Protecting Patron Privacy with Evergreen

Galen Charlton and Jeff Godin
2017 Evergreen International Conference

Privacy and security

"We protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted."

- Privacy
- Security

Protecting patron privacy in Evergreen

- During this workshop we will give you lots of tips for configuring and running Evergreen in a way to protect the privacy of your patrons
- But protecting patron privacy doesn't start or end with the ILS

General principles

- It takes the entire library to do it
- Protecting patron privacy is not a state of being, it is inherently an active process
- Incremental improvements are better than no improvements

Involving the entire library

- So, who is here? And what are your roles?

The plan

- Introduction
- What this is, and isn't
- Why privacy?
- Laws, ethics, and policy
- Types of attacks
- A bestiary of attackers
- Evergreen as a target
- Mitigations and defense
- Recommendations
- Revisiting general principles

What this workshop is

- An overview of patron privacy as a matter of law and professional ethics
- An examination of threats to patron privacy in the context of an Evergreen system
- A list of concrete steps you can take to improve how Evergreen protects your patron's privacy
- As a side benefit, an opportunity to share amongst ourselves

What this workshop is not

- An overview all system security issues and concerns — a DDOS may ruin your day, but does not necessarily have implications for patron privacy
- Focused purely on technical details of Evergreen server administration — protecting patron privacy is the concern of the entire library.
- Not focused on Evergreen development or coding practices for avoiding security pitfalls.

Why privacy?

Why protect patron privacy?

- Professional obligation
- It's often the law
- Intellectual freedom does not thrive in a panopticon
- Disclosing confidential patron records can cause concrete harm
  - Folks exploring their identity
  - Folks exploring controversial topics
  - Folks who have stalkers
- Financial liability: breaches can be expensive
- Libraries as trusted and trustworthy institutions
- Patrons expect it

ALA Code of Ethics

III. We protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted

http://www.ala.org/advocacy/proethics/codeofethics/codeethics

Legal requirements — Kentucky

You have requested that we clarify OAG 81-159 in which we said that a public library is not required to make available for public inspection its registration and circulation records. We said: "We think that the individual's privacy right as to what he borrows from a public library (books, motion picture film, periodicals and any other matter) is overwhelming." This conclusion was based on KRS 61.878 (1)(a) which exempts from the mandatory requirement of public disclosure "public records containing information of a personal nature where the public disclosure thereof would constitute a clearly unwarranted invasion of personal privacy."

Legal requirements — Ohio

(B) A library shall not release any library record or disclose any patron information except in the following situations:

(1) If a library record or patron information pertaining to a minor child is requested from a library by the minor child's parent, guardian, or custodian, the library shall make that record or information available to the parent, guardian, or custodian in accordance with division (B) of section 149.43 of the Revised Code.

Legal requirements — Ohio

(2) Library records or patron information shall be released in the following situations:

(a) In accordance with a subpoena, search warrant, or other court order;

(b) To a law enforcement officer who is acting in the scope of the officer's law enforcement duties and who is investigating a matter involving public safety in exigent circumstances.

(3) A library record or patron information shall be released upon the request or with the consent of the individual who is the subject of the record or information.

Legal requirements — Indiana

(b) Except as otherwise provided by subsection (a), the following public records shall be excepted from section 3 of this chapter at the discretion of a public agency: <...>

(16) Library or archival records:

(A) which can be used to identify any library patron; or

Library privacy policies — San Francisco PL

http://sfpl.org/?pg=2000001301

1. The Library will keep all such information that it purposefully or inadvertently collects or maintains confidential to the fullest extent permitted by federal state and local law, including the California Public Records Act1, the San Francisco Sunshine Ordinance2, and the USA PATRIOT Act3.

Library privacy policies — San Francisco PL

10. The Library does not maintain a history of what a library user has previously checked out once books and materials have been returned on time.

…

Library users may choose to opt in and enable My Check-out History. By doing so library users choose to give explicit consent to the storage of their Check-out History from the opt-in date. Library personnel will not access or release Check-out History unless required by law to do so. Library users may opt out of this service and delete Check-out History at any time. (Noted - November 30, 2011)

Library privacy policies — San Francisco PL

7. The Library protects library user account information by placing it on a secure server.

…

15. The only information stored on the RFID chip/tag will be limited to the item barcode or an encrypted number, as well as a security bit that indicates if the item is in or out of the library.

16. RFID technology will not be used for library cards.

Library privacy policies — Hennepin County

http://www.hclib.org/about/policies/disclosing-patron-data-staff-and-volunteer-responsibilities

Use and disclose patron data only to conduct the library business for which it was collected.

