



SPA MDR CDR SCE

Accelerating Regulatory Submissions with R: A Scalable and Compliant Approach

Priyanka Sawant
Sycamore Informatics



The Clinical Data
Science Conference

APAC 2026

19–21 February
Novotel Hyderabad International Convention Centre



Agenda



The Regulatory Landscape - FDA-R Consortium Pilots



Current Adoption Challenges - Why containerization matters



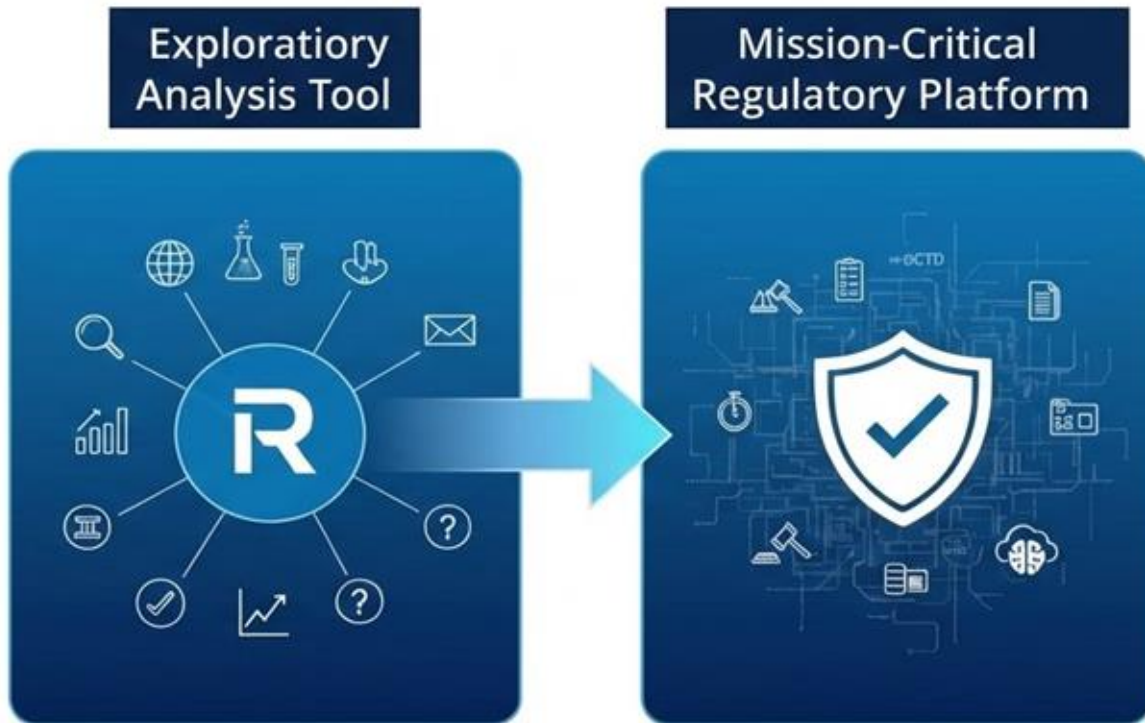
Modern Statistical Computing Environment - The holistic solution



