INTERPRETING
THE
RESULTS
OF THE TIME TRIAL FOR INDEXING A TEST REPOSITORY

Parameters:

Repository Name: dset-web-accessible-folder-dev (2293 XML Metadata Files)

Test repository is **cloned on disk prior to testing**, **but NOT indexed** before the 3 trials

Settings (Before):	Settings (After):
Harvester indexes what's available on disk	Harvester indexes what's new from GitHub

Before	After
Trial 1: • Harvester finds the pre-cloned test repository on disk • 2293 XML files are indexed • Time taken: 547 seconds	**Trial 1:** • Harvester finds local repository, no changes found after GitHub check • 2293 XML files are indexed at the start • Time taken: 530 seconds
Trial 2: • Harvester is rebooted. No check with GitHub is made. • 2293 XML files are indexed once more • Time taken: 540 seconds	**Trial 2:** • Update event triggered by web app. Harvester finds the pre-cloned test repository on disk, pulls the changes from GitHub. • 3 XML files indexed • Time taken: 5 seconds
Trial 3: • Harvester is rebooted once more. The same behavior is observed from first 2 trials. • 2293 XML files indexed again • Time taken: 540 seconds	**Trial 3:** • Push Event on Github triggers webhook event. Harvester finds the pre-cloned test repository on disk. Pulls changes real-time. • 1 XML file indexed • Time taken: 1 second