## Git & GitHub Venexia Walker, Winfred Gatua, Yi Liu

#### Agenda

Part 1: The Basics [Venexia Walker]

Part 2: GitHub Functionality [Winfred Gatua]

Part 3: IEU GitHub organisation [Yi Liu]

# Part 1: The Basics

Venexia Walker

#### What?

Git is a content tracking software.

GitHub is one online provider of Git.

Alternatives to GitHub include Bitbucket and GitLab.

You can sign up to GitHub for free at https://github.com/.

#### What?



https://support.nesi.org.nz/hc/en-gb/articles/360001508515-Git-Reference-Sheet

## Why?

Reproducible research.

Version control.

Code backup.

Work on code with collaborators.

Personal online code record.

#### How?

Online using https://github.com/	GitHub Desktop
Github command line tool (CLI)	Git integration into RStudio or other software

#### Online using https://github.com/



#### GitHub desktop



#### Git integration into RStudio or other software



#### Github command line tool (CLI)

#### •••

\$ gh pr status

Current branch There is no pull request associated with [develop]

#### Created by you

#1011 Update readme [readme-fix]
- Checks pending - Review required

## Summary

Something is better than nothing - if you use Git in anyway, that is great! 🎉

You will get into a horrible Git mess at some point. Don't panic. Backup everything somewhere else before you do anything radical like force a git push or git pull.

There is lots of help out there - the documentation is a great place to start.

# Part 2: Github Functionality

Winfred Gatua

## README

This is a useful place to highlight relevant project information, such as:

- How to use the repository
- How to access the relevant data
- Supplementary material for a project
- Project abstract and authors

A README.md file in the repository will automatically be rendered and the Markdown format means markdown links, tables, etc. can be included.

Example: <u>https://github.com/venexia/repurposing-antihypertensives-dementia</u>

### .gitignore

This hidden file tells Git what it should not track.

Try to keep your sensitive data in a specific directory and add the whole directory to the .gitignore. Alternatively, specific files can be added to the .gitignore.

Maintain directory structure by ignoring the directory, except a README.md:

raw/\*
!raw/README.md

## Git history

Everytime you commit something to Git, it is recorded in the Git history.

When you make a repository public, the history of the repository when it was private is shared.

If sensitive data has been accidentally included at any point, you can use tools, such as <u>BFG Repo-Cleaner</u>, to remove that part of the history. Worst case scenario: you can delete your repository but this will mean losing all your other history.

#### Relative file paths

Avoid including absolute file paths in your code for security.

The easiest way to do this is to use relative paths - for example, "raw/data.txt" rather than "E:/venexia/phd/project/raw/data.txt". You would then just need to set your working directory as "E:/venexia/phd/project/" before running anything.

If you are using RStudio, use the project functionality. An RStudio project will open where you left off in your working directory meaning you only ever need relative file paths.

#### Templates

You can create project templates that can be used when you setup new repositories or you use other people's templates,

To create a template, make a new repository with your preferred setup, then online go to 'Settings > Template repository'.

Example: <u>https://github.com/venexia/template</u>

#### Branches and pull requests



#### Make a branch

Online	<u>GitHub Desktop</u>
Click branch selector on repository main page Type branch name Click 'Create branch'	Click 'Current Branch > New Branch' Type branch name Click 'Create Branch'
GitHub CLI git checkout -b branch_name first time: git pushset-upstream origin branch_name once setup: git push -u origin branch_name Other useful commands: git branch -a	<b>Git integration</b> In RStudio Git tab, click Git logo Type branch name Click 'Create'

#### Make a pull request

Pull requests are used to merge a branch.

The request is made online by clicking 'Pull requests > New pull request' and selecting the relevant branch.

Pull requests allow you to compare what has changed since the branch was made - this is useful for you and and can be used by teams for code review.

The request will highlight merge conflicts.

