

Status

This project is sponsored by [Cancer.gov](#) to aid **sharing human biospecimens** with select diagnostic and treatment criteria.

Status:

2005: Approved for use at four Harvard affiliated teaching hospitals

2006: Initial open source release for Pathology Diagnoses ([Linux.com article](#))

2007: Completely rewritten API to improve performance, reproducibility, and hospital-specific customizations ([2.X branch](#)).

2008: Extended to support scrubbing other kinds of notes such as patient discharge summaries.

2009: Approved for use at two large HMO sites.

2010: Machine Learning work begins using millions of peer-reviewed publications to train "ham" (medical concepts) from "spam" (patient identifiers).

2011: [3.X development active](#)

2012: Research paper submission

2013: [3.0 software released as candidate](#)