The G-Node Infrastructure Services:

Safe, efficient and seamless data management for neuroscience

gin
gin.g-node.org

Christian Garbers, Achilleas Koutsou, Michael Sonntag, Thomas Wachtler

German Neuroinformatics Node, Department of Biology II, Ludwig-Maximilians-Universität München, Martinsried, Germany

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Federal Ministry of Education and Research

G-Node

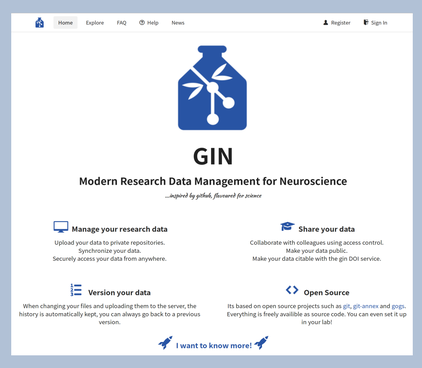
g-node.org
dev@g-node.org
g-node.github.io

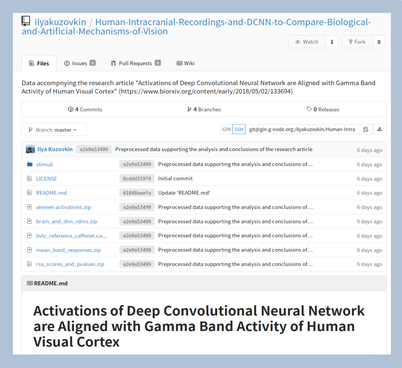
Introduction

Maintaining reproducible data workflows while keeping data in sync, backed up, and easily accessible from within and outside the lab is a key challenge in research. To minimize the time and effort required for these tasks, we present the GIN services [1](RRID:SCR_015864), a suite of tools designed for comprehensive, reproducible and versioned management of scientific data.

GIN extends and is fully compatible with established versioning tools (git [2] and git-annex [3]) and offers seamless data access with methods provided by those tools or with common web protocols (HTML, WebDAV). Furthermore, it integrates with tools for data and metadata management [4,5,6] and provides indexing and easy web-based editing of metadata. GIN combines the power of a repository management service (inspired by GitHub [7]) with data storage and offers easy to use interfaces for data management through a web browser, from the desktop file browser, from the command line, or in analysis scripts. GIN makes it straightforward to share data within a lab or with off-site collaborators and to work on it together. Finally, with GIN's metadata indexing and DOI services any dataset can easily be made findable or citable for publication.

In summary, GIN offers a convenient and powerful solution for the demands of reliable and efficient data management in the lab, combined with seamless data sharing with collaborators and the general scientific community - open, FAIR, and straightforward.

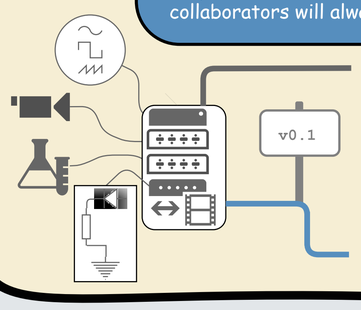




References: [1]GIN (https://gin.g-node.org/) [2]git (https://git-scm.com/) [3]git-annex (https://git-annex.branchable.com/) [4]odML (http://www.g-node.org/mix) [5]NIX (http://www.g-node.org/mix) [6]BIDS (http://bids.neuroimaging.io/)

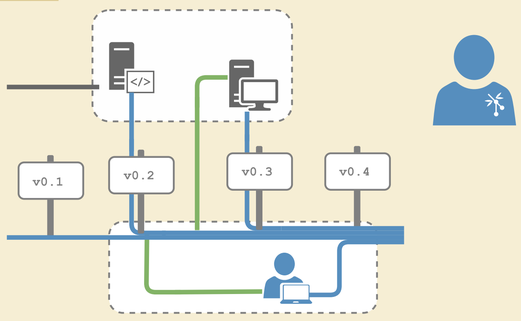
Argh! I sent Mark the fixed files ages ago, and you know what? He says his analysis failed because of me, but he never downloaded the new files!

Why don't you try the **GIN** toolset from **G-Node**? It's like a secure Dropbox designed specifically for scientific work, with a lot of extra features that will save you time and effort. Files are kept track of and in sync. With GIN, sharing data is easy, and your collaborators will always have the latest version.