## Basic formatting syntax

- 1. <u>https://commonmark.org/help/</u>
- 2. Markdown cheatsheet

Туре	Or	to Get
*Italic*	_Italic_	Italic
**Bold**	Bold	Bold
# Heading 1	Heading 1 =======	Heading 1
## Heading 2	Heading 2	Heading 2
[Link](http://a.com)	[Link][1] : [1]: http://b.org	Link
![Image](http://url/a.png)	![Image][1] : [1]: http://url/b.jpg	M↓
> Blockquote		Blockquote
* List * List * List	- List - List - List	<ul><li>List</li><li>List</li></ul>



# Part 3: IEU GitHub organisation

## https://github.com/MRCIEU



MRC Integrative Epidemiology Unit Packages & Teams 21 8 People 170 New repository O Find a repository... Type -Language -Sort mrcieu.github.io Private MRC IEU GitHub pages ● Jupyter Notebook ♀5 ☆1 ⊙0 \$\$0 Updated 1 minute ago upptime Public mm Use upptime to monitor web platform status uptime-monitor status-page upptime ● JSON 杯 MIT 学0 公0 ①0 第20 Updated 19 hours ago temmpo Public Text mining for mechanism prioritisation ~ AA mesh-terms ● Python \$ GPL-3.0 \$1 \$5 ① 0 \$10 Updated 3 days ago gwasglue2 Public AAA Connecting GWAS summary data to analytical tools ● R ♀ 1 ☆ 1 ① 13 1 0 Updated 3 days ago bp-drug-mr Public A . A Multi-ancestry MR of blood pressure drug targets ● HTML ♀1 ☆1 ①1 ♫0 Updated 3 days ago epigenetics-journal-club Public Repository for the IEU epigenetics journal club papers, presentations and scripts 

#### Become a member (if you aren't already)

Some of the community repos require you to be a member of the MRCIEU\_internal team to access



21 teams in the MRCIEU organization		Visibility -	Members -
Developers	<b>.</b> 2 <b>0 9  .</b>	20 members	0 teams
epigraphdb EpiGraphDB a database and platform for health data science	(a) (b) (c) (c) (c) (c) (c) (c) (c) (c) (c) (c	9 members	0 teams
External analyists	٩	1 member	0 teams
FATHMM	1	2 members	0 teams
Genetic epi tutors Futors for genetic epi practicals	g 🔟 🎗 🎂 🍣 🗑 🤁 🗉	40 members	0 teams
godmc-analysts SoDMC analysis	<b>8</b> 8040 <b>.</b>	55 members	0 teams
GoDMC2 SoDMC 2 analysis		16 members	0 teams
was_vcf_format eam for developing standard storage format or GWAS summary stats	<b>\$</b> @@# <b>*\$</b>	6 members	0 teams
eup4 EU programme 4	() () () () () () () () () () () () () (	23 members	0 teams
embers from IT services	<b>@</b> 🖕 🤁 🛞	4 members	0 teams
IELODI	🛞 🕅 🚼	3 members	0 teams
IRBase	🌔 💮 🎯 🈤 😱 🆃 🚷	20 members	0 teams
IRCIEU_internal nternal members of MRC IEU	<b>&amp; () &amp; () () + ()</b> -	109 members	0 teams
/VP analysts		5 members	0 teams
PracleCloud eam for Oracle Cloud development work	@ <b>t</b>	2 members	0 teams
henoSpD method to estimate phenotypic correlation nd multiple testing correction f	াল 😎	2 members	0 teams
Patture Coordination			

Welcome to MRC Integrative Epidemiology Unit GitHub pages.



