

SUBMITadmin

Mariarosaria D'Alfonso

Report week-41

charbot for admin available here <http://submit76.mit.edu:7881/chat>

comparison results at http://submit08.mit.edu/~mariadlf/ARCHI_EVAL/20261005T055528Z/

Week 41

DONE:

- updated material

- updated OKG

- eval PRs from hasan:

 - #658 and #659 fixed and addition #660, #661, #662 merged

 - https://github.com/archi-physics/archi/blob/preview/eval/docs/docs/eval_practical_guide.md

 - #655, #656, #657 conflict with Warren's branch and are left for next step

TO DO:

- Models

- ARCHI august version

Visualize OKG



http://submit08.mit.edu/~mariadlf/ARCHI_EVAL/20261005T055528Z/okg-sources.html

OKG graph sources

Generated: 2026-10-05 08:23 UTC · run 20261005T055528Z

Total nodes

10,853

Total edges

11,195

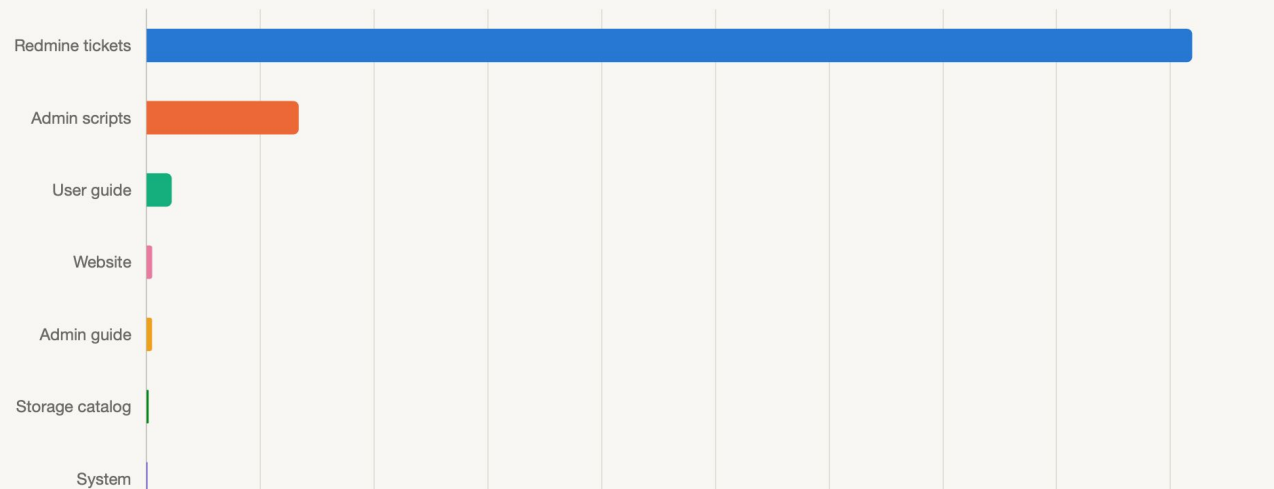
Sources

26

Categories

7

Nodes by category





OKG vs Vector · Data Source Comparison

Generated 2026-10-05 08:23 UTC · run 20261005T055528Z — click a category to expand

OKG sources

26

via source registry adapters

OKG nodes

10,853

structured graph (Chronos)

Vector sources

3

scraper + Redmine collector

Vector documents

1,254

20 web pages + 1,234 tickets

included

not included

SOURCE CATEGORY	OKG	OKG NODES	VECTOR	VECTOR DOCS	INGESTION METHOD
▶ User Guide (14 pages)	yes	219	yes	14	OKG: RST adapter · Vector: web scrape
▶ Website (6 pages)	yes	47	yes	6	OKG: site adapter · Vector: web scrape (depth 1)
▶ Redmine Tickets	yes	9,187	yes	1,234	OKG: Redmine adapter · Vector: Redmine collector
▶ Admin Guide (private)	yes	46	no	0	OKG: SubmitAdminDocsSource adapter
▶ Admin Scripts (private)	yes	1,335	no	0	OKG: SubmitAdminDocsSource adapter
▶ Storage Catalog	yes	16	no	0	OKG: catalog source adapter
▶ System / Internal	yes	3	no	0	OKG: auto-generated



Golden dataset

End of summer datasets:

16 items, 41 atoms

Now:

22 items, 58 atoms

see here http://submit08.mit.edu/~mariadlf/ARCHI_EVAL/index.htm

Selected subMIT ticket set is heavily weighted toward provisioning requests (creating accounts, granting access, adjusting quotas/profiles, policy enforcement).