Refrain from sharing information about patrons or patron activities that is unnecessary for the provision of library service. This includes, but is not limited to, postings about patrons on social media sites, and informal conversations.

Library privacy policies — Hennepin County

Immediately inform supervisor or person in charge when a request to disclose private data is made by law enforcement officials.

Supervisors and managers have the following additional responsibilities:

1. Ensure patron data is used appropriately and take disciplinary measures if direct reports are in violation of this policy.

Library privacy policies

- Should not be static
  - Particularly when the policy doubles as a disclosure mechanism

Consortial privacy policies — Evergreen Indiana

http://www.in.gov/library/files/Evergreen_Indiana_Patron_Record_Confidentiality_Policy_031009.pdf

6. In the event of the receipt of a valid third party Patron Record request, a Member Library shall only disclose or release the Patron Records of a person whose library membership originated in its particular library. No single Member Library shall disclose or release the Patron Records of any person whose library membership originated with a different Member Library.

Consortial privacy policies — Evergreen Indiana

10. Any Evergreen patron requesting by telephone a list of items checked out on a specific card must use the barcode number. Library staff shall not give out any specific information without the Evergreen barcode number; staff may only give out the number of items due and the due date.

Questions? Things to share?

Threats, attackers and exposures

Types of attacks - passive observation

- Observing data in transit
- Difficult/impossible to detect
- Information obtained can be used for other attacks
  - especially short-lived tokens or re-usable credentials
- Generally made possible by lack of encryption or by weak/flawed encryption

Types of attacks - active exploits

- Man-in-the-Middle attack
  - generally requires a privileged network position
  - often exploit TLS vulnerabilities and/or involves a user acknowledging a security warning
- Exploiting an application or operating system vulnerability
  - access control bypass
  - privilege escalation
- Other forms, different severities
  - full read access to data
  - full read/write access to Evergreen system
  - full access to servers at the OS level

Types of attacks - social engineering

- Can you tell me my password?
- No? Can you reset my password?
- No? Can you at least update my email address? I haven't been receiving my overdue notifications...

Attacks that aren't really attacks

Attacks that don't involve a malicious actor breaking any of "the rules" and cannot exclusively be addressed at a technical level, such as:

- Requests for data from law enforcement / court order
- Accidental disclosure of data in bulk, such as via unsecured backups
- Failure to properly dispose of printed records or storage media

Bestiary of attackers — law enforcement

- Local and state
  - Subpoenas
- Federal
  - PATRIOT act

Bestiary of attackers — crackers

- Attackers who care about library PII
  - Directory information
  - Payment information
  - Credentials
- Attackers who don't care
  - Script kiddies and vandals
  - Botnet operators

Bestiary of attackers — third parties

- Adobe
- Web analytics

Bestiary of attackers — other patrons

- Parents and guardians
  - Yes, this is tricky and delicate

Bestiary of attackers — library staff

- What does local celebrity X like to read?
- Fine-shamers

Questions? Things to share?

Evergreen as a target

Data Evergreen stores

- Patron name, address(es), phone number(s), email, password hash…
- User activity records, including type and time of activity
- Circulations, hold requests
- Billings and payments
- Lists
- Notes, standing penalties, alert messages, message center messages
- Record of notifications (email, sms, print, phone) including message bodies
- Record of password resets
- History of many of the above in auditor schema

Data Evergreen stores

- Don't forget about patron information that may slip into MARC records
  - E.g., don't use a 9XX field to record purchase requests
- Similar warning for item/copy level information
  - Don't record patron details in a copy alert message

Data Evergreen transmits

- Between servers
- Bulk data transfers
  - batch loading of patrons to a state resource-sharing system
  - offline patron (block) file
- Third party or library use of APIs
- Third party web/screen scraping

Mitigations and defenses

Protecting Evergreen — basic hygiene

- System administration
- Ongoing OS updates
- Apache
- XML processors
- Network security

Protecting Evergreen — network

- Know your ports
  - 210 — Z39.50
  - 80 — HTTP
  - 443 — HTTPS
  - 5432 — PostgreSQL
  - 5544 — MARC stream importer
  - 6001 — SIP2
  - 7680 — WebSockets
  - 7682 — Secure WebSockets
  - 11211 — memcached

Protecting Evergreen — networking

- Firewalls and filters
  - iptables
  - iptables wrappers like lfw and ufw
  - Firewall appliances
- Recursive security exposures
- Proxy servers
  - NGINX and HAProxy
- Hard shell, creamy filling

Protecting Evergreen — permissions

- Evergreen permissions
- Operating system permissions
  - Who is allowed to log into your server?
- Database permissions
  - pg_hba.conf
  - Limited-purpose users