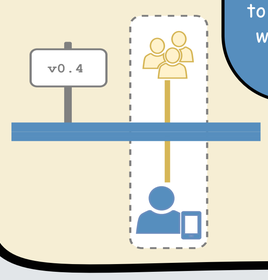
Oops, dang it. I overwrote my data with cute pics of my lab mice. Again.

Sounds like you need **versioned data**! With GIN, you can inspect any modifications and return to any previous state of a file.



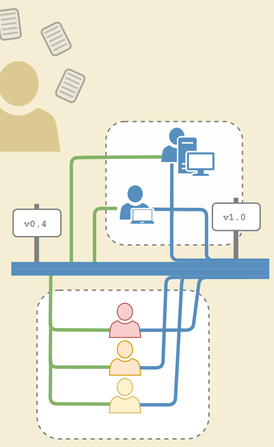
I'm taking care of my sick kids at **home**, and next week I'm flying to a **conference** in Japan. I'm never going to get any work done!

GIN allows you to access your data **remotely** from anywhere, using **any device**. With the GIN Web Server, you can **upload, download, edit, share, raise issues, make notes, search data, and much more** – no matter where you happen to be. You can even use it to keep track of what your students accomplish while you are away!



Umm... that's a lot of data. Guess I will have to buy a new hard drive. I really only need to work on that one file though.

With GIN, you and your collaborators can download only what you each require. Save time, space, and keep an overview by downloading only the **metadata**, such as name and history, of your **large files**. Then request their contents when you need it.



Why can't my collaborator figure out how to use my server with a cmd client? It's so easy!

Not everyone is that comfortable with computers. That is why GIN allows data to be managed in a variety of ways. You can use a **cmd client** and **scripts**, while your collaborator can interact with GIN using our intuitive **web interface**.

Nice! But can I integrate GIN into my processing **pipeline**?

Sure! GIN can be configured to send out a notification (webhook) to another server whenever there is an update. This can be used to kickstart your automatic processing steps.

Perfect. Can I **host** these tools **myself**?

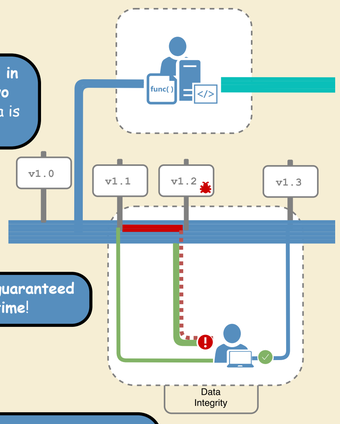
Yes, we even encourage it. Our **experts** will assist you with the installation if the need arises.

What if my data gets corrupted?

Your data is **cryptographically secured**. Files are routinely compared to their source, and you are **informed** if anything goes wrong.

Where is GIN hosted? What if your server breaks down?

GIN is hosted at the **LMU in Munich**, and there are **two backups** in place. Your data is **safe**.



If you lose funding, will all my data be erased?

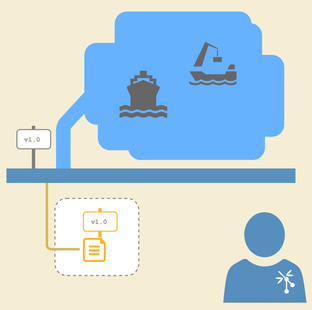
Repositories with a DOI are **guaranteed 10 years of data lifetime**!

What if aliens invade?

Well, if they are researchers, they are free to use GIN too.

I want to **publish** this awesome paper, but the journal is making it so hard! I have to get a **DOI** for my data? How do I do that?

Obtaining a DOI is easy with GIN. Prepare your repository following a simple guide, then hit the **DOI button**. That's it! Our **data experts** are there to assist you if need be. Once your data is published, researchers can ask questions and add to your dataset or create their own copy to work with. Your data can lead to new scientific breakthroughs, and you will get the credit!



Okay, I want GIN! Take all my money!

Woah, save that for a vacation! The **GIN tools**, access to the **web interface**, **storage space**, **DOI registration**, **detailed guides**, and even **expert assistance** are **free of charge**!

So many nights and weekends spent in the lab, but finally it's published and PhD is done! Whats next?

Here perhaps some gin for a change?

... Yeah. Thanks.