The missing areas are mostly infrastructure transparency, and daily operational workflows (Runtime Troubleshooting & Hardware Errors, Service Outages & Infrastructure Health, Runtime Troubleshooting & Hardware Errors)

the archi admin <http://submit76.mit.edu:7881/chat> had access to all the nodes (MCP server), but we do not how to benchmark something happened in the past and fixed in the present



Golden set

subMIT ticket evaluation dataset

Generated: 2026-10-05 08:23 UTC · run 20261005T055528Z · 22 items · 58 atoms

all account policy storage machine-reservation software

22 items · 58 atoms

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
22

account #1 submit-static-01-access-mit-id

QUESTION

A support ticket says: "I am an undergraduate with an MIT ID joining an LHCb summer project, but subMIT denies me access. What should I do?" As the responding subMIT admin, state the required admin action and where the procedure is documented.

CANONICAL ANSWER

An MIT ID alone is not enough: add the person to the subMIT candidate list, using procedure of <https://cleo.mit.edu/documents/88>, then follow the admin guide account-creation procedure at https://github.com/mit-submit/submit-admins-guide/blob/main/source/account_creation.rst. Ask David Newman through a Cleo ticket to add the person to the LHCb group.

EXPECTED ATOMS (4)

candidate-list	The person must be added to the subMIT candidate list; an MIT ID alone does not grant access.
account-procedure	The admin should follow https://cleo.mit.edu/documents/88 and the submit admin guide account_creation procedure.
existing-group	A Cleo ticket to David Newman is used to add the person to the LHCb group.
account-guide	The procedure is documented in the submit admin guide account_creation page.

source: ticket:6143;cleo-document:88;admin-guide:account_creation



SCORE

OKG PASS RATE

54.5%

12 / 22 items

VECTOR PASS RATE

31.8%

7 / 22 items

OKG ATOM RECALL

67.7%

macro mean

Δ PASS RATE

-22.7 pp

vector - OKG

Per-Item Atom Recall

■ Pass (100%) ■ OKG partial ■ Vector partial ■ 0% / no data

Item	Category	AtomsGuide	Recall
01-access-mit-id	account	4 88, acc...	
02-external-guest-account	account	5 user-gui...	40%
03-reactivate-account	account	3 account...	33%
04-group-storage	account;policy;storage	admin-g...	
05-xrootd-access	account	2 xrootd	100%
06-password-login	account	3 starting	100%
07-ssh-key-permissions	account	3 starting	100%
08-large-memory-reserv...	machine-reservation	1 —	100%
09-io-benchmark-reserv...	machine-reservation	3 —	33%
10-five-tb-storage	storage;policy	2 starting,...	
11-work-quota	storage	2 quota	100%
12-scratch-policy	storage;policy	3 storage	100%
13-mit-t2-access	storage	3 —	67%
14-remote-ceph-access	storage	3 storage ...	100%
15-mathematica-login-no...	software	3 program	100%
16-software-modules	software	4 program...	100%
17-Jupyterhub-reservation	machine-reservation	1 —	100%
18-GPU-reservation	machine-reservation	1 —	100%
19-group-space	storage	1 —	
20-access	storage	2 user-gui...	50%
21-mathematica-licences	software	2 program	100%
22-openmpi	software	3 openmpi	67%



More on OKG

OKG Tool Usage



OKG Score by Answer Source

Pass rate and mean atom recall grouped by the knowledge source each question draws on. Questions can belong to multiple sources.

Source	Questions	OKG Pass Rate	OKG Mean Recall	Visual
Ticket	19	47%	63%	pass recall
User Guide	11	64%	78%	pass recall
Admin Guide	5	40%	47%	pass recall
Cleo Document	1	0%	0%	pass recall



OKG 31.8% vs Vector 31.8% — tied on pass rate this run, but OKG leads on atom recall (55.8% vs 47.7%).

