Mission: All knowledge recorded for everyone's free use
Review, April Golden Gate Computer Society Meeting
By Mike Hancock, Newsletter Committee
May/June GGCS newsletter
www.ggcs.org
editor (at) ggcs.org

WITH A MISSION OF PROVIDING UNIVERSAL ACCESS to all knowledge forever, Internet Archive (www.archive.org), a San Francisco based non-profit, has archived 100,000 software titles, 2 million moving images, 3.5 million audio recordings, 5 million hours of television, 3 million e-books, and more than 327 billion web pages: a total of 40 petabytes archived.

John Gonzalez, Director of Engineering at Internet Archive, leads the building of a digital library of internet sites and other cultural artifacts in digital form and introduced his organization during the April 23 General Meeting.

Internet Archive's enormous archive, which has 2 million to 3 million visitors a day, is explored via the Wayback Machine, a search engine. Archive-It, another tool on the site, enables you to capture, manage, and search collections of digital content without any technical expertise or hosting facilities. No user I.D. or password is required for viewing, but a login is required for uploading.

The Internet Archive has only 150 staff members but runs one of the top 250 websites in the world. Reader privacy is critical, and it doesn't accept ads that track your behavior. Internet Archive doesn't even collect your IP address. It's also a member of the American Library Association.

John demonstrated access to the e-book library, where you can borrow e-books, one at a time, many of which are very old books that have been digitized. Although some books are torn apart to archive the digitizing, a super-scanning center in Hong Kong has a machine that enables the books to be laid open at 90˚ while the two open pages are scanned.

There are 28 scanning centers in eight countries that scan 250,000 books each year. The e-book collection now bridges a large gap in printed books that were restricted because of copyright limitations and e-book restrictions and includes some books in foreign languages.

Also available and demonstrated to us were 78 rpm records, old-time radio, and live music, including old and modern music. Another fun archive is that of feature films that you can access and play on your computer. Old and new television programs also are archived as well as maps.

Journalists use Internet Archive's Political TV Ad Archive in their fact-checking of candidates' claims. It includes apps for iPhones and Android devices. "Facets" assist you in your detailed searches.

While browsing the archives, you can capture a web page as it appears when viewing for use as a trusted citation in the future. This Save Page Now feature is only available for sites that allow crawlers.

"Let's Explore" has 14 archive.org websites that can be copied/posted for a lot of interesting information. These links are fun to spend time reviewing, and include Historical Software Collection, Amateur Radio Today, MIT press, Folkscanomy Library, Byte Magazine, NASA Images and Reports, 78 rpm records, and more.

The history of libraries is one of loss, so this year, Internet Archive has set a new goal: to create a copy of Internet Archive's digital collections in another country. They are building the Internet Archive of Canada because copies keep records safe.

This project will cost millions, which is why a tax-deductible donation is appreciated. About 60% of Internet Archive's income is earned from working with other institutions, such as the Library of Congress; up to 30% from donors and grants; and the balance from founder, Brewster Kahle.