The Path Forward - Strategic implications

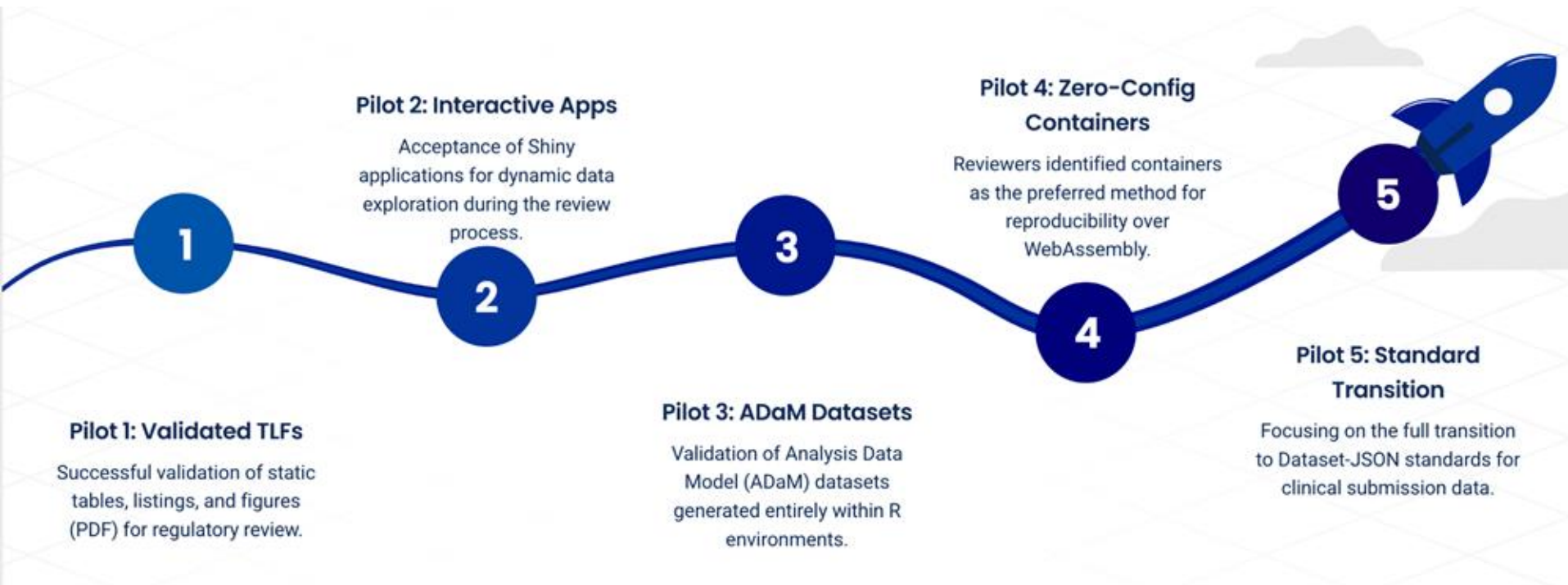


Pharmaceutical Industry at a Pivotal Point



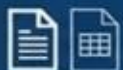
The FDA-R Consortium Pilot Evolution

Pilots progressed from static PDFs to interactive, containerized, reviewer-ready R workflows and modern data standards.



Regulatory Framework - Today & Going Forward

What is Accepted Today



R-scripts and .rds files in eCTD submissions



Package-based workflows with proper validation



Interactive and static outputs



What is Required Going Forward



✓ Complete reproducibility of all analyses



✓ Transparent documentation of computational environment



✓ Validation evidence for all software components



✓ Clear traceability linking code → environment → data → outputs

Why R Adoption Remains Challenging



Dependency Complexity Crisis



Validation Documentation Burden



Cross Platform Inconsistencies



Security & Compliance Concerns



Reviewer Accessibility Issues

Challenge: Dependency Complexity Crisis

The Ecosystem:

- 20,000+ packages on CRAN
- Thousands more on Bioconductor/GitHub
- Each package has multiple dependencies

The Problems:

- **Version Conflicts:** Package A needs C v1.5+, Package B needs C v1.4.x
- **Transitive Dependencies:** 10 direct packages → 50-100 total packages needed
- **Temporal Instability:** Today's "current" version ≠ next month's "current" version
- **Platform Dependencies:** System libraries vary across OS

Reality Check:

A modest analysis requiring validation for 50 packages = thousands of pages of documentation