#### **IEU GitHub Pages**

- Showcase the software, coding, and data science side of IEU research
- Showcase and promote your software!
  - R packages: very many!
  - Processing and application pipelines
    - MELODI-presto
    - MendelVar
  - Applications (web / command line / GUI)
    - MR dictionary
    - MELODI-presto
    - EpiGraphDB (web, API)
    - OpenGWAS (web, API)
  - Big software projects
    - OpenGWAS
    - EpiGraphDB

underlying causes of chronic disease Researchers at the MRC IEU have created a range of openly available software tools based on research methods developed within the unit. momentary assessment Mende ian randomization Heteroscedasticity EpiGraphDBDNA methylation Multilevel models systematic rating mining Regression OpenGWAS Deep learning Visualization Text embedding Evidence integration EWAS Catalog Machine learning Text embeddings Missing outcomes WR-phewAs RCT Data mining Phenome Scan GWAS Genetics EWAS Natural language processing Data integration vQTL OOV proteomics Missing-not-at-random

The MRC Integrative Epidemiology Unit at the University of Bristol (IEU) conducts some of the UK's most advanced population health science research. It uses genetics, population date and experimental interventions to look for the



#### How to contribute

- 1. By making a pull request: add an entry to the software.yml
- 2. Make suggestions on the repo's issues

See further details on the README

٢	master - mrcieu.github.io / data /		Go to file	Add file -	
٨	YiLiu6240 Add profile IEU GitHub wiki		• b0907ca 5 h	ours ago 🔞 F	listory
	profile-images	Add profile IEU GitHub wiki		5 hou	irs ago
	survey	Update		2 yea	ars ago
۵	.gitignore	Update		2 yea	ars ago
D	parent-projects.yml	add ewas catalog		2 yea	ars ago
Ľ	software.yml	Add profile IEU GitHub wiki		5 hou	ırs ago

> B	IN +282 KB data/profile-images/ieu-github-wiki.png 🖵		(*)
~	🕆 14 🖬 🖬 data/software.yml 🖵		
. <u>.</u> .	@@ -962,6 +962,20 @@ ieu-github-pages:		
962 963 964	- Python - Jinja2 - Website	962 963 964 965	<pre>- Python - Jinja2 - Website + ieu-github-wiki:</pre>
		966 967 968 969 970 971 972 973 974 975 976 977 978	<pre>+ label: IEU GitHub Wiki + description: Wiki and knowledge base for IEU members + maintainer_name: Yi Liu + contact: yi6240.liu@bristol.ac.uk + software_url: https://mrcieu.github.io/wiki + software_type: Application; + point_development: Active maintenance + code_public: false + collaborators: IEU software coordination team + image: ieu-github-wiki.png + tags: + - GitHub Actions + - Docker</pre>
965 966 967	<pre>xtmixediou-stata-package: label: xtmixediou Stata package description: Stata command to estimate the linear mixed effects integrated Ornstein-Uhlenbeck model</pre>	979 980 981	<pre>xtmixediou-stata-package: label: xtmixediou Stata package description: Stata command to estimate the linear mixed effects integrated Ornstein-Uhlenbeck model</pre>

