Scientific AI and the Future of OME-NGFF

Intelligent Bioimage Analysis Workflows

Matt McCormick, PhD

ƒideus labs

OME-NGFF Workshop 2025

🌐 HTML slides | 📄 PDF slides | 📂 GitHub repository

Real-life Bioimaging Lab Challenges

Scenario: A postdoc just collected 50 microscopy images

- Need to quantify cell populations across treatment conditions
- Segment nuclei, measure fluorescence intensity, compare to controls
- Generate statistical analysis and publication-quality figures

Real-life Bioimaging Lab Challenges

Without AI:

```
Week 1: Write analysis scripts, debug edge cases
Week 2: Batch process, fix failures, re-run
Week 3: Statistical analysis, figure generation
⏱️ 3+ weeks of precious postdoc time
```

With AI + OME-NGFF:

```
"Analyze these 50 images: segment DAPI-stained nuclei,
quantify GFP intensity per cell, compare treated vs control,
and generate summary statistics with publication figures"
✅ Results in hours, fully reproducible
```

Today's Vision

25 minutes of inspiration

1. **Scientific AI and Your Data** (5 min)
    - Why OME-NGFF for agentic AI
2. **Introduction to the ngff-zarr MCP Server** (10 min)
    - AI-powered bioimage conversion and batch processing
3. **Next Steps You Can Take** (10 min)
    - Improve your scientific throughput, reproducibility, and impact

Part 1: Scientific AI and Your Data

Why OME-NGFF for Agentic AI

The AI Revolution in Science

Large Language Models (LLMs) are transforming scientific computing:

- 💬 Natural language interfaces to complex tools
- 🧠 Reasoning over scientific data and workflows
- 🤖 **Agentic AI**: Models that can plan, decide, and execute

The Challenge: *How do we give AI meaningful access to your scientific imaging data?*

What is Agentic AI?

Traditional AI: **Question → Answer**
Agentic AI: **Goal → Outcome**

- 🧠 **Context** - Understanding your specific problem
- 🛠️ **Tools** - Access to your scientific software
- 🔄 **Reasoning** - Planning multi-step workflows
- ✅ **Execution** - Actually processing your data

Result: AI that understands your science and automates complex analyses

Why OME-NGFF is Perfect for Agentic AI

☁️ Cloud-Ready Architecture

- **Chunked storage** - access specific regions without downloading entire datasets
- **Hierarchical structure** - AI can reason about data at multiple scales
- **Standard web interfaces** - works everywhere your AI runs

Why OME-NGFF is Perfect for Agentic AI

📚 Rich Metadata

- **Spatial information** preserved and queryable
- **Standards-compliant** - AI knows what it's working with
- **Ecosystem of tools** - interoperability enables complex workflows

Why OME-NGFF is Perfect for Agentic AI

🌐 Truly Open

- **No vendor lock-in** - your data remains yours
- **Community-driven** - growing ecosystem of tools and support
- **Reproducible science** - data format ensures long-term accessibility

The Power of the Combination

```
Your Scientific Data (OME-NGFF)
          ↓
     AI Agent
    (with tools & context)
          ↓
Automated, Reproducible Workflows
```

What becomes possible:

- 🧪 Analyze petabyte-scale datasets in minutes
- 📊 Reproducible computational experiments
- 🎯 Intelligent parameter optimization
- 🚀 Batch processing without manual intervention

Part 2: Introduction to the ngff-zarr MCP Server

AI-Powered Bioimage Conversion and Batch Processing

What is the Model Context Protocol (MCP)?

Universal bridge between AI assistants and your scientific tools

```
Your AI Assistant (Claude, Qodo, etc.)
          ↕️
     MCP Client
          ↕️
  MCP Server (ngff-zarr)
          ↕️
Your Scientific Data & Tools
```

Why MCP?

- ✅ Standard protocol - works with multiple AI platforms
- 🔧 Easy integration - no custom coding required
- 🔄 Bidirectional communication - rich interaction model

The ngff-zarr MCP Server

What is it?

- **ngff-zarr** is an open-source toolkit for working with OME-NGFF data
- The **ngff-zarr MCP Server** exposes this toolkit to agentic AI systems
- **AI-ready interface** - natural language commands to scientific operations

The ngff-zarr MCP Server

What can it do?

- 🔄 **Convert** scientific images (NRRD, TIFF, Nifti, etc.) to OME-Zarr
- ✅ **Validate** OME-Zarr datasets for compliance and interoperability
- 🛠️ **Optimize** compression and chunking for your specific access patterns
- 📝 **Inspect** multiscale pyramids, metadata, and data structure
- 📦 **Batch process** large collections of images with reproducible settings

Real-World Example: Quantitative Analysis Pipeline

Remember our postdoc with 50 images?

