This paper appears at the IEEE European Symposium on Security & Privacy (2017).

Confidante: Usable Encrypted Email

A Case Study With Lawyers and Journalists

Ada (Adam) Lerner*, Eric Zeng*, Franziska Roesner
University of Washington
{lerner, ericzeng, franzi}@cs.washington.edu

***Abstract*—Email encryption tools remain underused, even by people who frequently conduct sensitive business over email, such as lawyers and journalists. Usable encrypted email has remained out of reach largely because key management and verification remain difficult. However, key management has evolved in the age of social media: Keybase is a service that allows users to cryptographically link public keys to their social media accounts (e.g., Twitter), enabling key trust without out-of-band communication. We design and prototype Confidante, an encrypted email client that uses Keybase for automatic key management. We conduct a user study with 15 people (8 U.S. lawyers and 7 U.S. journalists) to evaluate Confidante's design decisions. We find that users complete an encrypted email task more quickly and with fewer errors using Confidante than with an existing email encryption tool, and that many users report finding Confidante comparable to using ordinary email. However, we also find that lawyers and journalists have diverse operational constraints and threat models, and thus that there may not be a one-size-fits-all solution to usable encrypted email. We reflect on our findings — both specifically about Confidante and more generally about the needs and constraints of lawyers and journalists — to identify lessons and remaining security and usability challenges for encrypted email.**

1. Introduction

Encrypted email tools have long faced profound usability challenges. The first well-known paper on email encryption usability, "Why Johnny Can't Encrypt" [31], is from 1999, and numerous papers have since confirmed and expanded its main findings (e.g., [8, 11, 12, 21–26]). To date, the usability of email encryption has remained a challenge.

Indeed, due to the difficulty of email encryption, its use is limited even among those user populations with strong reasons to secure their communications, such as journalists [16, 19] and lawyers [15]. Although there are reasonably usable encrypted messaging applications for smartphones, such as Signal and WhatsApp, they require that users join a closed ecosystem. One of the benefits of traditional email, by contrast, is its openness and backwards compatibility. In our user study (Section 8), we find that email is a dominant communication channel for lawyers and journalists, and that securing it is critical.

*Co-first authors listed in alphabetical order.

Design and Prototype of Confidante. In this paper, we take another look at usable encrypted email. We design and implement Confidante, an encrypted email client that explores a new point in the design space. Motivating Confidante's design are two observations. First, we observe that many of the core usability pitfalls with encrypted email stem from problems with key management, sharing, and verification. As we find in our user study, and echoing past findings (e.g., [8, 11, 12, 25, 26]), asking users to reason about key management wastes time and risks critical mistakes.

Second, we identify an opportunity to make new progress on this long-standing challenge by leveraging a recently developed key management solution, *Keybase* (https://keybase.io). Keybase is a new take on key trust enabled by the public social media identities common on today's web. It allows users to post signed cryptographic proofs associating their public keys with public social media accounts (e.g., Twitter, Reddit), requiring other users only to check that these social media accounts belong to their intended communication partner. We describe Keybase and its security properties in more detail in Section 2. Section 4 describes the design of Confidante, a PGP email client that interfaces with a user's existing email account (Gmail, in our prototype) and leverages Keybase for key management. We use our prototype to explore whether, given key management via Keybase, we now have all of the building blocks for usable encrypted email.

User Study with Journalists and Lawyers. To evaluate Confidante's design choices, we conduct a qualitative user study with 8 U.S. lawyers and 7 U.S. journalists. We target these user groups because they may communicate sensitive information over email with their clients, sources, and colleagues, and thus may be motivated to use encrypted email — but currently do not frequently use it [15, 16, 19].

We design our user study to evaluate the design choices made in Confidante, including the role of Keybase, through a comparison with Mailvelope, an existing browser extension for encrypted email. We do not compare Confidante and Mailvelope as whole products, but instead explore individual design differences and surface recommendations for future designs. For example, we find that automated key management via Keybase makes using Confidante feel comparable to regular email for many users, but that some users also doubt its security properties due to this ease of use. Our study also highlights potential challenges associated with

using Keybase, including private key management and the social media linking requirement.

