# Technical Approach

Data Migration Strategy

### Phased Migration with Discovery and Validation

To meet the required deadline for modernization by July 2026, we propose a structured approach that begins with a dedicated Discovery and Planning Phase in collaboration with USGS, followed by three structured migration phases. This methodology ensures minimal disruption, full data fidelity, and compliance with federal IT and data security standards.

#### Discovery and Planning Phase (Q4 2025 – Q1 2026)

Before executing any data migration, we will partner with USGS stakeholders to perform a comprehensive Discovery Phase, which will include:

1. Reviewing the current SharePoint 2019 environment topology and infrastructure
2. Documenting content types, metadata schemas, and retention policies
3. Identifying sensitive content and compliance considerations
4. Validating site inventory, size estimates, and usage patterns (e.g., IndexUsageCount, TotalIO)
5. Defining user groups, permission structures, and access roles
6. Finalizing the ShareGate migration runbook and scheduling strategy

This phase ensures the foundation is aligned with USGS security, operational, and governance requirements.

#### Phase 1: Archival Content Migration (Q1 – Q2 2026)

1. Migrate the WSS\_Content\_IPDSv2\_Archive database (~3.6 TB) to SharePoint Online archive site collections.
2. ShareGate will preserve folder structure, metadata, versioning, and user/group permissions.
3. Bulk transfer will occur during low-activity periods to reduce user disruption.

#### Phase 2: Active Content Libraries Migration (Q2 – Q3 2026)

1. Migrate high-use content databases (Content0 to Content4, ~150–170 GB each).
2. Use ShareGate’s incremental sync capability to maintain live user access during migration.
3. Retain all permissions, document versions, and metadata fidelity throughout the transfer.

#### Phase 3: Home and Utility Site Migration (Q4 2026)

1. Migrate the WSS\_Content\_IPDSv2\_Home site (~1.4 GB), including navigation, landing pages, and branding assets.
2. Validate UI and user experience alignment in SharePoint Online.

#### Data Validation and Service Continuity

1. Utilize ShareGate’s comparison, reporting, and rollback tools to ensure data fidelity.
2. Implement snapshotting and retention strategies in source systems.
3. Execute migration tasks during non-peak periods with proactive communications and training.

### Required Permissions and Access

To conduct the migration securely and efficiently, we will require the following access permissions across environments:

#### Source: SharePoint 2019 (On-Premises)

1. Farm Administrator or Site Collection Administrator access:
   1. To enumerate sites and libraries
   2. To preserve metadata, permissions, and version history
2. SQL Server read access (optional, for usage analytics and planning)

#### Destination: SharePoint Online (Microsoft 365)

1. Global Administrator (initially, to set up permissions and access policies)
2. SharePoint Administrator:
   1. To provision new site collections
   2. To configure retention labels and compliance settings
3. Microsoft Entra ID (Azure AD) Admin:
   1. To configure federated identity, MFA, and conditional access
   2. To create and manage security groups for role-based access

#### Migration Tool Access

1. Licensed use of ShareGate with connection rights to both source and target environments
2. Network access and firewall rules to allow secure communication between environments

All access will be granted on a least-privilege, time-bound basis in accordance with DOI and USGS security requirements.

### Modernized Architecture

Our solution leverages the Microsoft ecosystem to provide a cloud-native, scalable, and compliant digital publishing environment.

#### Platform Components

1. SharePoint Online / OneDrive for document management
2. Power Platform (Power Automate, Power Apps) for workflows and business logic
3. Azure Services:
   1. Azure SQL and Blob Storage for data integration
   2. Azure API Management for interfacing with ScienceBase, SIPP, and DOI tools

#### Performance and Scalability

1. Supports 3,000+ users and 10,000+ annual submissions
2. Scalable architecture with auto-scaling, load balancing, and geo-redundancy
3. Schema optimizations based on usage analytics (I/O and IndexUsageCount)

Workflow and Compliance Automation

We will replicate and modernize all core IPDS workflows using the Microsoft Power Platform.

#### Modernized Workflows

1. Peer review and editorial tracking
2. Approval chains and supervisory sign-offs
3. Document publishing and public release processes

#### Compliance Rules and Business Logic

1. FSP validation and gating
2. Metadata-driven conditional routing
3. Role-based reviewer anonymity enforcement

#### Tracking and Accountability

1. Power BI dashboards for workflow and task status
2. SharePoint version control and audit logs for traceability

### Performance Optimization and Monitoring

#### Optimization Strategies

1. Re-index high-use libraries (Content0–Content4)
2. Optimize metadata schemas for searchability and speed
3. Store archival content separately to minimize retrieval time

#### Monitoring and Analytics

1. Use Azure Monitor, Power BI, and Microsoft 365 analytics to track:
   1. Submission volumes
   2. Workflow efficiency
   3. System latency and bottlenecks

#### Elastic Infrastructure

1. Auto-scaling compute and storage to handle future growth
2. No downtime during scaling or load fluctuations

### Security and Accessibility

#### Security Architecture

1. Federated identity via DOI Azure AD
2. Role-based access control with MFA and conditional access
3. Sensitivity labels and Purview DLP policies for compliance

#### Compliance with Section 508

1. All sites and apps built using Microsoft accessibility standards
2. Compatibility with screen readers, keyboard navigation, and high-contrast modes
3. Testing via automated tools and manual validation

# Innovation and Roadmap (Option Year 2+)

In Option Year 2, we propose enhancing the IPDS platform with AI-driven features aligned with Appendix G priorities.

## User Story–Driven Features

|  |  |  |  |
| --- | --- | --- | --- |
| Role | User Story | Feature | Benefit |
| Author | "Suggest peer reviewers based on topic and history." | AI Reviewer Finder (Graph API + Azure AI) | Accelerates setup, improves quality |
| Reviewer | "Maintain anonymity during review." | Anonymized Workflow Handling (Power Automate + SharePoint perms) | Ensures impartiality and privacy |
| Project Lead | "Track review progress and task owners." | Power BI Task Dashboards | Increases oversight and accountability |
| Compliance Officer | "Automatically check for FSP compliance." | FSP Rule Engine (Power Automate + AI Builder) | Reduces manual QA, ensures adherence |
| Metadata Curator | "Extract metadata automatically on upload." | Microsoft Syntex Integration | Reduces manual input, improves searchability |
| Site Admin | "Monitor system usage and plan for growth." | Azure Monitor + Power Platform Admin Center | Supports proactive scaling and maintenance |
| Records Manager | "Auto-tag and retain archival content per policy." | Microsoft Purview Labels | Ensures retention compliance and streamlines archiving |

## Conclusion

This technical approach offers a secure, scalable, and future-ready migration to Microsoft 365, anchored by a collaborative Discovery Phase with USGS and executed through a phased, risk-managed strategy. Our design prioritizes compliance, performance, user experience, and extensibility. Option Year innovations further align the IPDS platform with USGS’s science delivery mission through AI automation and user-driven enhancements.