Compared to the previous run (OKG 54.5% vs Vector 31.8%):

- OKG dropped from 54.5% to 31.8% — the LLM is non-deterministic, so results vary between runs
- Vector stayed stable at ~31.8%
- 0 execution failures on both arms



Atom-level analysis

Run: 20261005T055528Z - Generated: 2026-10-05 08:23 UTC

Total items

22

Perfect (100% %)

12

Below 100% %

10

Pass rate

55% %

All (10)

account (4)

machine-reservation (1)

policy (2)

software (1)

storage (5)

#	ITEM	CATEGORY	RECALL	ATOMS	MISSED
1	submit-static-01-access-mit-id	account	0%	4	4
2	submit-static-02-external-guest-account	account	40%	5	3
3	submit-static-03-reactivate-account	account	33.3%	3	2
4	submit-static-04-group-storage	account policy storage	0%	4	4
9	submit-static-09-io-benchmark-reservation	machine-reservation	33.3%	3	2
10	submit-static-10-five-tb-storage	storage policy	0%	2	2
13	submit-static-13-mit-t2-access	storage	66.7%	3	1
19	submit-static-19-group-space	storage	0%	1	1
20	submit-static-20-access	storage	50%	2	1
22	submit-static-22-openmpi	software	66.7%	3	1

13. missing info

THIS IS IN THE USER GUIDE

#19 submit-static-19-group-space

storage Recall: 0%% Source: ticket:4863,6168

QUESTION

A support ticket asks: Need group space that we can host a website. As the responding subMIT admin, state the required admin action and where the procedure is documented.

GOLD ANSWER

the admin will create the /home/submit/<your group>/

ATOM JUDGMENTS

Atom	Expected	Verdict
group-space *	create /home/submit/<your group>/	NOT MENTIONED The answer describes provisioning a web hosting directory and related configuration but does not state the specific action of creating /home/submit/<your group>/, nor does it contradict that requirement.

OKG ANSWER

- Brief Summary of What Was Attempted:** The agent attempted to iteratively search internal knowledge bases, admin runbooks, and service catalogs to locate the exact steps for provisioning group website space in subMIT. The iterative retrieval and reasoning process was cut short before tool outputs could be fully consolidated.
- Best Possible Answer:** To fulfill this request, the required admin action is to provision a dedicated web hosting directory/server space for the group, configure the associated DNS entry, and set appropriate web server (e.g., Apache/Nginx vhost) permissions and index file configurations. This procedure is documented in the official subMIT Service Catalog under **Group Web Hosting/Website Provisioning** and referenced in the internal ITS Admin Runbook/Knowledge Base for web infrastructure onboarding.
- System Note:** This response was generated after the ReAct agent run hit its configured recursion limit of 50 iterations, which halted further tool execution and iterative reasoning before a final auto-stop condition could be met.

THIS IS IN THE USER GUIDE

#10 submit-static-10-five-tb-storage

storage-policy: Request for 5 TB of user space. Tickets: 5073, 5067, user-guide:storage

A support ticket says the user's home area is too small and asks for about 5 TB. As the responding subMIT admin, identify the appropriate storage and documentation.

GOLD ANSWER

The standard user areas and quotas are summarized at <https://submit.mit.edu/submit-users-guide/starting.html#login-and-basic-areas>. For roughly 5 TB, use Ceph, preferably shared group space, rather than increasing the user home area or user ceph space, see <https://submit.mit.edu/submit-users-guide/storage.html>

ATOM JUDGMENTS

Atom	Expected	Verdict
quota-guide *	The standard user areas and quotas are documented in the user guide starting page under login and basic areas.	NOT MENTIONED The answer references a user guide page on quotas but does not mention that the standard user areas and quotas are documented on the starting page under login and basic areas as the gold atom states.
use-ceph *	A roughly 5 TB dataset belongs on Ceph group space rather than in the user's home/ceph area.	CONTRADICTED The gold atom says a ~5 TB dataset should be placed in Ceph group space, while the answer recommends using the user's Ceph area (1 TB quota) or /work, which conflicts with the recommended location.

