Scheduling

When running the Harvester using cron, add the following line to the beginning of the crontab file content to ensure the data are properly encoded:

LANG=en_US.UTF-8

```
0 */4 * * * /path/to/runIngest.sh
```

stub
The data harvester is convenient to run as a scheduled process utilizing cron. Once you have established your workflow for ingesting data publications cron can be utilized to schedule the task on a regular basis. For example the default run script can be added to cron to run every 4 hours and ingest and score new data.

The above example will execute the bash script every four hours. Alternatively, you could separate out the ingest process.

```
0 */4 * * * /path/to/runFetch.sh
0 */8 * * * /path/to/runTranslate.sh
0 17 * * * /path/to/runScore.sh && /path/to/runTransfer.sh
```

The above entries will call the fetch script once every 4 hours, translate once every 8 hours, and will run scoring and update VIVO every day at 5 o'clock PM local time. By breaking out the processes allows for even more complex workflows. If one wanted to run the above schedule for both PubMed and internal SQL data, runFetch.sh could be modified to fetch the data from multiple sources, etc.