VEDA Services

User Personas & Service Bundles Guide

Executive Summary

This document defines five distinct user personas that represent the primary audiences for VEDA Services. Each persona captures unique needs, workflows, and pain points that inform how VEDA services should be packaged and delivered. Following the persona definitions, we present recommended service bundles optimized for each user type.

High Level Summary Tags

Section 1: User Personas

Persona 1: Data Dana

The Data-Hungry Scientist

Role	Research Scientist / Data Analyst
Organization	University, Research Institution, or Federal Agency
Technical Level	Advanced (Python, R, GIS tools)
Primary Goal	Access and analyze large-scale satellite imagery for research publications

Profile Description

Data Dana is a mid-career scientist who spends most of her time analyzing Landsat and Sentinel-2 imagery to study environmental changes. She is comfortable writing Python code but frustrated by the hours spent downloading, reformatting, and preprocessing data before she can begin actual analysis. Dana needs programmatic access to analysis-ready data and scalable computing resources to process multi-decade time series.

Key Needs

- Cloud-optimized data formats (COG, Zarr) to eliminate download bottlenecks
- STAC catalog for discovering and querying available datasets
- JupyterHub environment with pre-configured libraries and scalable compute
- Example Jupyter Notebooks to accelerate her learning curve
- Help desk support when encountering technical issues

Pain Points

- Time lost to data wrangling instead of science
- Local computing resources insufficient for large-scale analysis
- Dependency conflicts when setting up analysis environments

Persona 2: Publisher Pete

The Mission-Driven Program Manager

Role	Program Manager / Data Steward
Organization	Federal Agency (NASA, NOAA, USGS)
Technical Level	Intermediate (understands data concepts, limited coding)
Primary Goal	Make agency datasets publicly accessible and discoverable

Profile Description

Publisher Pete manages a portfolio of Earth observation datasets that his agency is mandated to share publicly. He understands data management principles but relies on technical teams for implementation. Pete needs a streamlined pathway to transform archival data into modern, discoverable formats and host them reliably. He cares deeply about data quality documentation and ensuring users can actually find and use what he publishes.

Key Needs

- Data transformation services to convert legacy formats to cloud-optimized standards
- STAC catalog ingestion and reliable AWS S3 hosting
- Content curation support including data quality reports
- Automated pipelines to update datasets when new versions release
- Data recovery capabilities for business continuity

Pain Points

- Complex procurement processes to acquire technical services
- Lack of in-house expertise for cloud data infrastructure
- Pressure to demonstrate public value of agency data investments

Persona 3: Visualizer Vicky

The Storytelling Developer

Role	Frontend Developer / Data Visualization Specialist
Organization	Science Communication Team, NGO, or Media Organization
Technical Level	Advanced (JavaScript, React, mapping libraries)
Primary Goal	Build compelling interactive visualizations that communicate complex data to broad audiences

Profile Description

Visualizer Vicky transforms dense scientific datasets into engaging web experiences. She builds tools like the Fire Event Explorer and EMIT Methane Plume viewer that let non-technical audiences explore satellite data intuitively. Vicky needs robust APIs for tiles and statistics, flexible website infrastructure, and close collaboration with scientists to ensure accuracy while maintaining accessibility.

Key Needs

- OGC Tiles and WMTS APIs for performant map rendering
- Statistics APIs for on-the-fly data summarization
- Custom visualization tool development support
- Website design, development, and hosting with CMS flexibility
- Iterative stakeholder collaboration during development

Pain Points

- Slow or unreliable tile services that degrade user experience
- Limited documentation for geospatial APIs
- Disconnect between data providers and visualization requirements

Persona 4: Infrastructure Ian

The Platform Architect

Role	DevOps Engineer / Cloud Architect
Organization	IT Division of Research Institution or Agency
Technical Level	Expert (AWS, Terraform, Kubernetes, CI/CD)
Primary Goal	Deploy and maintain reliable, cost-effective cloud infrastructure for data services

Profile Description

Infrastructure Ian keeps the lights on. He manages the AWS resources, Terraform configurations, and CI/CD pipelines that power data services. Ian needs infrastructure that is well-documented, easily reproducible, and cost-transparent. He values automation, monitoring, and the ability to quickly deploy updates without downtime. Ian often serves as the bridge between scientists who need resources and organizational leadership who need cost accountability.

Key Needs

- AWS/Terraform/CI/CD management services with infrastructure-as-code
- Apache Airflow for orchestrating reusable data workflows
- Grafana integration for cost monitoring and management
- Modular, re-deployable components for rapid updates
- Access management via Keycloak or GitHub groups

Pain Points

- Unpredictable cloud costs that blow budgets
- Technical debt from poorly documented legacy systems
- Balancing security requirements with user accessibility

Persona 5: Community Carlos

The Collaborative Connector

Role	Graduate Student / Early-Career Researcher / Educator
Organization	University, Community College, or Citizen Science Group
Technical Level	Beginner to Intermediate (learning Python, new to cloud)
Primary Goal	Learn satellite data analysis skills and connect with the user community

Profile Description

Community Carlos is eager to learn but often overwhelmed by the complexity of geospatial data ecosystems. He benefits enormously from example notebooks, community forums, and accessible documentation. Carlos represents the next generation of data users and is often the canary in the coal mine for usability issues. He needs low barriers to entry, responsive help when stuck, and a welcoming community to ask questions without judgment.