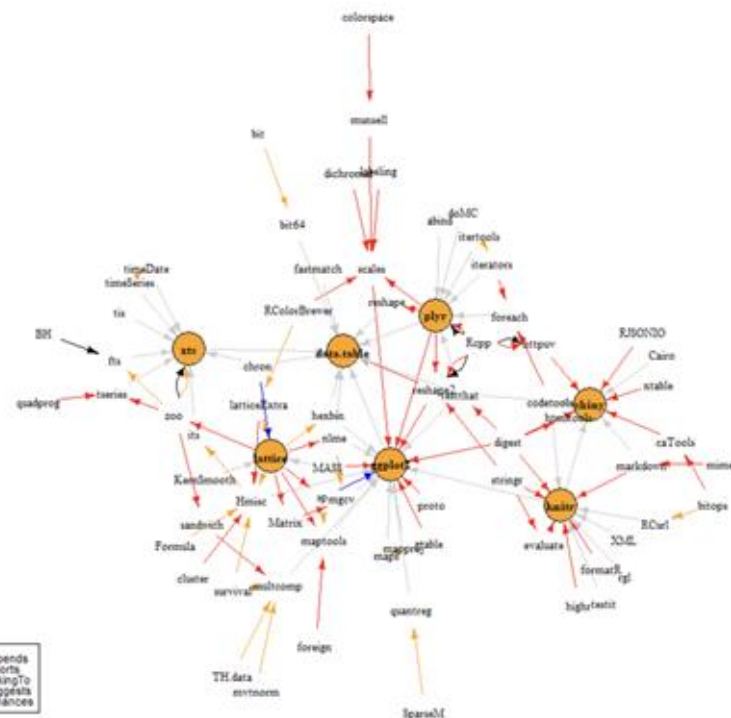


Image credits - CRAN



Challenge: Validation Documentation Burden

Traditional Validation Approach:

For EACH of 50 packages:

- Installation Qualification (IQ)
- Operational Qualification (OQ)
- Performance Qualification (PQ)
- Risk Assessment
- Traceability Matrix

Additional Burdens:

- **Revalidation Cycles:** Every package update requires review
- **Multi-Environment Qualification:** Validate R on dev, test, validation, submission environments
- **Audit Trail Management:** Manual tracking of versions = error-prone

The Scaling Problem: Traditional approaches designed for monolithic software (SAS) don't scale to R's modular ecosystem



Image credits - CRAN



Challenge: Cross-Platform & Cross-Team

"Works on My Machine" Syndrome:

- ✓ Development environment: Analysis runs perfectly
- X Validation environment: Fails due to subtle configuration differences
- 🕒 Hours/days spent troubleshooting environmental issues

Global Collaboration Challenges:

- CRO coordination: Matching sponsor environments requires extensive documentation
- Multiple continents: Different OS, IT policies, infrastructure
- Legacy maintenance: Reproducing 5-year-old analyses on modern infrastructure

The Cost: Discrepancies discovered late in timelines can delay submissions



Challenge: Security & Compliance Concerns

Open-Source Risks:

- **Supply Chain Risk:** Packages installed directly from CRAN without formal vetting
- **Provenance Verification:** How do we prove packages weren't modified in transit?
- **Vulnerability Management:** When CVEs discovered, must assess impact across all projects

Regulatory Requirements:

- **21 CFR Part 11 Compliance:** Audit trails, e-signatures, system validation
- **Data Retention:** 20+ year accessibility and reproducibility requirements

IT Department Reality: Risk-averse pharmaceutical IT struggles with community-driven, rapidly-evolving open-source ecosystem



Challenge 5: Reviewer Accessibility



The Reviewer's Perspective:

- Diverse IT environments with varying restrictions
- Security policies: No internet access, limited installation privileges
- Limited R programming/system administration expertise
- Strict review timelines