Finally, we encounter a surprising side-effect resulting from our choice of participants from specific user groups. The findings from our user study allow us not *only* to evaluate our research prototype, but *more generally* to shed light on the threat models and security and usability needs of lawyers and journalists for encrypted email. Most importantly, these findings suggest that lawyers and journalists have security and operational requirements divergent enough that they may require different tools entirely. For example, we find that the security goals of journalists and lawyers differ. Journalists must protect their communications, including metadata, indefinitely. U.S. lawyers, however, must only take "reasonable steps" to protect communications, since in practice, attorney-client privilege renders evidence produced by circumventing those "reasonable steps" inadmissible in court. This finding should give us hope: by targeting specific user groups, we can in some cases make abstract concessions on security that do not violate the *specific* security needs of those users, but that enable usability and adoption.

Contributions. We make the following contributions:

1) We design and implement Confidante, a backwards-compatible PGP email client that automates the cryptographic operations of encrypted email and uses Keybase to manage and verify keys.
2) We conduct a user study with 8 U.S. lawyers and 7 U.S. journalists to evaluate Confidante's usability and security design choices. We compare to design choices in Mailvelope, an existing browser extension for encrypted email.
3) Beyond evaluating Confidante, our user study sheds light more broadly on the threat models and use cases for encrypted email among journalists and lawyers.
4) Based on these findings, we make recommendations for future tools and research directions in usable encrypted email. For example, we suggest focusing on specific user populations separately.

2. Background and Motivation

2.1. Encrypted Email Usability

Usability and adoption have been well-known challenges for encrypted email since the first paper on the topic in 1999 [31]. Issues identified with PGP 5.0 [31] remain with PGP 9 [26], and other tools built on PGP, such as Mailvelope, also suffer from usability issues [24]. Recent work has explored the usability and user comprehension effects of automatic and manual encryption [2, 8, 23, 25] (i.e., whether users interact with the ciphertext before sending), with mixed results. Others have explored different degrees of integrating encryption with a user's existing email program [22]. We revisit these questions in the context of our prototype and compare to earlier results.

Key distribution and verification is central challenge of encrypted email [11, 26, 31]. Users of PGP-based email encryption are typically asked to manually import keys of intended recipients. Traditional solutions for PGP key management use key servers or other methods of publicly posting a key, whose authenticity is hard to verify [29]. The web-of-trust model [33], in which users sign each others' keys to support verification, has high manual overhead.

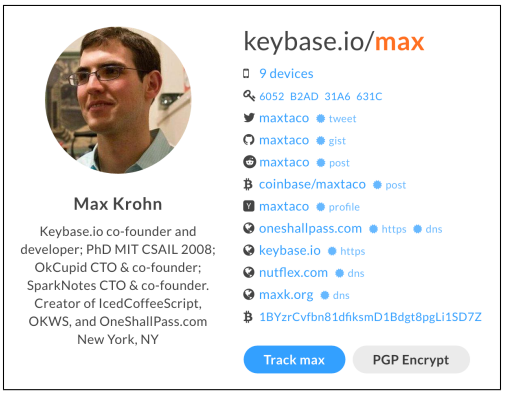

Figure 1. A Keybase profile, linked to the user's social media accounts.

Recent work in usable encryption attempts to make key management and verification increasingly transparent to users (e.g., [2, 8, 11]). Closest to our work, Atwater et al. [2] provide transparent key management by relying on a simulated version of Keybase. In Section 2.2 we discuss important differences between this simulation and Keybase.

Others have identified barriers to encrypted email adoption besides usability [12, 21]. We believe some of these barriers, such as social costs and the need for accurate mental models of security, would be reduced or eliminated by encrypted email that is as easy to use as regular email.

2.2. Keybase: Key Verification in a Social Age

Confidante uses Keybase (https://keybase.io), a new public key database and key trust protocol. Keybase users link public keys to their public social media identities by posting cryptographic proofs to their social media accounts (e.g., Twitter, Github). Proofs are two-way: signatures prove key ownership, and posting proves account ownership. Keybase's design ensures that attackers cannot impersonate users to substitute an attacker-controlled key without controlling *all* of the victim's linked social media accounts.

Client software can verify proofs independently, without trusting Keybase. Given proof verification by trusted software, users must check only that the social media accounts associated with a key belong to the person with whom they wish to communicate. This check is one that users already do naturally when communicating on social media.

Currently, Keybase keys can be searched by social media account, but not by email address, since email is not well suited for posting public proofs. This limitation poses challenges in our implementation of Confidante, and differs from Atwater et al.'s simulated version of Keybase [2], which retrieves keys by email address.

