

Proposal Report on Intellectual Property & Data Governance Framework for the Deep-time Digital Earth (DDE) International Big Science Program (Suzhou Center)

1. Background

The Deep-time Digital Earth (DDE) International Big Science Program is not merely a database; it is a fundamental shift in how humanity understands its home planet. By shifting from “Siloed Research” to “Big Data-Driven Discovery”, DDE aims to reconstruct 4.6 billion years of Earth’s history.

Data constitute the essential “fuel” of this engine. In the twenty-first century, however, data are accompanied by complex legal, ethical, and geopolitical considerations. This report sets forth a comprehensive governance framework for the management of these assets, ensuring that DDE remains a beacon of Open Science while fully respecting the legitimate rights and interests of scientists, institutions, and sovereign states.

2. The Five Pillars of DDE Data Philosophy

Building upon the original DDE Statement, released at the DDE Open Science Forum in November 2022, the following five non-negotiable principles shall govern all data on the platform:

Openness (Transparency): Science moves at the speed of light when data is visible. DDE prioritizes “Open by Default” for all research funded by public resources.

Sharing (Interoperability): Data must be not only accessible but also usable. DDE adheres to the FAIR Principles (Findable, Accessible, Interoperable, Reusable).

Traceability (Provenance): Every datum must carry its “DNA”—who collected it, when, and how. This ensures Scientific Integrity and Academic Credit.

Non-Discrimination (Equity): Access to data shall be determined by scientific merit, not by nationality, wealth, or institutional prestige.

Respect for Sovereignty (Compliance): DDE recognizes that geological data may

constitute national strategic resources. Global sharing shall be harmonized with local laws.

3. Structural Taxonomy of DDE Intellectual Property

To manage the chain of scientific contributions and the relevant intellectual property in collaboration with scientists all over the world, DDE categorizes all assets into three distinct tiers. This allows for flexible management of different types of rights.

3.1 Tier I: DDE Core IP (The Foundation)

Definition: IP created or commissioned directly by the DDE Legal Entity

Assets: DDE Brand/Trademarks, the DDE platform architecture, core Knowledge Graphs, and Data Networks (datasets derived from large-scale aggregation).

Ownership: The DDE Legal Entity (e.g., an international NGO¹).

Licensing: DDE grants a Universal Free License for non-commercial research but may charge for high-volume commercial API access to ensure operational sustainability.

3.2 Tier II: Institutional Background IP (The Tools)

Definition: Pre-existing software, proprietary algorithms, or databases owned by partner organizations (e.g., institutions, universities and companies).

Ownership: Remains with the Contributing Institution.

Licensing: Institutions grant DDE a “Right to Use” for research purposes. Commercial exploitation by a DDE user requires separate authorization from the IP owner.

3.3 Tier III: Contributed Scientific Data (The Content)

Definition: The individual datasets uploaded by scientists.

Ownership: The Original Author/Scientist.

Licensing: DDE utilizes the Creative Commons Attribution 4.0 (CC BY 4.0).

Significance: Global reuse is permitted, subject to mandatory attribution to the original contributor.

¹ Such an entity is planning to be established by the end of 2026.

4. Facilitation of Data Sharing

4.1 The Individual vs. The Institution

Researchers are often concerned about personal liability.

The Framework Solution: DDE implements an Institutional Indemnity Clause. Scientists uploading data act in their institutional capacity. Any legal dispute concerning copyright, accuracy, or related issues shall be handled between the institutions, thereby shielding individual scientists from personal liability.

4.2 Data Sovereignty and Cross-Border Flow

Many jurisdictions have strict laws controlling geological data export.

The Framework Solution: The “Federated Regional Node” Model. DDE will not store all data in one central server. Instead, data stays on Local Servers in the home country. DDE sends the Algorithm to the data (Compute-to-Data). Only the Computed Result (e.g., a “Heat Map” of potential minerals) will be shared globally. This ensures compliance with national security laws while enabling global collaboration ²

4.3 The “Prior Rights” Tangle

Geological maps are often “Derivative Works” based on decades of previous maps.

The Framework Solution: Blockchain-enabled Provenance.

DDE will adopt a Distributed Ledger (Blockchain) or high-fidelity Metadata Tags to track every version of a dataset.

When a user downloads a “Final Map”, the system automatically generates a “Reference List” of all contributors whose work fed into that map.

5. Recommended Licensing Strategy

To maximize scientific impact while protecting contributors’ rights, DDE recommends the following four-tier licensing ecosystem:

5.1 Raw Observations: Public Domain (CC0)

For foundational primary data (e.g., rock types, basic stratigraphic records, outcrop

² Need to confirm whether this will be implemented by the key platform providers of DDE.

coordinates), DDE recommends CC0 (Creative Commons Zero).

The Rationale for this recommendation is that basic data are the “atoms” of the platform. If every atom required a complex legal citation, large-scale integration would be impossible. CC0 provides zero barriers to entry and maximum interoperability.

5.2 Research Outputs & Models: Attribution (CC BY 4.0)

This is the Default License for research papers, maps, and refined models.

CC BY 4.0 allows worldwide reuse with mandatory attribution to the original author. This aligns with the “Academic Credit” economy, which provides transparent sharing while securing the contributor’s professional standing.

5.3 Strategic/Sensitive Data: Managed Access

For high-value energy or mineral data that cannot be fully public, DDE recommends a “Data Use Agreement” (DUA) according to which users must apply for access. Data shall be used within a virtual environment and cannot be re-shared. This keeps the balance between collaborative research and national security or commercial sensitivity.