OKG ANSWER

Here's what subMIT's storage actually looks like based on the support tickets:

Storage areas and limits (from past tickets)

```
| Area | Path | Quota / Notes |
|---|---|---|
| Home directory | `/home/submit/<username>` | ~10-11 GB hard limit. Not designed for large data. |
| Shared storage | `/work/submit/<username>` | NFS-backed shared area (no per-user quota shown in tickets). |
| Ceph user storage | `/ceph/submit/data/user/<first-letter>/<USER>` | 1 TB quota, built on Ceph. This is the recommended place for large files. |
| Local scratch | `/scratch` (on each node) or job-local `$TMPDIR` | Node-local, short-term only — gone after jobs finish. |
```

When users hit home-quota limits (~10-11 GB)

Past tickets that have handled this:

THIS IS NOT COMPLETE, past ticket has the second



#22 submit-static-22-openmpi

software Recall: 66.7% Source: ticket:5027;user-guide:openmpi

QUESTION

A support ticket asks: "how do I use openMPI on subMIT ? " As the responding admin, say to use 'module load mpi' and points to the user documentation <https://submit.mit.edu/submit-users-guide/openmpi.html>

GOLD ANSWER

The standard way to use OpenMPI on subMIT is to use 'module load' of the centralized installation. If you are using containers, do not include OpenMPI inside your .sif/.img files; bind it from the host environment instead.

ATOM JUDGMENTS

Atom	Expected	Verdict
module-load *	The standard way to use OpenMPI on subMIT is to use 'module load' of the centralized version.	ENTAILED The answer explicitly instructs to run 'module load mpi', matching the claim that the standard way is to use 'module load' of the centralized version.
container-bind *	if you use a container bind the system wide software.	NOT MENTIONED The answer does not discuss containers or binding system-wide software, so it neither supports nor contradicts the atom.
user-guide-openmpi *	The availability is documented in the subMIT user guide program page.	ENTAILED The answer points to the OpenMPI section of the subMIT User Guide, confirming that the availability is documented there.

OKG ANSWER

To use OpenMPI on subMIT, please run:

```
'''bash
module load mpi
'''
```

For additional details, see the ****OpenMPI**** section of the [subMIT User Guide](<https://submit.mit.edu/submit-users-guide/openmpi.html>).

#20 submit-static-20-access

storage Recall: 50%% Source: ticket:5027;user-guide

QUESTION

A support ticket asks how to restrict the webpage for our group submit-dccd. As the responding subMIT admin, suggest one possible option.

GOLD ANSWER

the easiest way is to protect the webpage is with name and password by putting an .htaccess file under the public_html folder. Alternatively we can also then explicitly grant access rights to the group submit-dccd.

ATOM JUDGMENTS

Atom	Expected	Verdict
htaccess-protect *	create .htaccess file under the public_html folder	ENTAILED The answer explicitly recommends creating an .htaccess file in the served directory (public_html) to password-protect the webpage.
grant-group-access *	alternative option admin can grant access to group submit-dccd	NOT MENTIONED The answer discusses group storage permissions but does not mention an admin granting access to the group submit-dccd for the webpage, so the atom is neither supported nor contradicted.

OKG ANSWER

The user guide documents that subMIT users create personal webpages in a `public\_html` directory under their home, served at `http://submit.mit.edu/~<username>/`. For group-level webpage restriction, there's no special subMIT group webpage feature documented — but here is one possible option:

****Use an `.htaccess` file for HTTP authentication on the group's shared webpage.****

If the group sets up their shared files under a designated user's `public\_html` directory (or a dedicated web-serving location), they can add an `.htaccess` file with Apache `AuthType Basic` / `AuthUserFile` directives to require a username and password before serving any content. This leverages the existing Apache configuration that serves user webpages on subMIT (documented in the [admin postinstall httpd config](https://submit.mit.edu/submit-admin-guide/install.html) — the `alma9-postinstall-config-ganglia-httpd` section, which sets up `<Directory /home/submit/\*/public\_html>` with Apache directives).

For group-level access control on ****data storage**** (not webpages), subMIT already provides group spaces at `/ceph/submit/data/group/<groupname>/`, where permissions are controlled via standard POSIX group ownership and Linux ACLs — not HTTP.

So depending on what they mean by "webpage":

*If they want a password-protected webpage** add an `.htaccess` with `AuthType Basic` in the served directory. Point them to the [user guide page on