#### Submit / update your software!

#### See results

/RC I	nteg	rative	Epide	miolo	gy Unit					+	lome	About	MR	C IEU s	oftwa	re brov	wser	Detail	s Wi	d		
					N	<b>IR(</b> wiki	CIE	U s	of	two	ire	br	ov	/se	r							
all	API	Active	develo	pment	Active	e maint	enance	Analy	tical c	and/or d	ata pro	cessin	g pipe	line	Andro	id Stu	dio	Applic	ation			
Artist	ic Lice	ense 2.0	Bas	h Bo	otstrap-	vue	C++ [	DNA met	thylat	ion D	ata inte	egratio	on I	Data m	nining	Dat	abase	De	ep lear	ning		
Dictio	nary	Djang	go D	ocker	EWAS	EW	'AS Cata	log E	arly p	rototype	Ela	sticsed	arch	EpiG	aphDB	B E	vidence	e integ	ration			
FastA	PI	Flask	GPL-3	GPI	Lv3 G	WAS	Gene E	xpressio	on Om	nnibus	Genet	tics	GitHu	ıb Acti	Actions Graph dat			ase	se			
Heter	osced	lasticity	Hug	o Illu	imina Be	adchip	Inter	action ef	fect	Java	Jinja2	2 Li	teratu	ure mining MIT MF		MR-B	ase	e MR-pheWA				
Mach	ine leo	arning	Mend	elian ra	ndomiza	tion	Metabo	lomics	Met	hodolog	y Mi	ssing	outcor	mes	Missir	ig-not	-at-rar	ndom				
Multil	Aultilevel models Nature		Natural langua		uage processing		y Neo	4j No	None Ope		GWAS Phenome Sco		Scan Phenotypes			s Py	Sam	Pyth	on			
RCT	Reg	gression	Shir	ny Sr	nakemak	e S	oftware p	oackage	St	ata S	ystemo	itic rev	/iew	Terro	form	Text	embe	dding				
Text e	mbed	ddings	Text r	nining	The E\	NAS C	atalog	Variabl	es	Varianc	e Vi	sualize	ation	Vue	Vue	e.js	WDL	We	b app			
Web :	servic	e We	bsite	XML	httr	metal	polomics	micro	biom	ics m	omento	ary as:	sessm	ent	proteo	mics	smo	king				
systei	natic	review	VQTL																			
						_	1															
NACTOR IN SUCTOR In Comparison In Comparison	Links Aspert Aspert Univer Since Aspert Aspe	s and resources was control to a way of the staty life Control for a second of the control of the control of the second of the control of the second of the control of the control of the second of th	u vago una case factores naciones tes tes tas tas espect	ngo fa guidens er tuk Nasi ber degy - dic d	C have been a series of the series of the series of the series of the series of the series of the series of the series of the series of the series of the series of the se	dange B																

#### MRC Integrative Epidemiology Unit

IEU software projects —Annotated Semantic Queries

-API for MRC-IEU OpenGWAS -BESIDEMR R package —bpbounds R package -bpbounds Stata package -cofunnel Stata package -DataWatch \_\_dmrff -dnamalci -EpiGraphDB -EpiViz -ewaff -FATHMM —geograbi —GLU R package -GWAS of trait variance -gwas2vcf -IEU GitHub Pages -IEU GitHub Wiki —iPVs —ivonesamplemr Stata packag -medrxivr

—meffil —MELODI —MELODI-Presto

#### IEU GitHub Wiki

Wiki and knowledge base for IEU members

MRC IEU Wiki		Q Search
MRC IEU Wiki	Links and resources	Table of contents
Home		University links
Data management		MRC IEU
Links and resources	Suggestions and contributions are very welcome (see the homepage for guidance on making	HPC help
Scripting	changes).	External resources
Software development		The University of Edinburgh Data Science Training in
Best practices in coding	University links	Health & Bioscience
Code Testing	MRC IEU	
Creating bgen files from plink	The IEU intranet (internal) - extensive resources for training in epidemiology - MRC IEU	
R on a Linux server using	software browser	
Conda	MRC IEU homepage and coding	
Mounting a network drive on Linux	Code clinic Teams link - Go here to get lots of peer support and for recorded videos from     and evening	
R tidyverse examples	past sessions	
Using Biomake		
Using Docker	HPC help	
Using Emacs	ACRC page	
	ACRC HPC docs (internal)	
	External resources	
	The University of Edinburgh - Data Science Training in Health & Bioscience	
	https://edcarp.github.io/Ed-DaSH/	
	Ed-DaSH is a Data Science training programme for Health and Biosciences. The team has	
	developed workshops using The Carpentries platform on the following topics. See	
	workshops for dates and registration details. All workshops will be delivered remotely.	
	List of upcoming workshop sessions: https://edcarp.github.io/Ed-DaSH/vorkshops.html	
laintainer(s) <sup>,</sup> Yi Liu		
ontact: yi6240.liu[d	at]bristol.ac.uk	
ollaborator(s): IEU	software coordination team	
oftware URL: http:	s://mrcieu.github.io/wiki	