2.3. Target Users: Journalists and Lawyers

Most prior work on encrypted email usability has focused on non-specific user groups. (An exception is Gaw et al. [12], who studied adoption challenges among employees of an activist organization.) In our work, we instead focus on specific professions that regularly engage in potentially sensitive communications: journalists and lawyers. Our user study in Section 8 evaluates the design of Confidante in the context of their specific threat models, use cases, and professional constraints.

Both journalists and lawyers are bound by professional ethics to protect communications with sources or clients. Journalists consider source protection a high priority [16, 17], and lawyers are bound by rules of client confidentiality — a requirement whose implementation is being questioned in the face of the increasing use of third-party email providers like Google (e.g., [1]). Despite concrete reasons to secure their communications (e.g., [14]), usability and adoption challenges have prevented even these users from widely adopting encrypted email (e.g., [15, 16, 19]).

3. Goals and Threat Model

3.1. Goals

We aim to design a secure and usable encrypted email client. We chose email over other messaging systems because it is widely deployed, commonly used, and interoperable. Indeed, our user study (Section 8) shows that lawyers and journalists communicate via email with clients and sources on a more-than-daily basis. While designing novel communication channels is also important, it is critical that we secure email.

We consider the following *functionality and usability goals* for an encrypted email system:

1) **Usability:** The system should be easy to use, with a user experience like ordinary email.
2) **Interoperability:** Users should be able to use their existing email addresses and exchange encrypted email with someone using a different encryption client.
3) **Minimal Configuration:** The system should require little or no configuration (e.g., key management) beyond ordinary email.

We simultaneously aim to achieve these *security goals*:

1) **End-to-End Encryption:** The system should end-to-end encrypt emails between senders and recipients.
2) **No Trusted Server:** Passwords, keys, and plaintext emails should only be accessible on the client device. No external server should be able to access this information, reveal it in response to a subpoena, or have it accidentally compromised. (We do assume that clients receive legitimate binaries; see Section 3.2)
3) **Difficult to Make Security Errors:** It should be difficult or impossible to make security-critical errors.

We do not provide **metadata protection** as part of our design, since our goal is to remain compatible with email. In particular, we do not attempt to hide sender and recipient information. We discuss the impact on users in Section 9.3.

3.2. Threat Model

We assume that users wish to protect the contents of their emails from any entities other than themselves or the intended recipient. Possible adversaries include the email service provider (who stores and transmit emails), a government organization (who may subpoena or otherwise access data from the email provider, or who may eavesdrop on the network), or hackers (who may compromise a user's account, the network, or the email service provider).

We designed the Confidante email client primarily as a native desktop application. We assume users' computers are uncompromised, and that the server used to distribute Confidante is not compromised to serve users a malicious binary. Threats of this nature have been in the news recently, e.g., the FBI's requests of Apple [32], but this issue arises for any software distributed from third parties. We consider approaches to verify this binary out-of-scope but complementary. (Note that we also prototyped a web version of Confidante; we describe the security and usability tradeoffs with such a design in detail below.)

We use Keybase in Confidante to retrieve public keys. Keybase requires limited trust, since the cryptographic proofs linking a Keybase user's key to their social media accounts can be publicly verified, independently of Keybase itself. For Keybase, or any attacker, to impersonate a user, it would need to compromise *all* of the user's linked social media accounts. We assume that users may host their private keys on Keybase. These keys are passphrase-protected, so Keybase cannot directly access them, but this model may enable key compromise via phishing or similar attacks. We discuss these issues and alternatives in Section 9.

4. Design: Confidante

The above goals guide the design of Confidante, a PGP email client. Confidante uses Keybase for key management and sends and receives encrypted email through existing mail providers. Our implementation (Section 5) uses Gmail, but it could support any email provider.

4.1. Core Design: Confidante Desktop Client

4.1.1. Native Desktop App

We design Confidante as a multi-platform desktop email client with a user interface similar to conventional email clients. It supports all basic functionality expected of email clients, like sending mail and viewing threaded conversations. Confidante uses an external email provider to store messages and Keybase to manage cryptographic keys.