Without agentic AI:

```
Traditional manual workflow:
1. Convert images to working format (troubleshoot compatibility)
2. Write segmentation script (tune parameters per image)
3. Extract measurements (debug data export)
4. Statistical analysis (wrangle data formats)
5. Generate figures (iterate on visualization)
6. Document everything (if you remember what you did)
⏱️ 2-3 weeks of iteration and debugging
```

With agentic AI + ngff-zarr MCP:

```
Natural language request:
"Analyze these 50 microscopy images: segment DAPI-stained nuclei,
quantify GFP intensity per cell, compare treated vs control groups,
and generate summary statistics with publication figures"

AI agent:
1. Converts to OME-NGFF (optimal format selection)
2. Performs segmentation (auto-tunes parameters)
3. Extracts measurements (structured metadata)
4. Runs statistical comparisons (appropriate tests)
5. Generates figures (publication-ready)
6. Documents entire workflow (fully reproducible)
✅ Results in hours, not weeks - fully automated and documented
```

🛠️ In Practice: AI-Powered Conversion

💬 Convert a bioimage with AI assistance

Put the agent to work!

In chat:

```
Convert the vs_male.nrrd file to OME-Zarr format and
find the optimal compression codec for this type of data.
```

✨ Watch the AI agent:

1. 🔍 Analyze the input file
2. 🎯 Select appropriate parameters
3. ⚙️ Execute the conversion
4. 📊 Report optimization results

💬 Examine OME-Zarr contents

Ask the AI to explore:

```
Examine the contents of carp.ome.zarr and tell me
about its structure, dimensions, and metadata
```

✨ The AI agent will:

- 🔍 Inspect multiscale levels
- 📏 Report spatial metadata
- 🧩 Analyze chunk structure
- ✨ Suggest next steps

💬 Generate batch script

Scale up with AI automation:

```
I have a folder of 50 similar NRRD files.
Generate a Python script to batch convert them all
to OME-Zarr with the same optimal settings
```

✨ The AI agent creates:

- 🐍 Complete Python script
- ⚠️ Error handling
- 📈 Progress reporting
- 🎯 Optimized parameters from previous analysis

Key Capabilities in Action

🤖 Intelligent Conversion

- AI analyzes your image data automatically
- Selects appropriate compression codecs
- Generates multiscale pyramids without artifacts
- Optimizes for your hardware and access patterns

Key Capabilities in Action

📊 Analysis and Reporting

- Inspect complex multiscale structures
- Generate processing scripts
- Plan batch operations with resource awareness
- Validate results automatically

The Future of Scientific Image Analysis

Today: AI agents help with format conversion and optimization

Tomorrow: AI agents assist entire scientific workflows

- 🧪 Multi-step analysis pipelines
- 🎯 Automated optimization and tuning
- 📈 Scalable processing of massive datasets
- 🔬 Knowledge discovery through AI reasoning

Your role: Guide the AI with scientific questions, not technical commands

Part 3: Next Steps You Can Take

Improve Your Scientific Throughput, Reproducibility, and Impact

Your Path Forward

🎯 Immediate Actions

1. **Explore** the ngff-zarr documentation and examples
2. **Try** converting a sample image to OME-NGFF
3. **Evaluate** whether OME-NGFF fits your workflow

Your Path Forward

🚀 Short-term Goals

1. **Pilot** OME-NGFF adoption in your lab
2. **Integrate** with your existing analysis pipelines
3. **Measure** improvements in efficiency and reproducibility

Your Path Forward

📈 Long-term Vision

1. **Scale** your analysis to larger datasets
2. **Leverage** AI and automation for complex workflows
3. **Contribute** back to the community

Tools for Every Scientist

🐍 ngff-zarr Python Library

For computational researchers:

- Direct programmatic access to conversion and optimization
- Integration with Jupyter notebooks and workflows
- Scientific Python ecosystem compatibility (NumPy, Dask, Xarray)
- Custom analysis pipelines built on OME-NGFF

📝 **Note:** *OME-Zarr in cloud environments* by Eric Perlman this afternoon will cover cloud-based imaging workflows in detail!

Tools for Every Scientist

ngff-zarr TypeScript Library

For web and visualization developers:

- Browser-based OME-NGFF exploration
- Web applications for image analysis
- Cloud-native deployment options

ƒideus labs

Who we are:

- 🧬 Biomedical imaging specialists
- 🌟 OME-NGFF ecosystem contributors
- 🤝 Open science advocates

What we offer:

- 📚 Training and consultation services
- 🔗 Integration support for your existing workflows

Connect: 📧 Email: info@fideus.io 🌐 Website: https://fideus.io

Key Takeaways

✅ **OME-NGFF** is cloud-ready, open, and built for agentic AI

✅ **MCP Servers** bridge your AI assistants and scientific tools

✅ **ngff-zarr** makes powerful bioimage analysis accessible to everyone

✅ **Your impact multiplies** when you combine AI, open data, and reproducible science

Questions & Discussion

💡 What we covered:

- Why OME-NGFF is essential for scientific AI
- How ngff-zarr MCP enables intelligent automation
- Concrete steps to improve your research

👥 Let's discuss:

- Your specific imaging challenges
- Integration questions for your workflow
- How OME-NGFF can accelerate your research

Getting Started with OME-NGFF

📚 Resources and Documentation

ngff-zarr Project:

- 📖 **Documentation:**: https://ngff-zarr.readthedocs.io
  - Comprehensive guides and API reference
- 📂 **GitHub:** https://github.com/fideus-labs/ngff-zarr

OME-NGFF Specification:

- 📋 **NGFF Standard:** https://ngff.openmicroscopy.org
- 📚 **Community:** Open Microscopy Environment (OME)