Key Needs

- JupyterHub access with pre-configured environments (no local setup required)
- Curated Jupyter Notebooks with step-by-step tutorials
- Help desk support for troubleshooting
- Community forums (GitHub Discussions, Discourse) for peer learning
- GitHub-managed documentation and learning resources

Pain Points

- Steep learning curve for cloud-native geospatial workflows
- Difficulty knowing where to start or what questions to ask
- Limited computing resources at their institution

Section 2: Recommended Service Bundles

Based on the persona analysis above, we recommend the following service bundles. Each bundle is optimized to address the core needs of its target persona while providing a coherent, integrated service experience.

Bundle 1: Researcher Accelerator

Target Persona: Data Dana

Bundle Overview

The Researcher Accelerator bundle eliminates data preparation bottlenecks so scientists can focus on analysis. It combines cloud-optimized data access, scalable computing environments, and expert support into a seamless research platform.

Included Services

Service Category	Specific Services
Data + Data APIs	Cloud-Optimized GeoTIFFs, Zarr stores, STAC catalog access, Statistics API
Cloud Computing	JupyterHub deployment with scalable storage/memory, library dependency resolution
Content Curation	Example Jupyter Notebooks for data access and analysis workflows
User Services	Help desk support, community forum access

Value Proposition

Researchers can go from hypothesis to analysis in hours instead of weeks. Pre-configured environments eliminate dependency issues, while cloud-optimized formats enable processing of datasets that would overwhelm local machines.

Bundle 2: Data Publishing Suite

Target Persona: Publisher Pete

Bundle Overview

The Data Publishing Suite provides end-to-end support for agencies needing to transform archival data into discoverable, cloud-hosted public resources. It handles the technical complexity of format conversion, catalog management, and ongoing maintenance.

Included Services

Service Category	Specific Services

Data Transformation	Conversion to COG, Geoparquet, native Zarr, virtual Zarr indexing
Ingestion & Hosting	STAC catalog creation, AWS S3 storage, data recovery
Content Curation	Dataset documentation, quality reports, automation pipelines for updates
Data APIs	OGC Tiles, WMTS, statistics endpoints for downstream users

Value Proposition

Agencies can fulfill open data mandates without building internal cloud expertise. Automated update pipelines ensure datasets stay current, while quality reports demonstrate data stewardship to stakeholders.

Bundle 3: Visualization Studio

Target Persona: Visualizer Vicky

Bundle Overview

The Visualization Studio empowers developers to build stunning, performant data visualization experiences. It combines robust APIs, custom tool development, and full-stack website support to turn complex scientific data into accessible public engagement tools.

Included Services

Service Category	Specific Services
Data APIs	OGC Tiles, WMTS, statistics APIs optimized for visualization
Custom Visualization	Tool development for station, raster, gridded, and point cloud data
Featured Tools	Fire Event Explorer, EMIT Methane Plume, Global Mangrove Distribution
Website Services	Design, development, CMS options (Drupal, Tina, Strapi, WordPress, GitHub)

Value Proposition

Science communicators can create world-class interactive experiences without building backend infrastructure from scratch. The iterative development process ensures final products meet both scientific accuracy and user experience standards.

Bundle 4: Platform Operations

Target Persona: Infrastructure Ian

Bundle Overview

Platform Operations provides the DevOps foundation for reliable, cost-effective data services. It includes infrastructure management, workflow orchestration, monitoring, and access control to keep platforms running smoothly.

Included Services

Service Category	Specific Services
Infrastructure	AWS/Terraform/CI/CD management, re-deployable components
Workflow Orchestration	Apache Airflow for reusable data processing workflows
Monitoring	Grafana integration for cost management and service health

Access Management	Keycloak and GitHub groups for authentication and authorization

Value Proposition

IT teams get production-ready infrastructure patterns without reinventing the wheel. Cost transparency through Grafana enables proactive budget management, while infrastructure-as-code ensures reproducibility and disaster recovery.

Bundle 5: Community Launchpad

Target Persona: Community Carlos

Bundle Overview

Community Launchpad lowers barriers to entry for newcomers to cloud-native geospatial analysis. It combines accessible computing environments, curated learning materials, and community support channels to help the next generation of users get started.

Included Services

Service Category	Specific Services
Cloud Computing	JupyterHub access with pre-configured environments
Content Curation	Tutorial Jupyter Notebooks, GitHub Pages documentation
User Services	Help desk support for troubleshooting
Community Forums	GitHub Discussions, Discourse boards for peer learning

Value Proposition

New users can start analyzing real satellite data on day one without installing software locally. Curated tutorials provide guided learning paths, while community forums create peer support networks that scale beyond one-on-one help.

Section 3: Service Bundle Summary Matrix

The following matrix provides a quick reference for mapping personas to their recommended service bundles and primary service categories.

Persona	Bundle	Primary Focus	Core Services
Data Dana	Researcher Accelerator	Analysis	Data APIs, JupyterHub, Notebooks
Publisher Pete	Data Publishing Suite	Distribution	Transformation, Hosting, Curation
Visualizer Vicky	Visualization Studio	Communication	APIs, Custom Tools, Website
Infrastructure Ian	Platform Operations	Reliability	AWS/Terraform, Airflow, Grafana
Community Carlos	Community Launchpad	Education	JupyterHub, Tutorials, Forums

Conclusion

These five personas and corresponding service bundles provide a framework for understanding how different user types engage with VEDA Services. While individual users may span multiple personas, this segmentation helps prioritize feature development, tailor documentation, and create coherent service packages that address real workflow needs.

The bundles are not mutually exclusive. Organizations often require elements from multiple bundles, for example an agency might need the Data Publishing Suite to share their data, the Visualization Studio to communicate findings, and Platform Operations to maintain the infrastructure. The bundle framework provides a starting point for conversations about which services best fit each engagement.