Logging and monitoring

- Logs - useful for so much more than just filling filesystems!
- Monitoring for availability and performance is a start, but you can also monitor for security/privacy related metrics
  - Report/graph successful / unsuccessful logins
  - Check TLS ciphers / security-related HTTP headers
  - Ensure that URLs that should require authentication continue to do so
- Know what "normal" looks like
- Automate checks and alerting, but beware alert fatigue

The overall library IT environment

- Lock workstations when they are unattended, especially in public areas
- Consider shortened "idle timeout" time on service desk computers
- Evaluate other workstation risks: missing patches, malicious USB devices...
- Have a password policy that encourages / requires good password practices
  - NIST 800-63-3 draft looks promising
- Deactivate accounts when staff become former staff
- Don't use shared role accounts with widely-known passwords
- Deactivate accounts when staff become former staff
- **Don't use shared role accounts with widely-known passwords**

The overall library IT environment

- Ensure that your shared role accounts with widely-known passwords have the minimum privileges required
- Change the password on your shared role accounts from time to time
    - When staff start referring to a given account by its password as opposed to its username, it may be time to change the password

How Evergreen protects privacy out of the box

- SSL/TLS for staff and patron account functions
- Reasonable password storage mechanisms and password reset functionality
- Permissions system
  - Patron OU opt-in
- Holds Aliases, useful for public hold shelf
- Redaction of credentials in Evergreen's logs
- Evergreen security team
  - Following good security coding practices as best we can
    - SQL injection protection
    - sensible defaults — we try not to ship too many foot-cannons
  - Updating as security knowledge increases — e.g., updating the password-hashing algorithm

How Evergreen protects privacy out of the box

- Limits on things like easily-accessible previous circ history of an item
- Aged Circulations
  - including preservation of select statistics
- Opt-in checkout/holds history
  - patron opt-in, not staff opt-in
  - patron viewable, not staff viewable
  - checkout history is user-editable and user-purgeable

TLS and certificates

- SSL is dead, long live TLS
- Certificates: Important protection against MITM
  - today, these certificates are rarely used with SSL traffic, but people still call them "SSL Certificates"
- Certificates establish a chain of trust between your client/browser and the server by way of a Certificate Authority (CA)
- "Add SSL exception" in XUL client = TOFU
- Web browsers are intentionally making it more difficult to bypass / override certificate errors
- As a site owner, you can force HTTPS and require a key from a set of CAs using HSTS and HPKP headers **and some caution**

TLS and certificates - LetsEncrypt

- LetsEncrypt - free domain-validated certificates
- certbot is wonderful
- certbot doesn't run everywhere -- other ACME clients
- CT logs = come and get it!
- Might still need/want an EV cert
- no wildcard certs, but SANs aplenty

SIP2 and tunnelling

- SIP was designed to pass over serial cables
- If you must pass SIP2 traffic over untrusted networks, please tunnel/encrypt
  - SSH forwarding
  - SSL (TLS!) tunnel
- If using SIP2 with an external vendor, this requires coordination
  - Some vendors are better at this than others

General issues with vendor authentication

- SIP was designed for performing staff actions
- Granting a vendor SIP access is in many cases equivalent to granting staff credentials to your ILS, and may provide more access than is strictly necessary for the task of patron validation
- Vendors usually live on the other side of the Big Bad Internet -- tunnel!

Protecting Evergreen — tunneling Z39.50

In this case, this would be protecting *external* users searching your catalog via Z39.50

(Not catalogers looking for records to import)

https://galencharlton.com/blog/2015/10/securing-z39-50-traffic-from-koha-and-evergreen-z39-50-servers-using-yaz-and-tls/

Data retention

- Evergreen lets you store a lot of information about patrons
  - How to contact them
  - What they request and check out
  - How much money they accrue and pay in fines
- Evergreen servers gather a lot of information about patrons, "incidentally"
  - Their home IP addresses
  - What browsers and other technology they use

Data retention

Computers are *great* at making copies of data

- Backups
  - How long do you need to retain backups?
  - For what kinds of data? E.g., consider auditing requirements for acquisitions records
- Test instances and databases
  - Consider taking steps to prevent login for all but your test users

You can't retain what you don't collect

- Social Security Numbers?
  - Let's not.
  - No.
  - Just no.
  - ABSOLUTELY NOT.
- Proofs of identity and residence
  - You may need to check a driver's license or state ID when registering a patron…
  - … but do you *really* need to store the DL number online?
- DEBATE
- Information about family relationships