#### IEU wiki

- Curated knowledge base and wiki for IEU (GitHub) members
- Wiki site
- source wiki

MRC IEU WIKI		Q Search
KRC IEU WINI	Hame	Table of contents
tone	Home	Contri ellinia
Duta management		Making changes
links and resources	Welcome to the MRCHEU wiki. It has information about the code clinic and also is a place to	
kripting	post and share code snippets that others might find useful.	
icftware development		
kriicks v	0.000	
Best practices in coding	Code cinic	
styles	EU members - make sure you're on the Code clinic Teams channel where you get -	
Code lesong		
Creating Open roles maniparea	<ul> <li>Peer support and discussion for coding and data related issues</li> </ul>	
R on a Linux server using	<ul> <li>Bi-weakly meetings (Tuesdays 12pm)</li> </ul>	
Conda	· Recorded videos on different aspects of scripting, writing software and managing data	
Mounting a network drive on- Linux		
R tidyverse examples	Makino channes	
Using Biomake	Noning Changes	
Using Docker	If you are a member of the MRCIEU GitHub organization, you can -	
Using Ewacs	<ul> <li>Make edits and contributions to the wiki directly at the source wiki</li> </ul>	
	<ul> <li>Make discussions at the discussions page</li> </ul>	
	<ul> <li>Contribute to the doublesment of the ultimeter</li> </ul>	

Home	Edit New pag
Welcome to the MRCIEU wiki. You are looking at the internal source wiki (on the public facing site this page will be replaced by "SITE README"). This wiki is aimed at building a curated	• Pages 🚯
knowledge base for IEU members to share knowledge.	tind a page .
As a member of the MRCIEU GRHub organization, you can	+ Home
· Make edits and additions to the wiki directly at the source wiki	Making changes
<ul> <li>Make discussions at the discussions page</li> </ul>	Differences between the source wiki and the public wiki website
Contribute to the development of the wiki repo	Technicals
Making changes	Best practices in coding styles
NOTE: The following covers steps to make changes to the content pages of the will she. If you	Code Testing
are hoping to make changes to the will site itself (i.e. to the the repol's code), you should consult the documentation of the wiki repo.	Creating been files from plink
Creating a new page: On this will page click 'New page'	<ul> <li>Data management</li> </ul>
<ul> <li>Give the page a title, at the top of the body of the page put something like sutter: Your Name. If you want to claim ownership), and then just write the snippet and</li> </ul>	* Installing a specific version of R on .
relevant into below in Markdown format.	<ul> <li>Links and resources</li> </ul>
<ul> <li>To edit an existing page: simply go to a page and click on edit.</li> <li>Once you gave the gave the Cittle &amp; Articles will adversifially under the</li> </ul>	* Mounting a network drive on Linux
https://mcieu.github.io/wki/ website. There's nothing else you need to do.	* R tidyverse examples
NOTE: Aside from some exceptions (including this one), EVERY visible page on this source	<ul> <li>Scripting</li> </ul>
which will be visible on the public facing intercompression action provides on the solution information that might be sensitive or confidential (e.g. server address, patient information	* SITE_READWE
etz).	Software development
Differences between the source wiki and the	* Using Biomake
public wiki website	Using Docker
(V) (L) (2) (L) (L) (L) (L) (L) (L) (L) (L) (L) (L	Using Emacs
<ul> <li>Terminology</li> </ul>	

\* The repo of the wiki site is https://github.com/MRCIEU/wiki

#### How to contribute



#### Wiki for MRC IEU users

This is the code repository of MRCIEU's GitHub wiki site, where the wiki of this repo is served as the page content of the rendered site.

- If you want to make changes to the wiki's content, please check the documentation here.
- . If you want to make changes to the wiki site itself, please continue reading the docs.

