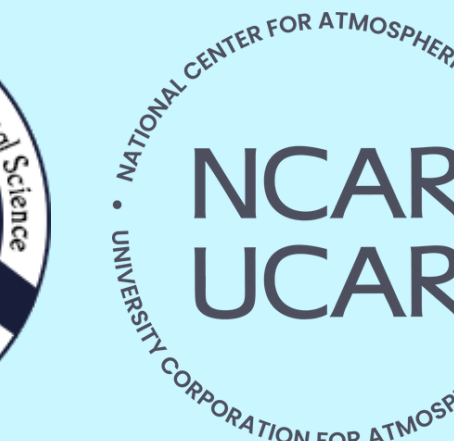


# Harvester Automation for Metadata Search Web Application

SIParCS Interns: Sama Manalai, Terry Yuan

SIParCS Mentors: Nathan Hook, Saquib Aziz Khan, Eric Nienhouse, Christy Grant



## Abstract

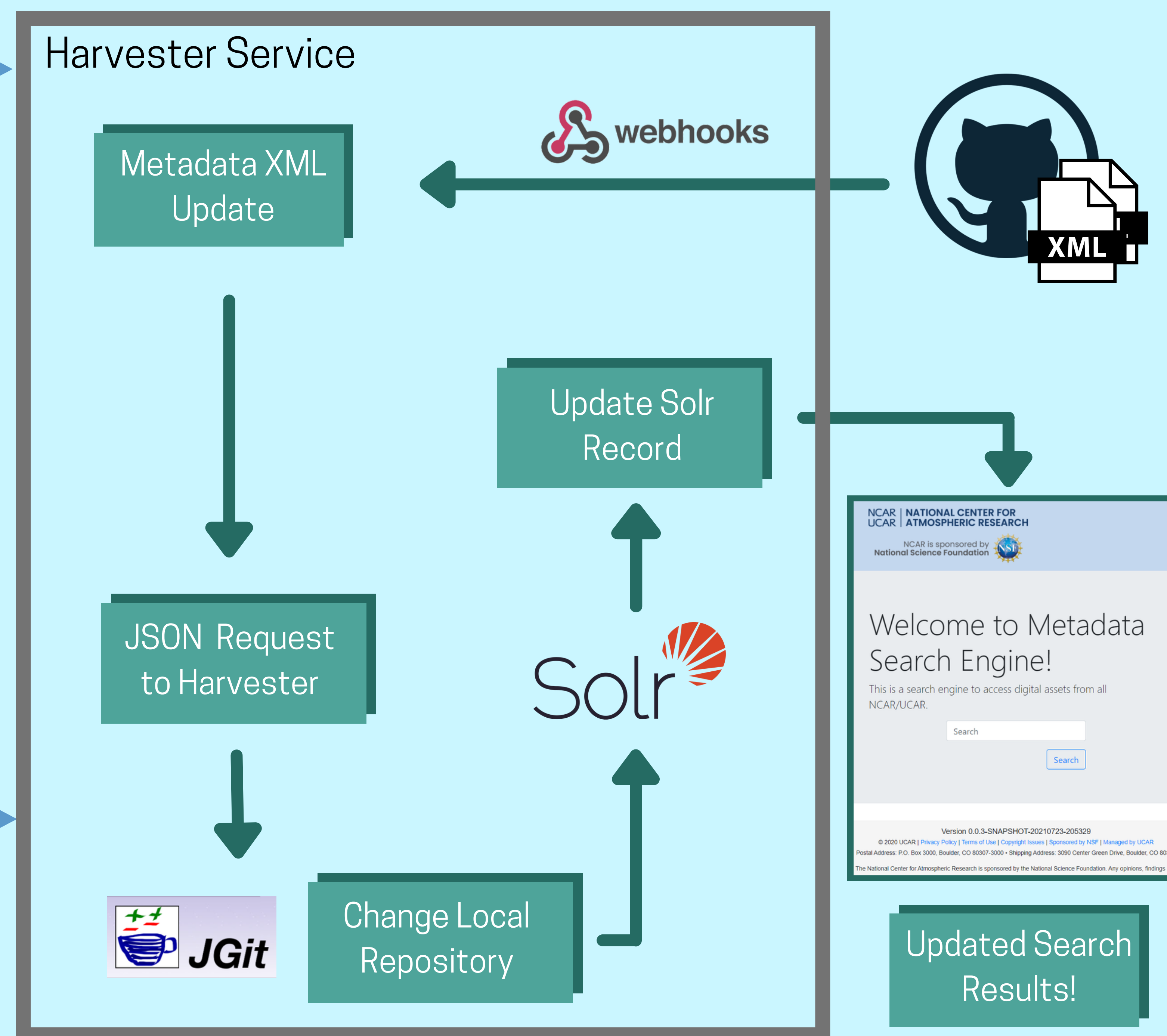
The Metadata Search Web Application is a new service application currently being developed within NCAR as a means to provide a search tool suitable for accessing NCAR-created digital media and establishing a maintainable codebase based on NCAR specifications. One of the main services offered in the web application is the Harvester service, which has been given increasing levels of autonomy to provide rapid metadata updates and reflect those changes on the Search service without manual intervention. The following discusses the methodologies and technologies employed to automate the harvester.

## Methodologies

- Agile Scrum
- SOLID Principles
- Layered Architecture
- Pair Programming
- Refactoring Code



## Harvester Service Automation Visualized



## Results/Time Trial for Indexing Test Repository

Total Indexed Files: 2293 XML Metadata Files

Before	After
Index: 547 secs	Index: 530 secs
Index: 540 secs	Check Update: 1 sec
Index: 562 secs	Check Update: 3 secs

## Technologies



## Future Work

- Treating GitHub repositories instead of local disk repositories as the authoritative source when getting metadata updates
- Implementation of autocomplete for search results
- Login and two-factor authentication (2FA) feature for Harvester controls
- Allowing developers to see a preview of ISO metadata when hovering over the ISO metadata link

## Acknowledgments

**SIParCS Mentors:** Nathan Hook, Saquib Aziz Khan, Eric Nienhouse, Christy Grant  
**SIParCS Leads:** AJ Lauer, Virginia Do, Jerry Cycone, Max Cordes Galbraith