Protecting data *outside* of Evergreen

- If you must correspond about patrons with other staff, refer to them strictly by barcode or patron ID
- Use the patron alias field for slips
- Refer patrons to the "My Account" page where possible instead of emailing details about their circulation
- Limit creation and distribution of detailed reports
- Log files from devices and services that interact with Evergreen

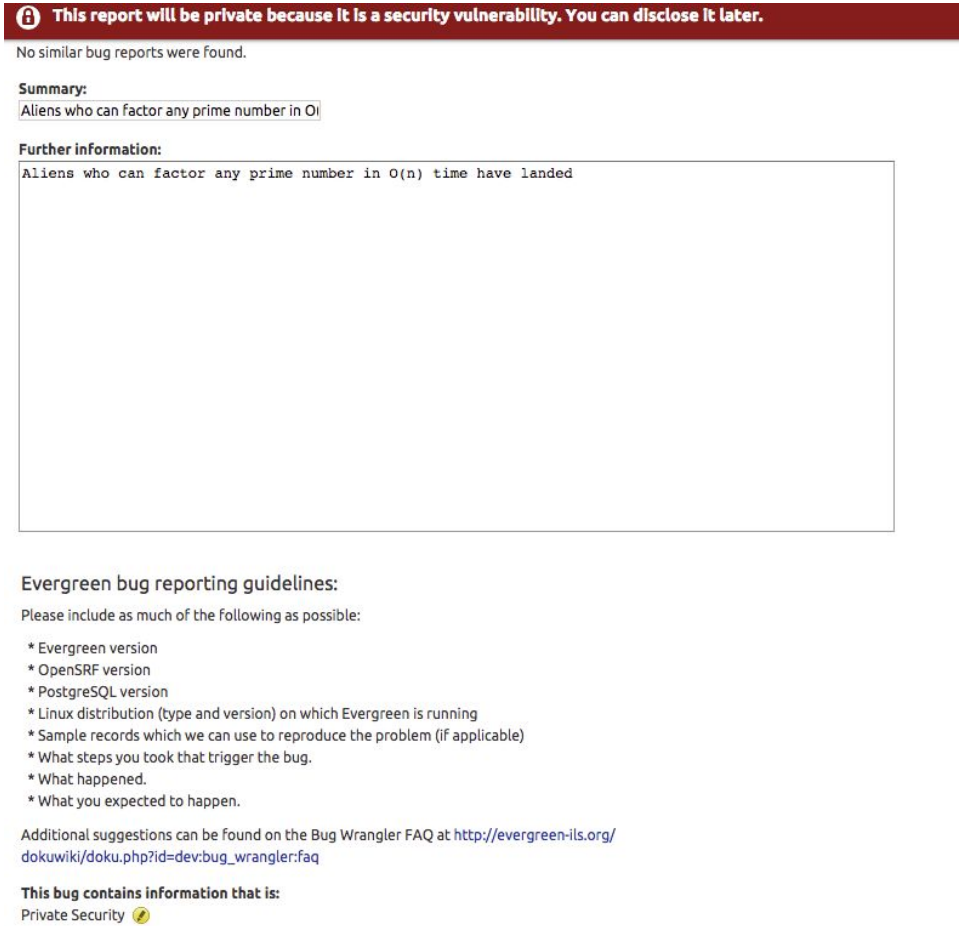
Education of third parties as a mitigation

- It's not necessarily clear how to exchange data with Evergreen securely
- Some third parties know about patron confidentiality concerns… others, not so much
- The more you can help educate third parties, the better
- Evergreen's open source nature can help

Questions?

Reporting security bugs

- Launchpad
- Please try to avoid publicly disclosing how a vulnerability can be exploited by mistake



Recommendations

- Stay on top of operating system and application updates
- Know who has access to your Evergreen system, how, and for what purpose
- Upgrade Evergreen as frequently as you can
  - And apply Evergreen security updates ASAP
- Have written privacy policies
- Have written data retention policies
- Turn on HTTPS across the board
- Encrypt data exchanges
- Treat test systems as seriously as production systems

Staff Training and Documentation

- Staff training and documentation is essential
- Remember to incorporate patron privacy into existing training/docs
- Consider having privacy-specific training
- Include scenarios that demonstrate social engineering attacks

Resources

- ALA/LITA Library Privacy Checklists: http://www.ala.org/lita/advocacy
  - Data Exchange
  - E-Book Lending
  - ILSs
  - Library Websites
  - Public Access Computers
  - Students

General principles redux

- It takes the entire library to do it
- Protecting patron privacy is not a state of being, it is inherently an active process
- Securing your Evergreen system is a necessary act, but not sufficient
- Incremental improvements are better than no improvements

Thanks!

Galen Charlton, Equinox Open Library Initiative

Jeff Godin, Traverse Area District Library