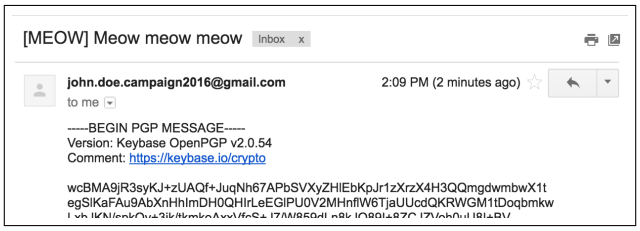

Figure 2. View of a message sent via Confidante, as viewed from the recipient's normal Gmail inbox.

PGP Encryption and Signing. Confidante supports end-to-end encryption and digital signatures that are backwards-compatible with ordinary email and PGP. This design omits certain desirable properties (e.g., forward secrecy or metadata protection) that would be possible with new or alternative protocols. However, it enables interoperability with existing PGP clients: emails encrypted and signed via Confidante can be decrypted and verified by any PGP software.

Dedicated Encrypted Email Client. We make a key design choice: only encrypted communications are possible within the Confidante client. Emails sent from Confidante are encrypted and signed locally, before being sent to the mail provider. Encrypted emails are retrieved from the mail provider and decrypted locally in the client. Emails sent through Confidante show up in other mail clients as encrypted messages — for example, Figure 2 shows the view of a message sent via Confidante as views from the recipient's normal Gmail inbox. Users can still use other mail clients (e.g., the ordinary Gmail web client) to send and receive unencrypted messages.

We chose to keep Confidante's interface separate from the user's normal email client, in contrast to integrated browser extensions like Mailvelope, to prevent errors like accidentally sending sensitive material in plaintext. This approach mimics apps like Signal's desktop client, which cannot send plaintext messages [28].

An additional benefit of a dedicated client, as opposed to a browser extension that modifies an existing webmail client, is that extensions must be designed separately for each mail provider and can break whenever the site updates.

Automatic Encryption. We convey the presence of encryption through user interface choices, without requiring users to manually encrypt or decrypt. For example, our "Send" button reads "Encrypt and Send" and decrypted emails in the client are shown with an option to view the encrypted version ("Show Encrypted"). This choice tries to balance the feeling of safety provided by interacting with ciphertexts with the convenience of automation. Prior work presented differing results comparing automatic versus manual encryption [2, 8, 23, 25], and we revisit this question in our user study (Section 8).

4.1.2. Integration with Keybase

Confidante leverages Keybase for public key discovery, private key storage, and key ownership verification.

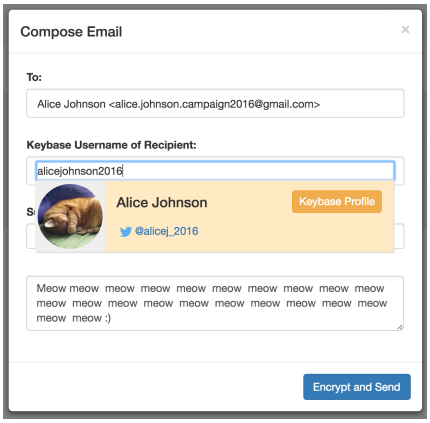

Figure 3. Selecting a Keybase user for whom to encrypt the email in Confidante's "compose" dialog.

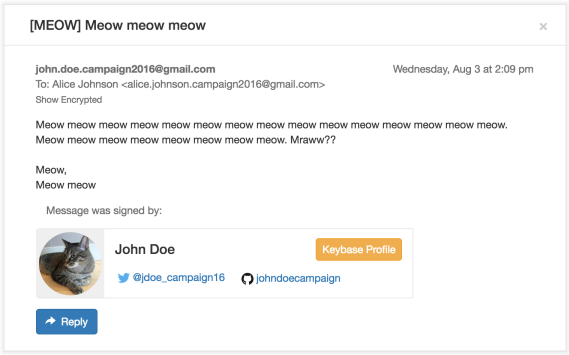

Figure 4. Confidante annotates signed email messages with a Keybase-populated user card.

Public Key Management. Confidante uses Keybase as a public key directory. To send email, users provide the Keybase usernames of recipients in addition to their email addresses. When the "Encrypt and Send" button is clicked, Confidante retrieves the public keys of those Keybase users and encrypts the message using all of those keys (as well as the sender's own public key).

