced targets
- Find ideal genotypes with beneficial consumer and industry attributes (ideotypes)
- Diversify future products (food, bio-renewable, therapeutic)
- Address emergent unknown targets and threats

[7] https://divseekintl.org/commons-landscape-matrix/; https://divseekintl.org/divseek-commons-the-matrix-is-born/
[8] https://divseekintl.org/wp-content/uploads/2019/12/WhitePaperDivSeek.pdf
[9] https://divseekintl.org/strategic-plan/
[10] https://divseekintl.org/regional-hubs/
[11] https://divseekintl.org/commons/

Democratizing Access to PGR Information through the Crop Value Chain

CONTEXT

- The AgFood and wider bioeconomy depend on reliable supplies of cultivated plants adapted to regional environments and markets.

- Historically, access and systematic characterization of PGR has led to transformational outcomes[12] (eg - from 1960s onwards, dwarfing wheat and rice; canola introduced as new oil crop for temperate regions in 1970s; soybean adaptation to tropical regions in South America; discovery and use of natural disease and pest resistances conferred by major genes in many crops).

- It is recognized that a tension exists between local and global food systems. Minor and underutilized crops are mostly too small and limited in human and physical resources to generate cohesive communities of practice (CoP), resulting in poor access to relevant education, modern biotechnology and information management tools.

- There are large imbalances in investment between PGR conservation and downstream utilization.

- investment in PGR information systems is lacking in strategically important and regional traditional crops that are often regarded as minor or underutilized[13], including many horticultural crops.

- Although there has been good progress in international training and exchange of expertise, ongoing digital inequalities[14] represent a major barrier to the conservation and utilization of PGR in much of the world.

- Documentation about PGR accessions is often patchy or lacks visibility about the history of conserved material, limiting traceability and attribution.

ADDRESSING BARRIERS TO COMMUNICATION

The full socio-economic potential of crop PGR is limited by barriers to communication and information flow between different domains that include technological, social, economic and political factors.

Sociological barriers include legal and intellectual property issues, especially around Access & Benefit Sharing; privacy and security concerns; limited awareness of the impact of plant genetic diversity; lack of incentives and resistance to change; exacerbated nationalism, language barriers and cultural and organizational differences.

- Clarifying ownership, usage rights, and licensing agreements is essential for integrating data while respecting legal boundaries.

Technical barriers are primarily associated with data compatibility and standardization, data quality and consistency, and the need for representative sampling of PGR collections and ability to access genetically relevant and explicit data (traceability and heritability). This indicates a requirement for consistent vocabularies, measurement systems and integration of other data-sharing standards.

[12] https://library.oapen.org/bitstream/handle/20.500.12657/61485/9781786768827_web.pdf?sequence=1&isAllowed=y

[13] https://doi.org/10.1007/s00425-019-03179-2

[14] doi: 10.1080/01436597.2022.2079489

Vision for a crop *PGR Digital Commons* to harness the value of PGR

We propose **a global re-valorization of PGR-derived information flow and engagement to establish a PGR *Digital Commons*.** This will require more inclusive development, availability and adoption of best practices amongst different players in the crop research value chain.

Based on the economic, environmental and sociological imperatives outlined in sections 1-3, a key strategic requirement is to **raise awareness** of the value of cultivar and PGR derived information amongst Policy makers across the sectors of Health, Environment, Agriculture and the wider bioeconomy, including biorenewables.

We argue for timely and targeted investment for infrastructure development, and specifically where this supports information flow that adds value to *ex situ* PGR collections through open sharing of evaluation and trait characterization information.

The global R&D community is well placed to extend and connect networks that engage with specific target communities within the wider bioeconomy beyond Ag production and food supply industries.

In particular, there is a need to collate evidence demonstrating added-value in human nutrition, fiber and renewables. In order to engage a wider audience and contribute to benefit sharing there are opportunities to decentralize conservation strategy, by extending the success of consumer engagement and participatory citizen science programs[15] .

Wider cooperation requires economists and data scientists alongside crop and commodity specialists. We anticipate a series of prioritized and concerted actions to address social, legal and technical issues.

A key priority is to increase policy awareness of the role of PGR in health (nutritional security), defense/food security, and climate-change mitigation.

To reach a wider set of user communities, DivSeek represents an existing community of practice and are well placed to broker services to develop more cohesive and flexible governance frameworks to fill existing gaps outlined here.

