

Product Requirements Document (PRD)

Acrux MVP - Research Data Management Platform

Version: 1.0

Last Updated: October 3, 2025

Document Owner: Product Team

Status: Active

Table of Contents

- [1. Executive Summary](#)
 - [2. Product Vision & Goals](#)
 - [3. Target Users](#)
 - [4. User Stories](#)
 - [5. Feature Specifications](#)
 - [6. Success Metrics](#)
 - [7. Release Planning](#)
 - [8. Dependencies & Constraints](#)
 - [9. Open Questions](#)
-

Executive Summary

Acrux MVP is a research data management platform designed to streamline the organization, analysis, and collaboration of research data. The platform addresses the critical challenges faced by researchers, data scientists, and research teams in managing complex datasets, ensuring data integrity, and facilitating collaborative research efforts.

Problem Statement

Research teams currently struggle with:

- Fragmented data storage across multiple platforms
- Lack of standardized data organization
- Difficulty in tracking data provenance and versions
- Limited collaboration tools for distributed teams
- Time-consuming manual data processing tasks

Solution Overview

Acrux provides a centralized platform that:

- Organizes research data with intelligent categorization
- Tracks data lineage and version history

- Enables real-time collaboration
 - Automates routine data processing tasks
 - Ensures data security and compliance
-

Product Vision & Goals

Vision Statement

“To empower researchers with a seamless, intelligent platform that transforms how research data is managed, analyzed, and shared across teams and institutions.”

Product Goals

Primary Goals

1. **Data Organization:** Provide intuitive tools for organizing and categorizing research data
2. **Collaboration:** Enable seamless collaboration among research team members
3. **Data Integrity:** Ensure data accuracy, versioning, and auditability
4. **Accessibility:** Make research data easily discoverable and accessible

Secondary Goals

1. Integration with popular research tools and platforms
2. Advanced analytics and visualization capabilities
3. Compliance with data governance standards
4. Scalability to support growing research teams

Success Criteria

- **Adoption:** 100+ active research teams within 6 months
 - **Engagement:** 80% weekly active users among registered researchers
 - **Satisfaction:** Net Promoter Score (NPS) > 50
 - **Performance:** 99.9% platform uptime
 - **Data Growth:** Support for 10TB+ data storage per organization
-

Target Users

Primary Personas

1. Research Scientist (Dr. Sarah Chen)

- **Role:** Principal Investigator
- **Goals:**
 - Organize multi-year research projects
 - Monitor team progress
 - Ensure data quality and compliance
- **Pain Points:**
 - Too much time spent on data management
 - Difficulty tracking collaborator contributions
 - Concerns about data security

- **Needs:**
- Dashboard overview of all projects
- Access controls and permissions
- Audit trails for data changes

2. Data Analyst (Alex Rodriguez)

- **Role:** Research Data Analyst
- **Goals:**
- Process and analyze research datasets
- Create visualizations and reports
- Identify patterns and insights
- **Pain Points:**
- Inconsistent data formats
- Lack of processing automation
- Limited visualization tools
- **Needs:**
- Data transformation pipelines
- Interactive visualization tools
- Export capabilities for analysis tools

3. Graduate Student (Maria Kim)

- **Role:** PhD Researcher
- **Goals:**
- Contribute to research projects
- Learn data management best practices
- Collaborate with advisors and peers
- **Pain Points:**
- Steep learning curve for new tools
- Limited guidance on data organization
- Difficulty accessing relevant datasets
- **Needs:**
- Intuitive user interface
- Documentation and tutorials
- Search and discovery features

Secondary Personas

- **Lab Manager:** Handles administrative tasks, budget tracking, resource allocation
 - **IT Administrator:** Manages platform deployment, security, and integrations
 - **Research Collaborator:** External partners with limited access needs
-

User Stories

Epic 1: Data Management

Story 1.1: Upload Research Data

As a research scientist

I want to upload multiple data files at once

So that I can quickly populate my research project

Acceptance Criteria:

- Support batch upload of up to 100 files
- Accept common file formats (CSV, Excel, JSON, PDF, images)
- Show upload progress and completion status
- Validate file types and sizes
- Maximum file size: 500MB per file

Priority: P0 (Must Have)

Story Points: 5

Story 1.2: Organize Data into Collections

As a data analyst

I want to organize related datasets into collections

So that I can maintain logical groupings of my research data

Acceptance Criteria:

- Create custom collections with names and descriptions
- Add/remove datasets to/from collections
- Support nested collections (subcollections)
- Bulk operations for moving multiple datasets
- Visual hierarchy display

Priority: P0 (Must Have)

Story Points: 8

Story 1.3: Tag and Categorize Data

As a graduate student

I want to add tags and metadata to datasets

So that I can easily find and filter relevant data later

Acceptance Criteria:

- Add multiple tags to any dataset
- Auto-suggest existing tags while typing
- Filter datasets by single or multiple tags
- Edit and delete tags
- Tag analytics (most used, recent tags)

Priority: P1 (Should Have)

Story Points: 5

Epic 2: Search and Discovery

Story 2.1: Search Across All Data

As a research scientist

I want to search across all my project data

So that I can quickly find specific datasets or information

Acceptance Criteria:

- Full-text search across dataset names, descriptions, and metadata
- Search within file contents where applicable
- Real-time search results (< 2 seconds)
- Highlight matching terms in results
- Filter results by date, type, collection

Priority: P0 (Must Have)

Story Points: 13

Story 2.2: Advanced Filtering

As a data analyst

I want to apply complex filters to narrow down datasets

So that I can focus on specific subsets of data for analysis

Acceptance Criteria:

- Filter by multiple attributes simultaneously
- Date range filters
- File type filters
- Tag-based filters
- Save filter presets for reuse

Priority: P1 (Should Have)

Story Points: 8

Epic 3: Collaboration

Story 3.1: Share Projects with Team Members

As a research scientist

I want to invite team members to collaborate on projects

So that we can work together on research data

Acceptance Criteria:

- Send invitations via email
- Set role-based permissions (viewer, editor, admin)
- View list of all collaborators
- Remove collaborators when needed
- Pending invitation status tracking

Priority: P0 (Must Have)

Story Points: 13

Story 3.2: Comment on Datasets

As a graduate student

I want to add comments and discussions to datasets

So that I can communicate with team members about specific data

Acceptance Criteria:

- Add threaded comments to any dataset
- Mention team members with @username
- Email notifications for new comments
- Edit and delete own comments
- Mark comments as resolved

Priority: P1 (Should Have)**Story Points:** 8**Story 3.3: Activity Feed****As a** research scientist**I want to** see recent activities in my projects**So that** I can stay updated on team progress and changes**Acceptance Criteria:**

- Display chronological activity feed
- Show uploads, edits, comments, shares
- Filter by activity type
- User profile pictures in feed
- Link to affected datasets

Priority: P2 (Nice to Have)**Story Points:** 5**Epic 4: Data Analysis****Story 4.1: Preview Data Files****As a** data analyst**I want to** preview data files without downloading**So that** I can quickly verify content before analysis**Acceptance Criteria:**

- Preview CSV/Excel files in table format
- Display first 1000 rows
- Preview images, PDFs in browser
- Basic sorting and filtering in preview
- Download option from preview

Priority: P0 (Must Have)**Story Points:** 8**Story 4.2: Create Data Visualizations****As a** data analyst**I want to** create charts and graphs from datasets**So that** I can visualize trends and patterns**Acceptance Criteria:**

- Support bar, line, scatter, pie charts
- Interactive chart configuration
- Save charts to project
- Export charts as images
- Responsive chart display

Priority: P1 (Should Have)

Story Points: 13

Epic 5: Administration

Story 5.1: Manage User Permissions

As a research scientist

I want to control who can view or edit specific data

So that I can maintain data security and privacy

Acceptance Criteria:

- Set permissions at project and dataset levels
- Role-based access control (RBAC)
- Inherit permissions from parent collections
- Override inherited permissions when needed
- Audit log of permission changes

Priority: P0 (Must Have)

Story Points: 13

Story 5.2: Monitor Storage Usage

As a lab manager

I want to view storage usage across projects

So that I can manage budget and resources effectively

Acceptance Criteria:

- Dashboard showing total storage used
- Breakdown by project
- Historical usage trends
- Alerts for approaching limits
- Export usage reports

Priority: P2 (Nice to Have)

Story Points: 5

Feature Specifications

Feature 1: Project Dashboard

Description: Central hub for researchers to view and manage all their research projects.

Components:

- Project cards with thumbnails and key metrics
- Quick actions (upload, search, share)
- Recent activity summary
- Storage usage indicator
- Create new project button

User Flow:

1. User logs in to platform
2. Lands on dashboard showing all projects

3. Can filter/sort projects by various criteria
4. Clicks on project card to enter project workspace

Technical Requirements:

- Load dashboard in < 3 seconds
- Support pagination for 100+ projects
- Real-time updates for activity feed
- Responsive design for mobile/tablet

Design Considerations:

- Clean, modern interface
- Color-coded project status indicators
- Accessible navigation
- Dark mode support

Feature 2: Data Upload System

Description: Robust file upload system supporting batch operations and various file formats.

Components:

- Drag-and-drop upload zone
- Progress indicators for each file
- File validation and error handling
- Upload queue management
- Metadata form for uploaded files

User Flow:

1. User navigates to project workspace
2. Clicks upload button or drags files to drop zone
3. Selects files from device (supports multiple selection)
4. System validates files and shows preview
5. User adds metadata (optional)
6. Confirms upload and monitors progress
7. Receives confirmation upon completion

Technical Requirements:

- Chunked upload for large files
- Resume capability for interrupted uploads
- Virus scanning for all uploads
- Automatic thumbnail generation for images
- Support for 50+ file types

Design Considerations:

- Clear visual feedback during upload
- Error messages with actionable guidance
- Mobile-optimized upload experience
- Keyboard accessible

Feature 3: Advanced Search Engine

Description: Powerful search functionality to find datasets across projects quickly.

Components:

- Search bar with auto-complete

- Filter sidebar
- Results list with previews
- Sort options
- Saved searches

User Flow:

1. User enters search query in search bar
2. Results appear in real-time as they type
3. Can apply filters to narrow results
4. Clicks on result to view dataset details
5. Option to save search for future use

Technical Requirements:

- Full-text search across metadata
- Faceted search with multiple filters
- Search results in < 2 seconds
- Support for Boolean operators
- Typo tolerance and fuzzy matching

Design Considerations:

- Prominent search bar placement
- Clear result count and filters
- Highlighted search terms in results
- Empty state guidance

Feature 4: Collaboration Workspace

Description: Suite of tools enabling team collaboration on research projects.

Components:

- Team member list with roles
- Invitation system
- Comment threads on datasets
- Activity notifications
- Permission management interface

User Flow:

1. Project owner invites team members via email
2. Invited users receive notification and accept
3. Team members can view/edit based on permissions
4. Users add comments and tag others
5. Activity feed shows all team actions
6. Owner can modify permissions as needed

Technical Requirements:

- Real-time collaboration updates
- Email notification system
- Role-based access control
- Conflict resolution for simultaneous edits
- Activity log retention (90 days)

Design Considerations:

- Clear permission indicators

- Intuitive comment threading
- Prominent notification badges
- User avatars for personalization

Feature 5: Data Visualization Builder

Description: Interactive tool for creating charts and graphs from research data.

