



Sandboxed coding agents

LC user meeting
October 6, 2026

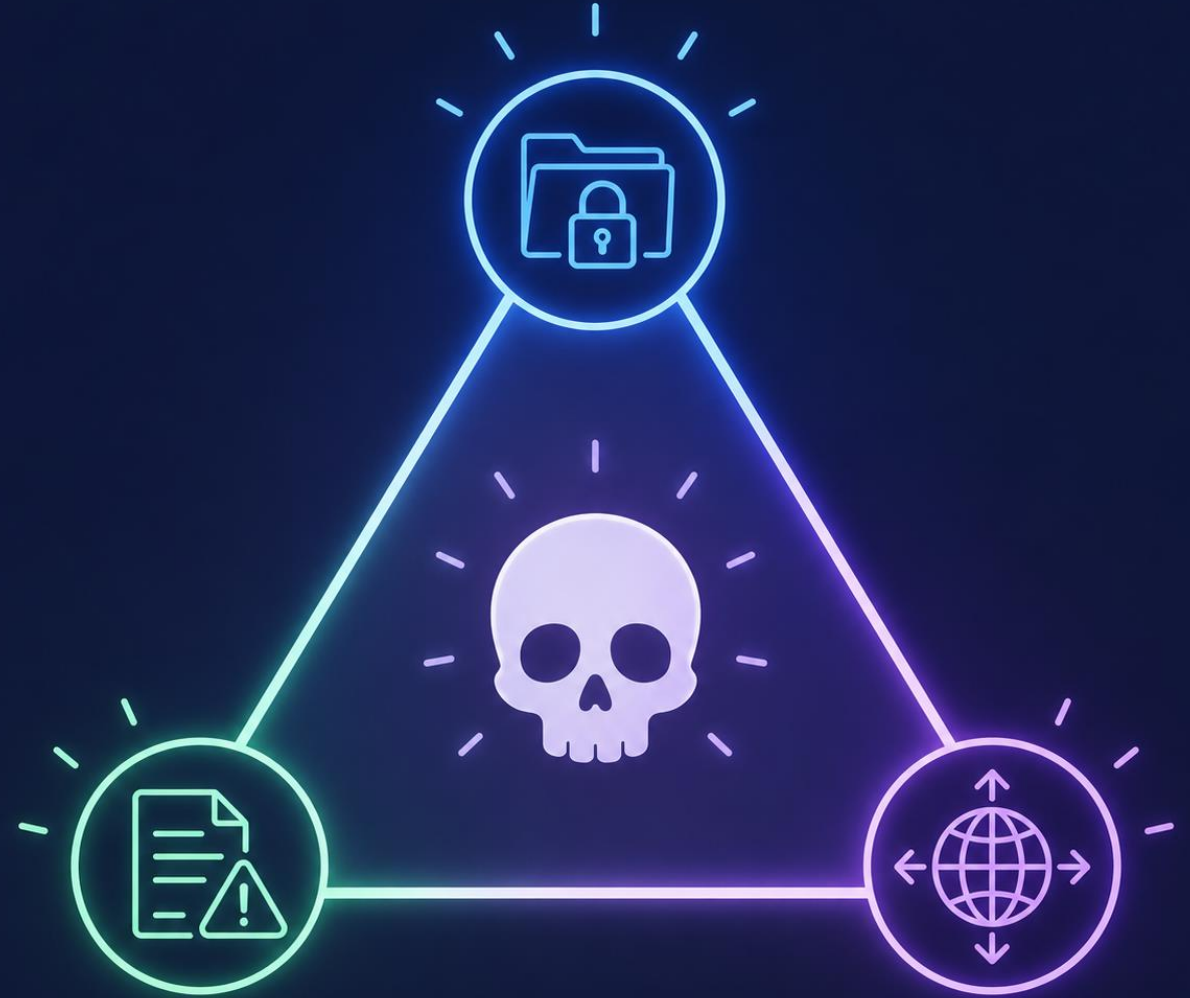
Prepared by LLNL under Contract DE-AC52-07NA27344.

Jane Herriman | Livermore Computing

Motivation

Avoid “Lethal trifecta”:

- Access to private data
- Exposure to untrusted content
- Ability to externally communicate



Sandboxing agents

- Sandboxing contains the agent
 - Restricts network access
 - Restricts filesystem access
- On LC, we recommend you run coding agents (claude, codex, opencode, etc.) in a blackhole sandbox



What's available?