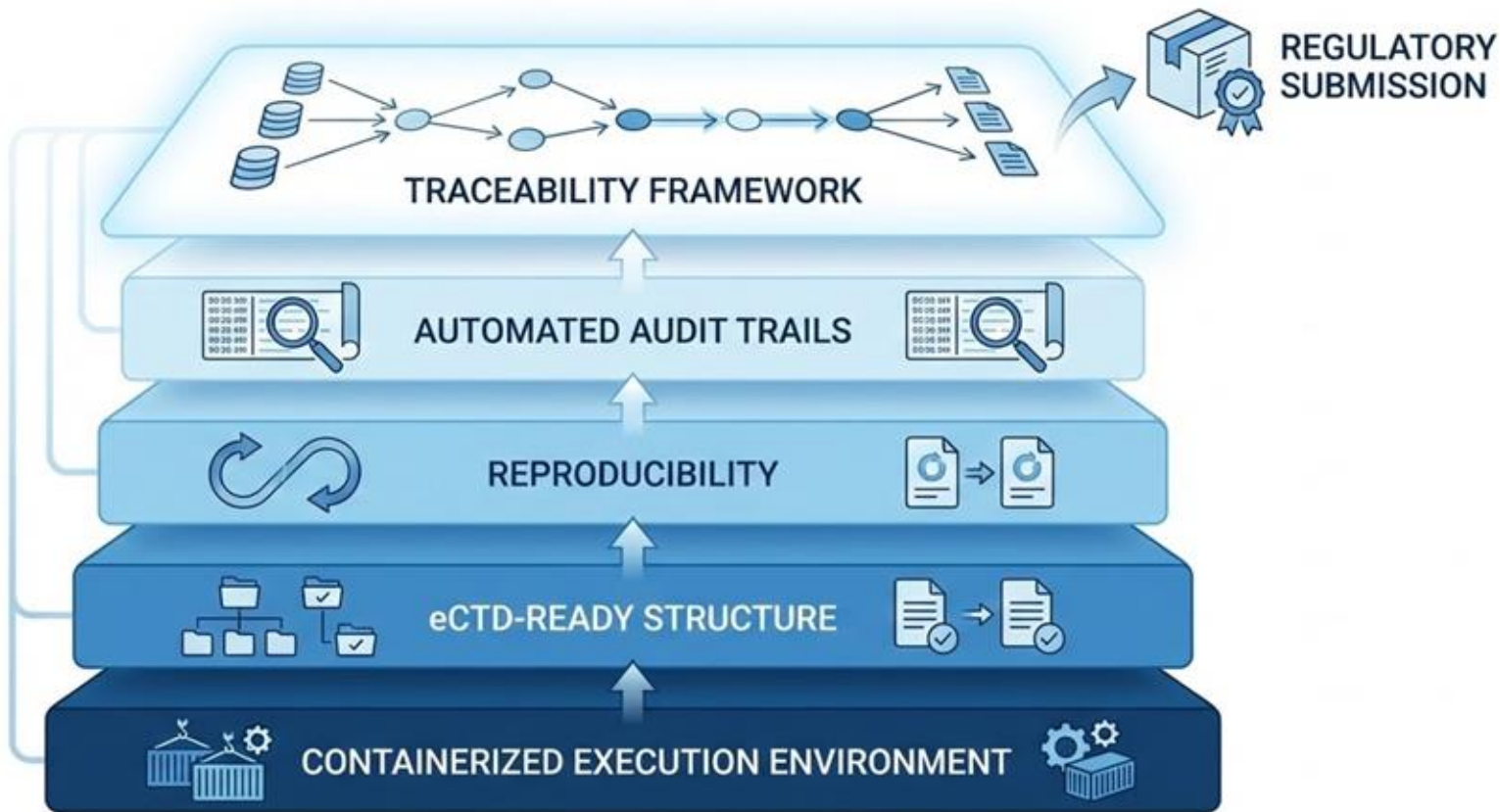
The Sponsor's Dilemma:

- How to provide analyses that work across different reviewer systems?
- How to avoid hours of environment setup consuming review time?
- How to make R-based submissions as easy to review as traditional submissions?