As the user types in the Keybase recipient field, Confidante searches for Keybase users by real name, username, key fingerprints, or linked social media accounts, and uses autocomplete to ease filling in account names. Confidante shows matching Keybase profiles in a dropdown UI (see Figure 3), displaying the user's full name, profile picture, connected accounts, and a link to their Keybase profile. This UI is intended to help users verify that they are encrypting to the correct person by inspecting the linked social media accounts. Though not yet implemented in our prototype, a production version of the Confidante client should verify the public cryptographic proofs linking these accounts, and (for example) warn users of discrepancies, rather than directly trusting the information provided by the Keybase server.

Confidante also supports digitally signing messages. If the user receives a signed message, Confidante uses the key ID embedded in the signature to look up the Keybase user

who signed the message. Confidante retrieves the relevant public key from Keybase to verify the signature and displays the Keybase profile details of the signer at the bottom of the message (Figure 4).

Private Key Management. Confidante also relies on Keybase for private key storage. We allow users to store private keys, encrypted with their passphrase, on Keybase's servers. This design choice allows the app to work seamlessly across devices, as the key can be downloaded onto each device on demand. This design choice embodies a tradeoff of security for usability — a production version of Confidante can also support manual key management for advanced users to avoid storing any private key material on a third-party server, as we discuss in later sections.

When the user logs into their Keybase account in Confidante, the client downloads the user's password-protected key and decrypts it locally in the Confidante client. When users view encrypted messages, Confidante automatically decrypts the message using the private key cached in the client. When users send messages, Confidante automatically signs them using the cached private key as well.

4.2. Alternate Designs: Web and Mobile

Though our core design discussion of Confidante focuses on the desktop client model, it can be implemented as a web or mobile application as well. As we describe below (Section 5), our decision to implement Confidante entirely using web technologies allows us to easily port it to other platforms (e.g., to desktop using the Electron framework and to mobile using the React Native framework).

Indeed,we implemented an initial prototype that runs as a web application in the browser. It has the identical user experience as the desktop client and has several usability advantages related to portability: it does not require installation, the user interface can adapt to mobile devices, and it is accessible on any platform with a modern browser.

However, our choice to focus on the desktop client version results from the security tradeoffs we must make in the web version. The web application requires us to run a backend server to serve the page content, requiring users to place limited but non-zero trust in that server. In particular, we find that the server must proxy certain web requests to Gmail and Keybase, which are permitted on the desktop client but forbidden on the web client due to the browser's cross-origin resource sharing restrictions. As a result, our server has access to Gmail and Keybase authentication tokens, which the user must trust the server not to misuse or fail to protect. Though the consequences of this access are limited if all of a user's emails are encrypted — the server can access only the encrypted private key from Keybase and encrypted emails from Gmail — the desktop client version removes the need for an intermediate server entirely.

5. Implementation

The Confidante prototype consists of 4506 lines of JavaScript, excluding libraries, stylesheets, and HTML.

Multi-Platform Technology Stack. Confidante is a single page web application that runs as a native desktop application using GitHub's Electron framework (http://electron.atom.io/). The frontend is built with React.js, a popular UI library. We use Keybase's kbpgp.js cryptography library, a JavaScript-based implementation of PGP. The backend, a Node.js application, is responsible for retrieving data from Gmail and Keybase, and for persistent storage.

Our web technology-based approach allows us to deploy Confidante to multiple platforms while sharing substantial parts of the codebase. In the web application variant, the backend is run as a Node.js server, and serves HTML, JavaScript, CSS, and data to the browser. The frontend and backend can be run natively using Electron on the desktop, and React Native (https://facebook.github.io/react-native/) on mobile, which provide a JavaScript runtime. As discussed above, native applications provide a security benefit by removing the need for an intermediate Confidante server.

Email Provider. We use Gmail as the email provider in our initial implementation because it provides a convenient API, and we predict that many potential users would either use or have experience with Gmail. However, our design is agnostic to the email provider; we could create integrations for other email providers or general IMAP/SMTP support given enough engineering resources.

Authentication. Users log into Confidante by authenticating with Keybase and Gmail. Since Keybase doesn't currently offer OAuth, password salting and hashing is performed client-side, so that the user's plaintext credentials are never exposed in transit. Users authenticate with Gmail using OAuth, which provides Confidante with a temporary authentication token for accessing the user's emails. In the native application versions of Confidante, this authentication information is only accessible in the client-side software; no additional Confidante server is involved.