#### How to contribute

#### A MRCIEU / wiki Private ⊙ Unwatch 59 - ♀ Fork 1 -Star 0 -<> Code 💿 Issues 🏌 Pull requests 🖓 Discussions 🕑 Actions 🖽 Wiki 🕕 Security 🗠 Insights 🚳 Settings Home Edit New page YiLiu6240 edited this page on Jan 16 · 12 revisions Welcome to the MRCIEU wiki. You are looking at the internal source wiki (on the public facing site this page will be replaced - Pages 15 by "SITE\_README"). This wiki is aimed at building a curated knowledge base for IEU members to share knowledge. Find a page ... As a member of the MRCIEU GitHub organization, you can Home · Make edits and additions to the wiki directly at the source wiki Making changes Make discussions at the discussions page Differences between the source · Contribute to the development of the wiki repo wiki and the public wiki website Technicals Making changes Best practices in coding styles NOTE: The following covers steps to make changes to the content pages of the wiki site. If you are hoping to make changes Code Testing to the wiki site itself (i.e. to the the repo's code), you should consult the documentation of the wiki repo. Creating bgen files from plink · Creating a new page: On this wiki page click 'New page' · Give the page a title, at the top of the body of the page put something like Author: Your Name (if you want to Data management claim ownership), and then just write the snippet and relevant info below in Markdown format. To edit an existing page: simply go to a page and click on edit. Installing a specific version of R on ... Once you save the page the Github Actions will automatically update the https://mrcieu.github.io/wiki/ website. There's Links and resources nothing else you need to do. Mounting a network drive on Linux NOTE: Aside from some exceptions (including this one), EVERY visible page on this source wiki will be visible on the public facing https://mrcieu.github.io/wiki/ site. So please do not add information that might be sensitive or confidential (e.g. R tidyverse examples server address, patient information etc). Scripting Differences between the source wiki and the public wiki SITE README

#### See results

Editing Post practices in adding styles	Page bistory New page Delete page	🇯 MRC IEU Wiki		Q Search
Editing Best practices in coding styles	Page history New page Delete page			
Best practices in coding styles Write Preview		MRC IEU Wiki Home Data management Links and resources	Best practices in coding styles	Table of contents Coding styles Formatters
In1     h2     h3     ⊘     [⊠]     B     i     ↓ [Ξ]     iΞ     iΞ     IΞ     IΞ     Markdown     Φ		Scripting Software development Articles ~	Maintaining a good and consistent coding style is important in the sense that it improves the readability of your code.	
# Best practices in cooling styles Author: Yi Liu		Best practices in coding styles Code Testing	Coding styles	
Maintaining a good and consistent coding style is important in the sense that it improves the **readability ## Coding styles	y∗∗ of your code.	Creating bgen files from plink Installing a specific version of R on a Linux server using Conda	Below are some of the commonly practiced coding styles in various programming languages: <ul> <li>R: tidyverse style</li> </ul>	
Below are some of the commonly practiced coding styles in various programming languages:		Mounting a network drive on Linux R tidyverse examples	R: google     Python: PEP8     Python: coople	
<ul> <li>[R: Lagyerse style](https://style.lagyerse.org/)</li> <li>[R: google](https://google.github.io/styleguide/)</li> <li>[Python: PEPB](https://peps.python.org/pep-0008/)</li> <li>[Python: engle](https://engle.github.io/styleguide/)</li> </ul>		Using Biomake Using Docker Using Emacs	Shell: google	
<pre>= [Shell: google](https://google.github.io/styleguide/shellguide.html)</pre>			Formatters	
## Formatters			There are formatters in many programming languages that support formatting your code into a consistent style.	
Attach files by dragging & dropping, selecting or pasting them.	[13]		R: formatR	
Edit message			Python: black	
Write a small message here explaining this change. (Optional)			<ul> <li>web development (js/ts, ntmi, css): prettier</li> </ul>	