Pilot 4 Finding: Reviewers prefer self-contained "bundle" environments with **zero configuration required**



The Modern Statistical Computing Environment



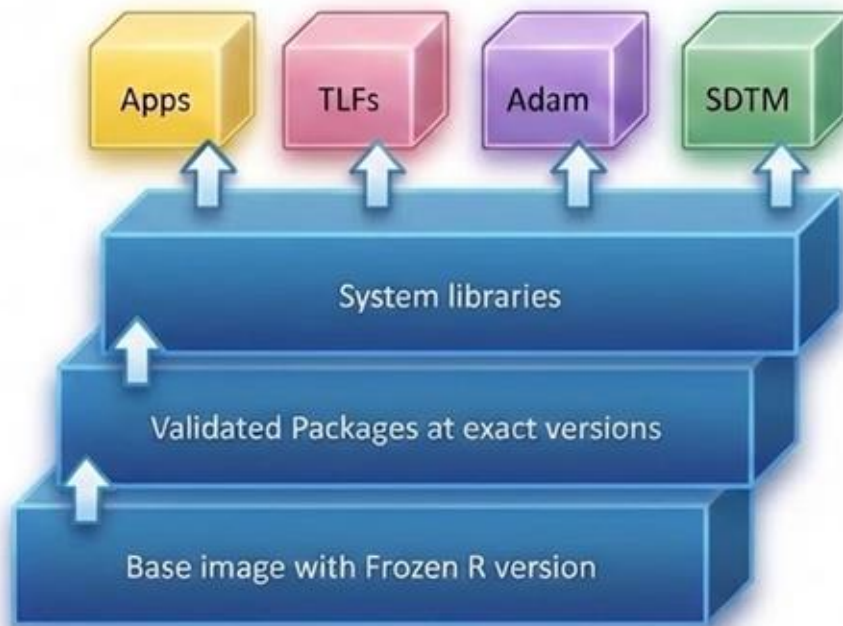
Component 1: Containerized Execution Environment

What Containers Provide:

- Frozen R version (e.g., R 4.3.2)
- Validated package libraries at exact versions
- Complete dependency resolution (system libraries include)
- Isolation from host system
- Platform portability (works on any compatible system)

Benefits:

- ✓ Eliminates "works on my machine" syndrome
- ✓ Guarantees identical execution across environments
- ✓ Enables zero-configuration reviewer access
- ✓ Supports 20+ year archival requirements



Component 2: Inbuilt Audit Trails & Traceability

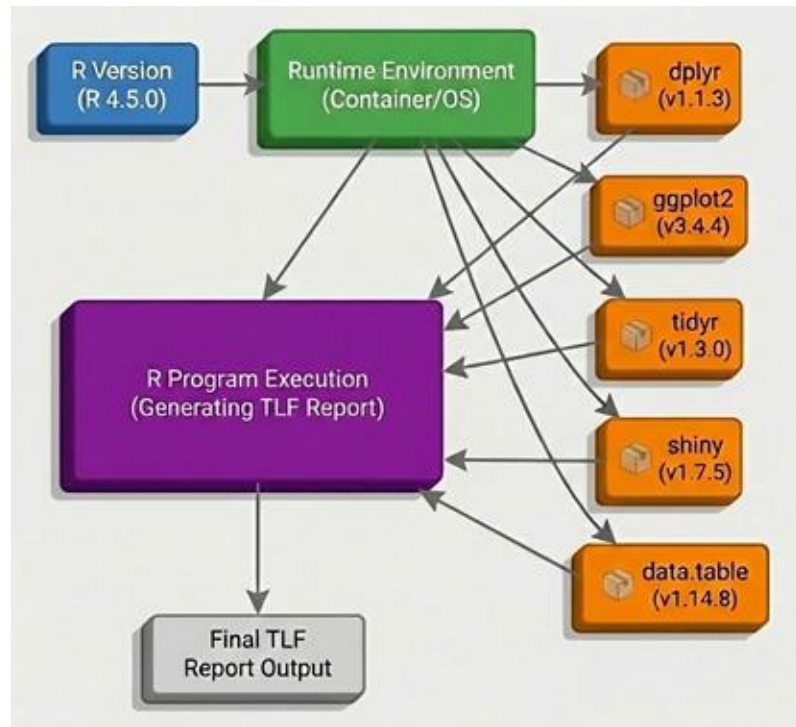
Automated Capture:

- Every code execution with timestamp and user attribution
- Data transformations with complete provenance
- Environment specifications and computational context
- Result generation linked to specific code versions

Structured Traceability:

- **Data Provenance:** Source data → transformations → analysis datasets
- **Code Versioning:** Git integration with review records
- **Specification Linkage:** SAP → code → outputs
- **Validation Evidence:** Automated collection of test results and approvals

Benefit: Complete audit trails without manual documentation effort

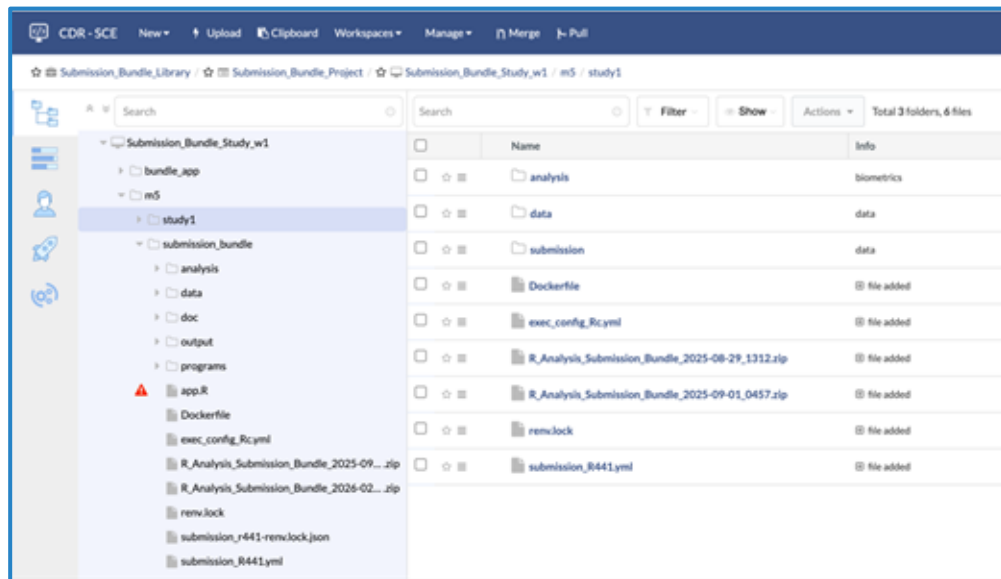


Component 3: eCTD-Ready Structure

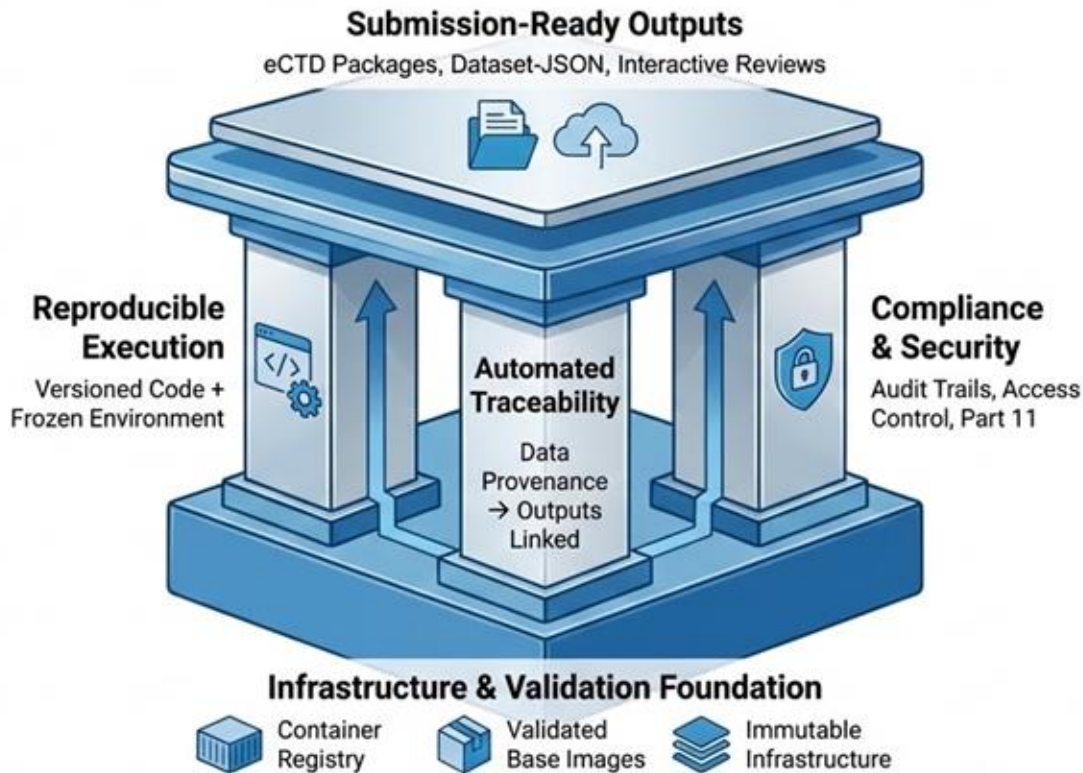
eCTD Templates:

- Pre-configured folder hierarchy (Module 5)
- Standardized naming conventions (automated, error-free)
- Integrated documentation in eCTD-compatible formats
- Submission metadata for automated processing

Impact: Submission preparation transforms from labor-intensive final phase to automated output



GxP Compliant Scalable Submission Container



Real World Impact

Development Efficiency:



- 30-50% reduction in statistical programming timelines
- Environment provisioning: Days → Minutes
- Cross-team coordination: Weeks → Days

Quality & Compliance:



- Up to 85% reduction in data integrity findings during inspections
- Zero environment-related discrepancies in validation
- Complete audit trails without manual documentation effort

Reviewer Experience:



- Environment setup: Hours → Zero (container execution)
- Reproducibility: 100% (identical results guaranteed)