For users without Keybase accounts, we streamline the onboarding process by allowing users to sign up using Keybase's account creation APIs through Confidante.

6. Performance Evaluation

We benchmarked Confidante's decryption performance, which is important for adoption because users want to access their email quickly. These metrics were measured with the web version of Confidante on a MacBook Pro (Retina, 13-inch, Late 2013) with a 2.4 GHz Intel Core i5 processor and 8GB RAM, using Google Chrome version 50.0.2661.102. Each test was conducted with 30 trials.

We measured the time for the app to decrypt (1) a single message and (2) an inbox of 30 messages in parallel. These approximate the delay when reading a new message that appears in the inbox, and the delay before the inbox is available when launching the app.

Our single message was 828 bytes plaintext (2884 bytes armored ciphertext), and the mean time to decrypt and display the message was 1261ms (σ=171ms). We also measured the time to complete three sub-tasks: the time to per-

form the cryptographic operations in the browser (μ=699ms, σ=11 ms), the time to fetch a public key from Keybase for signature verification (μ=382ms, σ=134ms), and the time to fetch a Keybase profile to display the signature card (μ=180ms, σ=84ms).

Our 30 message inbox contained 6 messages from 5 different senders, with an average size per message of 1928 bytes in plaintext (3967 bytes ciphertext). The client decrypts all of the messages and fetches data from Keybase in parallel. The mean time to decrypt the whole inbox was 5214ms (σ=375ms).

For modern web applications, an initial 5 second delay for inbox decryption and a 1 second delay for subsequent messages is within the realm of acceptable response times [20]. Webmail clients like Gmail and Outlook load for several seconds before emails are accessible.

7. User Study

To evaluate our design decisions, we conducted an in-person between-subjects test with two encrypted email tools: Confidante and Mailvelope (https://www.mailvelope.com/), a browser extension that integrates PGP into Gmail.

We do not intend to assert or prove that Confidante is "better" than Mailvelope, or to compare them as complete products. Mailvelope has known usability issues [24], and Confidante is a research prototype. We study them for their differing design choices. For example, unlike Confidante, Mailvelope is integrated into Gmail, requires manual key import, and automates less of the encryption process. We picked Mailvelope over tools like Protonmail or Virtru to compare apples to apples: we are evaluating interoperable PGP email tools, as per our goals in Section 3.

We conducted our study with 15 subjects: 8 U.S. lawyers and 7 U.S. journalists. We chose participants from these user groups because they may engage in sensitive communications with client or sources, respectively. They may therefore be motivated to use email encryption tools or otherwise attempt to protect their communications. As we discuss below, we find that our focus on specific user groups teaches us *not only* about our design decisions but also *more generally* about the unique needs and use cases of those user populations. We explore their reactions to the design decisions of Confidante, Mailvelope, and Keybase, illustrating how these tools might meet or not meet these users' real threat models and operational constraints.

Our study is primarily qualitative, not quantitative, and we use a grounded theory based approach (see Section 7.2) to surface key themes that arise in our interviews [5].

Human Subjects / Ethics. Our user study was reviewed and approved by our organization's human subjects institutional review board (IRB), and we obtained informed consent from all subjects. Though we asked participants about potentially sensitive communications, we did not ask them to disclose sensitive or identifying information. We did not ask participants to use their own Gmail or Keybase accounts but provided fresh accounts for their use during the study. With each participant's permission, we audio recorded the interviews and recorded screen captures of their activities on the laptop we provided for the study. Consent to recording was not a requirement to participate in the study, though all participants did agree to be recorded. We stored these recordings only in encrypted form, and informed participants that we will delete these recordings upon request.

Recruitment. We targeted local (U.S.) journalists and lawyers in our recruitment. We sent recruitment blurbs via relevant mailing lists, posted on our social media accounts, and asked lawyers and journalists in our own professional networks to forward our recruitment blurbs. Participants received $30 Amazon gift cards.

7.1. Study Design

Our study took 60-90 minutes and consisted of three parts: a pre-task interview, a roleplaying task in which the participant used either Confidante or Mailvelope to send and receive encrypted emails, and a post-task interview. Two researchers participated in each study: one asked interview questions and responded to emails during the task, and the other took notes.

To minimize the effects of differences between the names of the tools, we called Confidante "Mailsafe" and served it from https://mailsafe.io/ for the purposes of the study. We conducted the study with an earlier web-based version of Confidante, whose UI is identical to the desktop application version described in Section 4.