- Sandboxing tools are in ``/collab/usr/global/tools/sandbox``
- Tools include:
 - blackhole :
 - This is the in-house underlying tool used to create the sandbox
 - claude-sandbox & codex-sandbox :
 - These start claude/codex inside a blackhole instance
 - sandbox :
 - This is just blackhole running a shell
 - You can run Claude/Codex/OpenCode/etc. inside
 - This requires some manual configuration (relative to e.g. ``claude-sandbox``)
 - Codex/OpenCode/Claude binaries (versioning options)

Getting started (first time)

```
[testnewm@oslic8 ~]$ module load claude
[testnewm@oslic8 ~]$ claude-sandbox
Missing /g/g0/testnewm/.livai-api-key.txt file.
Please add your livai api key to /g/g0/testnewm/.livai-api-key.txt
[testnewm@oslic8 ~]$ vi .livai-api-key.txt
[testnewm@oslic8 ~]$ claude-sandbox
Missing /g/g0/testnewm/.claude directory or /g/g0/testnewm/.claude.json file.
Please '/bin/mkdir -p /g/g0/testnewm/.claude && cp /collab/usr/global/tools/sandbox/etc/claude
/claude.json /g/g0/testnewm/.claude.json && cp /collab/usr/global/tools/sandbox/etc/claude/set
tings.json /g/g0/testnewm/.claude/settings.json'
[testnewm@oslic8 ~]$ /bin/mkdir -p /g/g0/testnewm/.claude && cp /collab/usr/global/tools/sandb
ox/etc/claude/claude.json /g/g0/testnewm/.claude.json && cp /collab/usr/global/tools/sandbox/e
tc/claude/settings.json /g/g0/testnewm/.claude/settings.json
```



Getting started: no top-level directories

```
[testnewm@oslic8 ~]$ claude-sandbox
time=2026-09-27T19:03:48.852-07:00 level=WARN msg="duplicate hostname already present" hostnam
e=livai-api.llnl.gov ip=128.15.130.110
time=2026-09-27T19:03:48.852-07:00 level=WARN msg="duplicate hostname already present" hostnam
e=llamame.apps.czapps.llnl.gov ip=134.9.65.24
time=2026-09-27T19:03:48.853-07:00 level=WARN msg="duplicate hostname already present" hostnam
e=livai-api.llnl.gov ip=128.15.130.110
time=2026-09-27T19:03:48.853-07:00 level=WARN msg="duplicate hostname already present" hostnam
e=llamame.apps.czapps.llnl.gov ip=134.9.65.24

allowlist entries:
  128.15.130.110 tcp:[80, 443]
  134.9.65.24 tcp:[80, 443]
2026/09/27 19:03:48 invalid source path: not on allowlist: /g/g0/testnewm
```

For example, avoid \$HOME, /p/vast1/\$USER, /usr/workspace/\$USER, /p/lustre*/\$USER

Getting started: once key & config are in place

- Enter a subdirectory
- Run
 - ``module load claude``
 - ``claude-sandbox``
- Or
 - ``module load codex``
 - ``codex-sandbox``

```
[testnewm@oslic8 ~/newdirectory]$ pwd
/g/g0/testnewm/newdirectory
[testnewm@oslic8 ~/newdirectory]$ module load claude
[testnewm@oslic8 ~/newdirectory]$ claude-sandbox
```

```
[testnewm@oslic8 ~/newdirectory]$ module load codex
[testnewm@oslic8 ~/newdirectory]$ codex-sandbox
```

- By default, you'll only be able to see your present working directory!

Getting started: directory access

- At runtime, give `rw` (read/write) or `ro` (read only) access

```
[testnewm@oslic8 ~/newdirectory]$ pwd
/g/g0/testnewm/newdirectory
[testnewm@oslic8 ~/newdirectory]$ claude-sandbox -- --add rw:$HOME/otherdirectory --add ro:$HOME/.codex
```

- Or list directory access in `sandbox-config.yaml` per project

```
mounts:
  paths:
    - "rw:$HOME/otherdirectory"
    - "ro:$HOME/.codex"
"sandbox-config.yaml" 5L, 75C
```

<https://hpc.llnl.gov/about-livermore-computing/ai-ml-lc/running-ai-agents-lc>

Recommendations

- Consider creating dedicated spaces for coding agent work
- Networking: agent can only reach LivAI & LLamaMe endpoints, so
 - Download what you need from a second terminal
 - Reach out to OISSOs for exceptions (additional URL access)

<https://hpc.llnl.gov/technical-bulletins/bulletin-608>