- Archival compliance: 20+ years maintained effortlessly



The Question Is No Longer 'Whether' But 'How Fast'

From using R → to mastering R delivery

The Transformation:

FROM



Documentation exercises



Manual compliance efforts



Siloed solutions



Project-specific validation



TO



Demonstrations of computational excellence



Natural properties of well-designed systems



Organizational infrastructure



Reusable validated frameworks



The Future:

Organizations implementing **modern statistical computing environments** position themselves to **deliver life-saving therapies more efficiently** while ensuring **analytical insights remain credible, reproducible, and accessible decades into the future.**



Key Takeaways



Regulatory acceptance is established

FDA pilots validate R across the submission lifecycle



Challenges remain substantial

Dependency complexity, validation burden, cross-platform issues



Containerization alone ≠ Complete solution

Need integrated modern SCE with audit, traceability, eCTD-ready outputs

Benefits are quantified



30-50% timeline reduction



85% fewer data integrity findings



Zero environment discrepancies

Strategic window is open



Proactive adoption builds competitive advantage; reactive approach means playing catch-up



The Path Forward: Adopt containerized workflows as organizational infrastructure, not project-specific solutions



Questions?

Priyanka Sawant

Sycamore Informatics

Email: Priyanka.sawant@sycamoreinformatics.com

LinkedIn: [linkedin.com/in/priyankasawant](https://www.linkedin.com/in/priyankasawant)

Scan & Connect !



Thank you !