# The Evolution of Geospatial Data Discovery at MIT
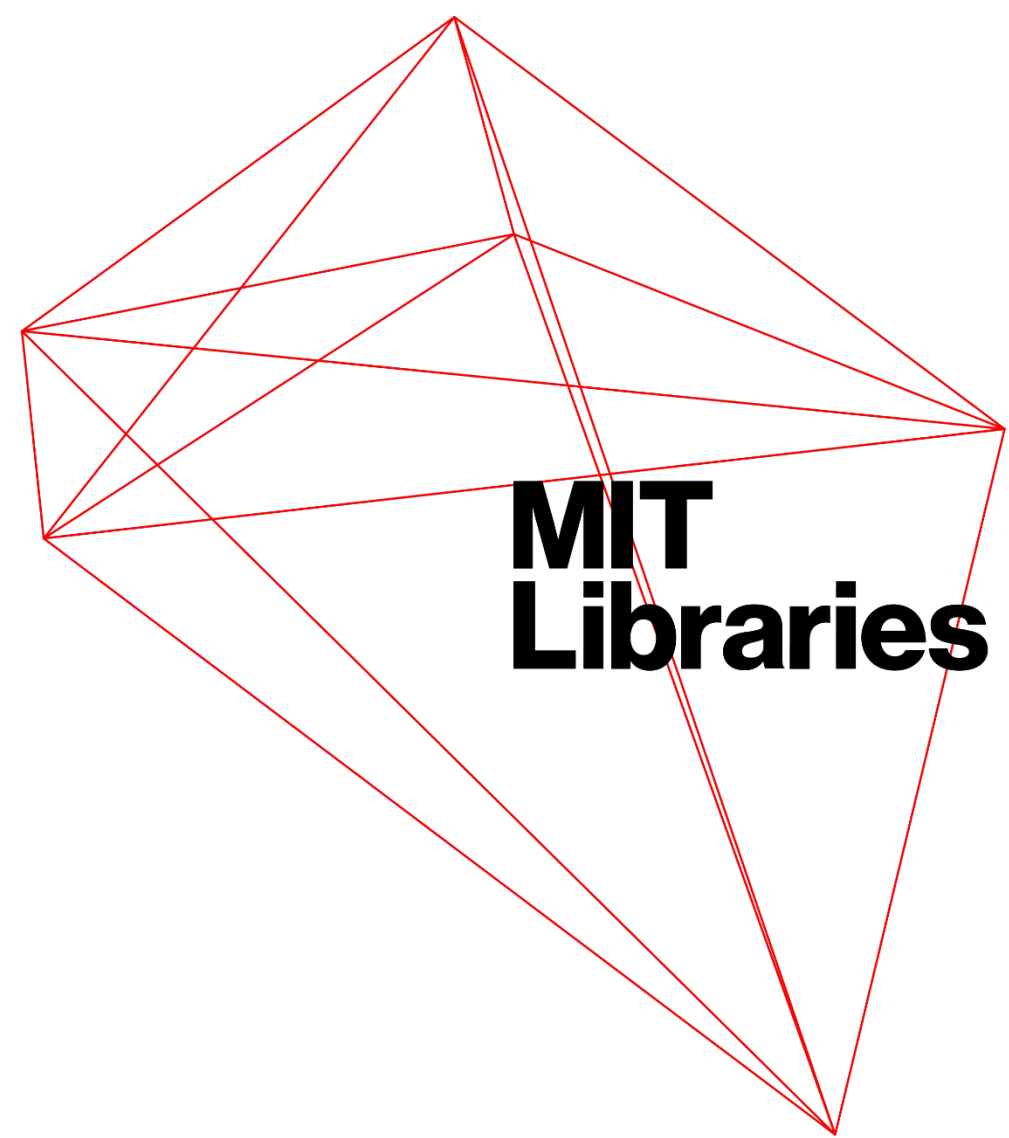
Paxton LaJoie - GIS Specialist in Education and Practice - MIT Libraries

**MIT Libraries**

The MIT Libraries has made many changes in the technology used to house our geospatial data collection over the 22 years we have been offering geospatial data discovery systems to the campus community. The Libraries first created access for users to our geospatial data collection in 2002, with an Esri geodatabase and a desktop search interface written in Visual Basic. Subsequent generations of the collection and search tool evolved from locally coded search interfaces using open source tools to OpenGeoportal and eventually to GeoBlacklight, with MIT Libraries engineers contributing to GeoBlacklight development. Throughout all of this, our geospatial collection remained separate from the rest of the MIT Libraries collections. Because of staffing and skill changes, technical limitations, and a desire to integrate all of our collections, MIT Libraries is currently moving from GeoBlacklight to a system directly integrated with other library systems and collections. The initial change will lose spatial search capabilities but will integrate our geospatial collection with other collections and will be powered by MIT Libraries' TIMDEX search API. We are currently planning the next steps beyond the TIMDEX text search, which will include both spatial search capabilities and the ability to examine geospatial data before downloading it.

## 2002

Development of a geospatial data discovery tool began in 2002 with conversations between MIT Libraries, the MIT Department of Urban Studies and Planning, and the Information Systems & Technology Group. The initial team created a tool for ArcView using Avenue, then moved it to ArcMap using Visual Basic. Geospatial data was stored in ArcSDE with metadata records stored in Oracle tables. This search tool functioned to add stored layers directly into ArcView and ArcMap projects.

## 2008

In 2008, MIT Libraries and the Dean for Undergraduate Education began developing a pilot project for a web-based interface for the geospatial data collection. The collection remained stored in ArcSDE and Oracle, and the web interface was written in JavaScript with a map interface in OpenLayers; developed by a student intern and the GIS Team.

## 2016

Since Open Geoportal was designed for Tufts University and its authentication systems, MIT experienced ongoing issues. During this time, Stanford University and others proposed a geospatial extension to Blacklight; Geoblacklight. MIT Libraries engineer Mike Graves joined the initiative to develop GeoBlacklight, which was implemented as MIT Libraries' new geospatial data search tool in 2016. It was written in JavaScript and Ruby on Rails, used a SOLR index for searches, GeoServer for data storage, and Leaflet plus OpenStreetMap for the map display.

## 2012

In 2012, the Libraries adopted the Open Geoportal system. This tool used a JavaScript interface coupled with an OpenLayers map interface. The Open Geoportal search tool moved the extraction of metadata, and the full metadata files, to a SOLR index while the data remained in ArcSDE and Oracle.

## 2023

The current geospatial data discovery tool stack is comprised of GeoBlacklight, GeoServer, and an MIT-built slingshot application. Functionality has decreased and it needs to be migrated from our Libraries legacy AWS account. To address these problem areas, a team of engineers and GIS professionals at the Libraries was created to execute a two-pronged solution. Starting in Fall 2023, the team began developing a Minimum Viable Project (MVP) using TIMDEX; which interfaces with an OpenSearch backend and exposes a GraphQL endpoint for data search. MVP goals include restoring the ability of the GIS team to ingest and index new geospatial data, provide download access to datasets for the MIT community, host text-based search and indexing of geospatial data and metadata together, and the ability to harvest and index geospatial data from other institutions.

## 2024

Now, in 2024, we are finishing development of the TIMDEX MVP project and moving on to user testing. The future of geospatial data discovery at MIT Libraries will continue to execute the second part of the MVP project: the development of a long-term tool meant to be actively maintained and populated with new geospatial data. Current plans for the upcoming project revolve around building on the functionality of the TIMDEX tool, including the implementation of metadata editing functions, creating a map interface for geographic data searches, and implementing a feature for visualizing previews of data layers stored in the repository.