5.4 Software & AI Tools: Permissive Open Source (Apache 2.0 / MIT)

For the processing scripts, AI models, and data-cleaning tools developed within DDE, these licenses encourage broad adoption and technological evolution without “copyleft” restrictions. It also helps to build a self-evolving technological ecosystem.

6. Governance: The DDE IP & Ethics Committee

To ensure the framework is followed, a Permanent IP and Ethics Committee shall be established with the Composition of 1/3 Geoscientists, 1/3 IP Lawyers and 1/3 Data Ethicists. Its main roles include:

- Dispute resolution concerning data or algorithm ownership.
- Review of take-down requests for stolen or sensitive data.
- Compliance audits to prevent unauthorized cross-border data transfer.

Appendix:

Administrative Measures for Data and Intellectual Property Management

(Preliminary Draft)

Chapter 1: General Provisions

Article 1: Purpose

These Measures aim to standardize the management of data and intellectual property (IP) within the DDE Big Science Program, ensure the scientific data security, protect the legitimate rights of contributors, and promote the global sharing and deep processing of geological knowledge.

Article 2: Scope of Application

These measures apply to all DDE Participants, including: Founding and Partner Members (Institutions); Working Groups (WGs) and Task Forces (TFs); Individual Scientists contributing data or using DDE resources; Third-party Service Providers (Cloud hosting, AI developers).

Article 3: Governing Authority

The DDE IP & Data Committee (the "Committee") serves as the supreme operational authority for these measures. It is responsible for reviewing data contribution agreements, arbitrating ownership disputes, and authorizing commercial licensing of DDE core assets.

Chapter 2: Data Classification and Ownership Registration

Article 4: The “Three-Source” Ownership Model

DDE-Generated Assets: IP created using DDE’s central budget or administrative staff belongs to the DDE Secretariat (Legal Entity).

Member-Contributed Assets: Pre-existing data or tools brought into DDE remain the property of the Original Member.

Co-Developed Assets: IP created through collaboration between multiple DDE

members shall be Jointly Owned based on a pre-signed “Project Contribution Statement”.

Article 5: Mandatory Registration

All datasets uploaded to the DDE platform must complete Metadata Registration. No data shall be anonymous. Registration must include: Name of the Principal Investigator (PI); Funding source/Project ID; Applicable restrictive laws (e.g., National Security, GDPR); The chosen DDE License (e.g., CC BY 4.0 or Managed Access).

Chapter 3: Data Lifecycle & Access Control

Article 6: Data Submission and Quality Audit

Before publication on the DDE platform, the data must pass a Technical Validation (ensuring the format matches DDE standards) and an IP Clearance Check (ensuring the contributor has the right to share).

Article 7: The Protection Period

To protect the “First-to-Publish” rights of scientists, contributors may request an Embargo Period of up to 24³ months. During this period, only Metadata are publicly visible. The raw data is encrypted and inaccessible to the public. The Committee may grant extensions for complex long-term field studies.

Article 8: Tiered Access Levels

DDE implements the following tiered protection based on data characteristics:

Level 1 (Open): Fully downloadable by the public (e.g., basic lithology).

Level 2 (Registered): Downloadable only by verified DDE members.

Level 3 (Restricted): Requires a signed Data Use Agreement (DUA) and Committee approval (e.g., high-resolution mineral deposit data).

Chapter 4: Derivatives and Artificial Intelligence (AI) Outputs

Article 9: Deep-Processed Data

³ Need to confirm whether the length is appropriate for DDE participants.

When DDE AI tools process raw data to create Derivative Works (e.g., an AI-generated global tectonic map):

The Original Data Provider retains rights to the raw data.

The DDE Secretariat owns the "Processing Algorithm."

The Resulting Map is owned by DDE but must digitally link back (via DOI) to all original data sources.

Article 10: AI Training Rights

By contributing data to DDE, the contributor grants the DDE AI models the Right to Learn. AI outputs must not disclose restricted or confidential raw data.

Chapter 5: Traceability, Citation, and Rewards

Article 11: The "Digital Pedigree"

The DDE Platform shall implement a Provenance Tracking System. Each time a dataset is used or downloaded, the system logs the event.

Article 12: Mandatory Citation Protocol

Any publication (journal article, conference poster) using DDE data must include the following information:

A formal citation of the original dataset's DOI;

An acknowledgment: "*This research was supported by data provided through the Deep-time Digital Earth (DDE) Program.*"

Article 13: Annual Recognition

DDE will publish an Annual Data Contribution Index and provide awards and priority computational access to high-impact contributors

Chapter 6: Data Sovereignty and Cross-Border Security

Article 14: National Compliance

DDE shall respect the sovereign rights of every member nation over their geological data. If a nation classifies certain data as secret, DDE will implement in-situ computing. The data stays on the nation's local server and only the non-sensitive mathematical model is allowed to cross borders.

Article 15: Red-Line Protocol

If a participant is found to have exported restricted data in violation of domestic controls using DDE tools, DDE shall suspend access and cooperate with relevant authorities and cooperatively provide audit logs to the relevant national authorities.

Chapter 7: Dispute Resolution and Liability

Article 16: Safe Harbor⁴

If a contributor shares data in good faith, DDE provides a Liability Shield. DDE (the organization) will defend the researcher against third-party copyright claims, provided the researcher followed the registration protocols in Chapter 2.

Article 17: Dispute Arbitration

Any IP dispute (e.g., two members claiming ownership of the same algorithm) shall be submitted to the DDE IP Committee. The Committee's decision is final within the DDE ecosystem. Unresolved disputes shall be submitted to the WIPO (World Intellectual Property Organization) Arbitration and Mediation Center.

⁴ This is a very strong clause. DDE committee may need to discuss whether such a clause is necessary for attracting contributors and whether the planned NGO has the resources to provide such a shield.