Components:

- Chart type selector
- Data mapping interface
- Customization panel
- Preview pane
- Export options

User Flow:

1. User selects dataset for visualization
2. Chooses chart type (bar, line, scatter, etc.)
3. Maps data columns to chart axes
4. Customizes colors, labels, legends
5. Previews chart in real-time
6. Saves chart to project or exports

Technical Requirements:

- Support 10+ chart types
- Handle datasets up to 100K rows
- Responsive and interactive charts
- Export to PNG, SVG, PDF formats
- Chart embedding capability

Design Considerations:

- Intuitive drag-and-drop mapping
- Live preview as options change
- Professional color palettes
- Accessibility compliance (WCAG 2.1)

Success Metrics

North Star Metric

Weekly Active Research Teams: Number of research teams actively using the platform each week

Key Performance Indicators (KPIs)

Adoption Metrics

Metric	Target	Measurement
New User Signups	50/week	Weekly
Team Activation Rate	70%	Monthly
Time to First Upload	< 10 min	Per user
Projects Created	200+/month	Monthly

Engagement Metrics

Metric	Target	Measurement
Weekly Active Users (WAU)	80%	Weekly
Daily Active Users (DAU)	45%	Daily
Avg. Session Duration	15+ min	Daily
Data Uploads per Week	1000+	Weekly
Collaboration Invites	5/project	Monthly

Quality Metrics

Metric	Target	Measurement
Platform Uptime	99.9%	Real-time
Page Load Time	< 3 sec	Real-time
Search Response Time	< 2 sec	Real-time
Error Rate	< 0.1%	Daily
Customer Support Tickets	< 50/week	Weekly

Satisfaction Metrics

Metric	Target	Measurement
Net Promoter Score (NPS)	50+	Quarterly
Customer Satisfaction (CSAT)	4.5/5	Post-interaction
Feature Satisfaction	4/5	Per feature
Onboarding Completion	85%	Monthly

Business Metrics

Metric	Target	Measurement
Customer Acquisition Cost	< \$100	Monthly
Customer Lifetime Value	> \$2000	Quarterly
Conversion Rate (Free to Paid)	25%	Monthly
Churn Rate	< 5%	Monthly
Monthly Recurring Revenue (MRR)	Growth trend	Monthly

Analytics Implementation

Tools:

- Google Analytics for user behavior tracking
- Mixpanel for event-based analytics
- Hotjar for user session recordings
- Custom dashboard for business metrics

Key Events to Track:

- User registration and login
 - Project creation
 - File uploads and downloads
 - Search queries
 - Collaboration invitations
 - Feature usage (visualization, comments, etc.)
 - Error occurrences
 - Subscription upgrades
-

Release Planning

Phase 1: MVP Launch (Months 1-3)

Goal: Launch minimum viable product with core features

Features:

- ☒ User authentication and authorization
- ☒ Project creation and management
- ☒ File upload and storage
- ☒ Basic search functionality
- ☒ Simple data preview
- ☒ Team collaboration (invite members)
- ☒ Role-based permissions

Success Criteria:

- 50+ active research teams
- 500+ datasets uploaded
- 70% user activation rate
- NPS > 40

Phase 2: Enhanced Collaboration (Months 4-6)

Goal: Strengthen collaboration features and user engagement

Features:

- Comment threads on datasets
- Activity feed and notifications
- Advanced search with filters
- Data tagging and categorization
- Collection organization
- Enhanced file preview
- Mobile responsive design

Success Criteria:

- 100+ active research teams
- 3000+ datasets uploaded
- 80% weekly active user rate
- NPS > 50

Phase 3: Analytics & Insights (Months 7-9)

Goal: Provide powerful analytics and visualization capabilities

Features:

- Data visualization builder
- Chart export functionality
- Data transformation tools
- Automated insights
- Dashboard customization
- Integration with analysis tools
- API access for developers

Success Criteria:

- 200+ active research teams
- 10000+ datasets uploaded
- 500+ visualizations created
- NPS > 55

Phase 4: Enterprise Features (Months 10-12)

Goal: Scale platform for enterprise research organizations

Features:

- Single Sign-On (SSO)
- Advanced security features
- Compliance reporting
- Custom branding
- Dedicated support
- SLA guarantees
- Advanced user analytics

Success Criteria:

- 10+ enterprise customers
 - 500+ active research teams
 - 99.95% uptime
 - NPS > 60
-

Dependencies & Constraints

Technical Dependencies

Infrastructure:

- Cloud hosting provider (AWS/Azure/GCP)
- Object storage service (S3 or equivalent)
- Database service (PostgreSQL)
- CDN for global content delivery
- Email service provider

Third-Party Services:

- Authentication service (NextAuth)
- File processing libraries
- Charting libraries (Plotly, Chart.js)
- Search engine (Elasticsearch - future)
- Analytics platforms

Integration Requirements:

- REST API development
- OAuth2 authentication
- Webhook support
- Data export APIs

Resource Constraints

Budget:

- Cloud infrastructure costs
- Third-party service fees
- Development team salaries
- Marketing and customer acquisition

Timeline:

- 12-month roadmap with quarterly releases
- 2-week sprint cycles
- Buffer time for bug fixes and technical debt

Team:

- 2 Full-stack developers
- 1 UI/UX designer
- 1 Product manager
- 1 QA engineer (part-time)

Regulatory & Compliance

Data Privacy:

- GDPR compliance for EU users
- CCPA compliance for California users
- Data encryption at rest and in transit
- Regular security audits

Research Standards:

- Support for data management plans
- Metadata standards (Dublin Core)
- Research data citation formats
- Institutional repository integration

Assumptions

- Users have stable internet connectivity
- Primary usage on desktop/laptop devices
- Research teams range from 3-20 members
- Average dataset size is 10-100MB
- Users are familiar with basic data management concepts

Risks

Risk	Probability	Impact	Mitigation
Slow user adoption	Medium	High	Aggressive marketing, free tier, user onboarding
Technical scalability issues	Medium	High	Load testing, scalable architecture, monitoring
Data security breach	Low	Critical	Regular audits, penetration testing, insurance
Key team member departure	Medium	Medium	Documentation, knowledge sharing, backup resources
Competitive pressure	High	Medium	Unique features, superior UX, community building

Open Questions

Product Questions

1. Should we support real-time collaborative editing like Google Docs?
2. What level of data analysis should be built-in vs. relying on integrations?
3. Should we develop mobile apps or focus on responsive web only?
4. How much storage should be included in free vs. paid tiers?

Technical Questions

1. Which cloud provider offers the best balance of cost and performance?
2. Should we build a microservices architecture or monolithic initially?
3. What's the strategy for handling very large file uploads (>1GB)?
4. How do we ensure data consistency in real-time collaboration scenarios?

Business Questions

1. What's the optimal pricing strategy for academic vs. commercial users?
2. Should we pursue partnerships with research institutions?
3. What's the customer acquisition strategy beyond digital marketing?
4. Should we offer white-label solutions for enterprises?

User Experience Questions

1. What onboarding flow will ensure highest activation rate?

- 2. How much guidance/tutorials should be built into the interface?
- 3. Should we support keyboard shortcuts for power users?
- 4. What’s the right balance between features and simplicity?

Document History

Version	Date	Author	Changes
1.0	Oct 3, 2025	Product Team	Initial PRD creation

Appendix

Glossary

- **Dataset:** A collection of related data files and associated metadata
- **Collection:** A logical grouping of datasets within a project
- **Project:** A workspace containing datasets, collections, and collaborators
- **Collaborator:** A user with permissions to access and/or edit project data
- **Metadata:** Descriptive information about a dataset (tags, description, etc.)


References

- Research Data Management Best Practices (Nature, 2023)
- FAIR Data Principles (GO FAIR Initiative)
- Data Management Plan Guidelines (NSF, NIH)
- Competitive Analysis: OSF, Figshare, Mendeley Data

Feedback & Contributions

This is a living document. Feedback and suggestions are welcome:

- **Email:** product@acrux.dev
- **Slack:** #product-feedback channel
- **GitHub:** Create an issue with label `prd-feedback`

Document Status:  Active
Next Review: January 2026
Owner: Product Manager - Acrux Team