PROPOSED APPROACHES and ACTIONS

A series of resourcing levels can be used in achieving the vision outlined. Areas for investment may include small-scale projects (~$200k), including postgrads; medium-scale projects < $2m or larger-scale programs and capacity building < $20m.

1. The formation of federated CoPs with a wider view of PGR value requires a trans-disciplinary cohort of practitioners and R&D disciplines.
    - Opportunity to connect diverse communities to establish and disseminate best practices by

[15] https://www.pulsesincrease.eu/

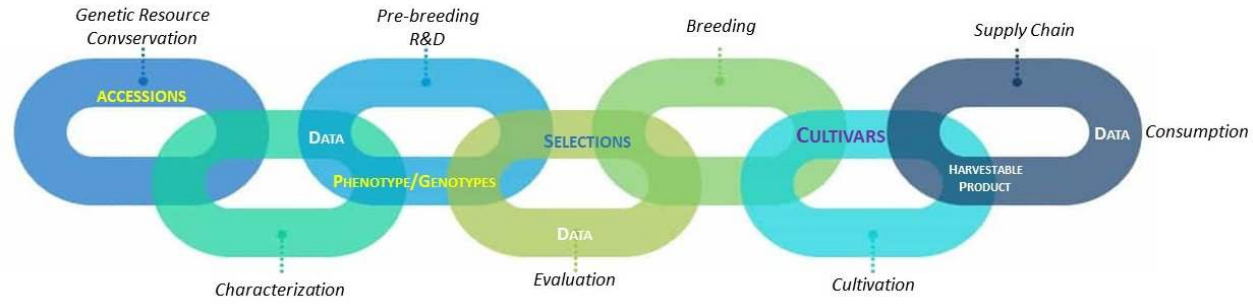
extending the DivSeek Hub model[16] with a globally-federated set of independent regional hubs.

2. The lack of a coordinated approach to interoperability for PGR related information presents a significant digital barrier, and an imbalance between advanced and developing economies. Many of the gaps identified may be addressed by broadening the remit of a dedicated crop PGR *Digital Commons*.
    - Extend the the *DivSeek Commons* PGR information ecosystem 'matrix' as a first step to Build on *DivSeek Commons* matrix[17] to fill gaps for a federated and interoperable PGR-information ecosystem.
    - Develop an open training environment to benefit PGR-rich, but often R&D-poor countries.
    - Catalyze development and access to information systems that leverage PGR research in developing countries.
    - Develop and pilot demonstrations of big data, AI and IoT[18] for PGR.

3. Work with national/international informatics infrastructure providers for provision of lookup services to encourage consistent and persistent use of object identifiers, names, synonyms, etc.
    - identify incentives for traceability compliance through provision of online registries and services that embed traceability.
    - Development of training programs, including peer to peer help-desks and international support.
    - Establish an active GitHub network for digital system developers supporting PGR and downstream communities.
    - Encourage investment in generic software tools that are applicable to a broad base of species/end-use attributes/regions.
    - Stimulate development of "automation" to simplify and improve compliance, lower transaction costs and enable usage.

4. Develop scalable sets of best practices and operating procedures, particularly at the interfaces between domains in the crop value chain that recognize the value of PGR variation.

5. Stimulate creation of new R&D and industry outreach centers that dock into genebanks to manage phenotyping of core-collections and connect with user communities via specialized projects.



6. Economic analyses are required to quantify the research value chain arising from PGR.
    - Regular global and regional assessments of the wider value and socio-economic impact of *ex situ* PGR conservation and use are required in the context of traceable provenance, evaluation and characterization data through the research and wider crop value chain.

[16] https://divseekintl.org/regional-hubs/
[17] https://divseekintl.org/commons-landscape-matrix/; https://divseekintl.org/divseek-commons-the-matrix-is-born/
[18] Internet of Things - https://en.wikipedia.org/wiki/Internet_of_things

7. Extend available examples for traceability of germplasm from source to consumer. Ensure engagement throughout the PGR value chain, using regional and global initiatives to raise public interest in PGR.
    - Develop narratives using demonstration projects to engage the wider bio-economy in understanding the value of cultivar and PGR germplasm variation.
    - Connect PGR and associated R&D with development of robust and consistent specifications through the crop value chain, developing compatible data exchange standards.
    - Identify incentives for PGR accession traceability compliance by down-stream users - through provision of online registries and services that embed traceability.

8. Increase knowledge and awareness of the value of PGR and derived information for the wider public in terms of health, environment and culture. Use information technologies to encourage decentralization of conservation strategy, by stimulating consumer engagement and participatory citizen science programs.

Specific actions to address Data Standards for PGR conservation and utilization

The past two decades have seen a transformation of human and economic behaviors arising from highly interconnected web services and intuitive GUIs. Such activities are enabled by **data standards.** In many sectors (eg automobile, food supply), proprietary platforms benefit from incorporating open standards and dispersed developer communities, enabling traceability in a competitive commercial environment.

CONTEXT

PGR characterization and evaluation involves diverse data types and data systems. There are major limitations for regional implementation with different human languages. To date interoperability and adoption of common controlled vocabularies in the PGR, breeding and R&D sector has been limited[19].

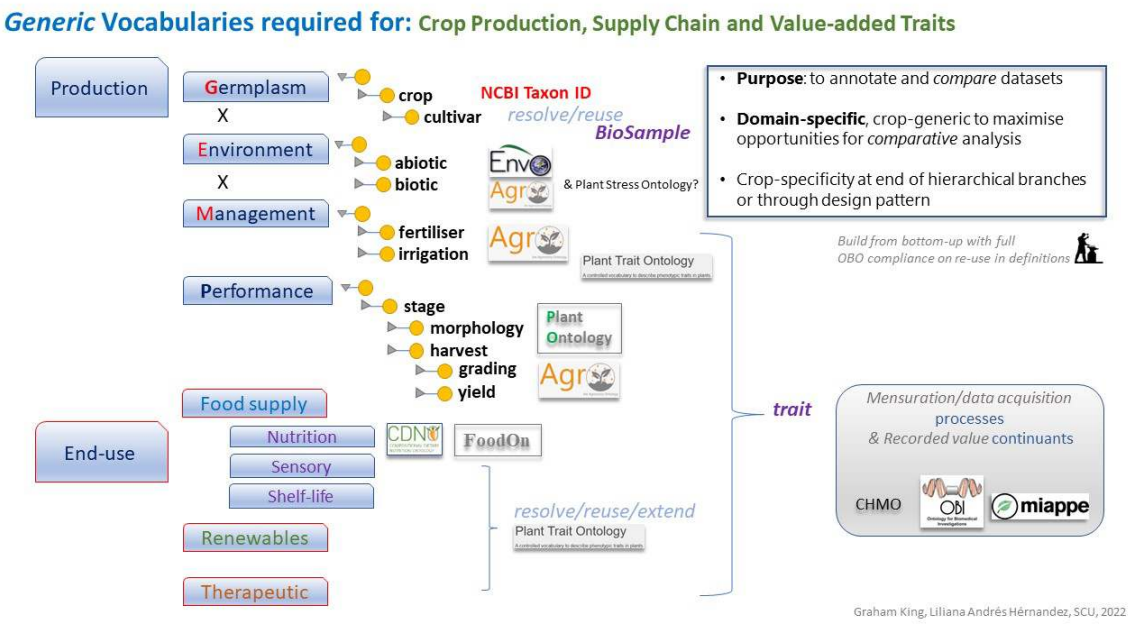
- Consistent systems of organization are lacking for description of attributes (phenotypic trait) data, with mostly *ad hoc* solutions on a crop by crop basis.
- A lack of generic standards reduces the ability of farmers, market makers and researchers to compare crops and their harvested products.
- The role of AI presents opportunities for development and adoption of trait ontologies across languages and crops.

PROPOSED ACTIONS

9. Given the reach of PGR (Sections 1, 3) a concerted effort with demonstration projects is required to address and highlight the value of adopting relevant, generic, scalable standards.

10. Tailor ontology use for specific production, supply chain and end-use applications.

[19] https://doi.org/10.1093/database/baab028



11. Apply generic ontologies to describe traits of minor and under-utilized crops, demonstrating value by focusing initially on areas such as nutritional composition, adaptation to future environments, and food system resilience.

12. Identify incentives and innovative approaches to stimulate adoption, completing and sustaining standards by engaging a broader community in development of PGR-relevant descriptive ontologies[20].
    - This will require careful stewardship to foster inter-community connections and manage disruption.

13. Establish a PGR Standing Group for standards authorisation and deployment.

[20] An ontology is a structured set of concepts and categories in a subject area or domain that shows their properties and the relations between them.