Pre-Task Interview. To ground our evaluation of Confidante in the context of our participants' workflows, we first asked participants about their current email practices, threat models, and use or non-use of encryption or privacy enhancing tools. These questions (and related follow-ups) aimed to uncover operational constraints and threat models for the participants in their own work and for their profession as a whole, including:

- Frequency of email, who emails who, with which tools.
- Adversaries, threats and assets.
- Use/non-use of email security tools.
- Familiarity with technology/computer security.

Task. Before starting the task, we provided a brief overview of public key cryptography. This background helped participants assigned to Mailvelope, who needed to reason about using public and private keys. We explained it to participants assigned to Confidante as well to avoid introducing extra variables into the study. Next, we explained Keybase and walked them through a researcher's Keybase profile.

Participants were then asked to role-play a work-related scenario. Lawyers were asked to email opposing counsel to make an informal discovery request for documents related to a case. Journalists were asked to email a staffer on a presidential campaign. In both scenarios, the participant was instructed to send an encrypted email to their contact, whose information (name, email address, and Keybase account) was listed on a worksheet we provided.

The researchers responded to the emails from the participants in-character, as a lawyer or campaign staffer, using pre-written responses. The task was deliberately designed such that participants would need to go through the process of contacting a new person twice, so that we could observe participants as they learned to use the tool. Thus, when participants emailed the first contact, we responded with a message redirecting them to an alternate contact (providing their email address and Keybase account).

When participants made errors that interrupted the email exchange, like encrypting the email with the wrong public key, the researcher improvised responses in-character to help diagnose the participant's error, but did not provide detailed instructions on how to complete the task.

Post-Task Interview. After the task, we asked participants about what they understood about the tool they had used, the underlying security it provided, and their feelings on its utility and usability, including questions about:

- When, how often, and with whom they would use such a tool.
- The security provided by the tool, and the role of Keybase and keys in its operation.
- How safe they would feel using the tool.
- How they felt about Keybase.

Participant Response Bias. Participant response bias in user studies is a known issue. Participants are more likely to prefer or praise technological artifacts which they believe were created by the researcher [7]. To mitigate this, we described each tool in the third person, not informing participants who had created the tool until the end of the study. Additionally, in the results below, we avoid reporting vague, positive results about whether participants liked the tools or Keybase, discussing only concrete statements in which they explained how they might use such tools.

7.2. Analysis

Our interview analysis follows a qualitative, grounded theory based approach [5]. Specifically, we analyzed our interviews through an iterative coding process, in which we identified key themes, or *codes*.

First, all three researchers independently analyzed the interviews, each making a broad list of topics which participants brought up. Then we discussed the list to narrow it down and combine closely related themes. We formed a consensus on themes, or codes, of interest. We iteratively revisited the coding manual and interviews until no new codes were added and all interviews were coded. Each interview was independently coded by two researchers. If the two researchers disagreed, they discussed in person and, where possible, reached consensus.

We report Cohen's kappa (κ) as a measure of inter-coder agreement [6]: the average kappa for all results in the paper is 0.97. Fleiss et al. rate values of kappa over 0.75 as excellent agreement and between 0.40 and 0.75 as intermediate to good agreement [10].

ID	Profession	Gender	PGP Exp.	Tool
L1-C	Lawyer	M	1	Confidante
L2-MV	Lawyer	M	2	Mailvelope
L3-MV	Lawyer	M	1	Mailvelope
L4-C	Lawyer	F	1	Confidante
L5-C	Lawyer	M	3	Confidante
L6-C	Lawyer	M	5	Confidante
L7-MV	Lawyer	F	2-3	Mailvelope
L8-MV	Lawyer	M	3	Mailvelope
J1-MV	Journalist	F	2	Mailvelope
J2-C	Journalist	M	1	Confidante
J3-C	Journalist*	M	1	Confidante
J4-MV	Journalist*	M	1	Mailvelope
J5-C	Journalist*	M	1	Confidante
J6-C	Journalist	M	1	Confidante
J7-C	Journalist	M	1	Confidante

Figure 5. Summary of participants and their assignments to experimental conditions. PGP experience was self-reported on a scale of 1 (low) to 5 (high). Journalists marked with an asterisk are students. We encode tool assignment as part of participant ID (e.g., L1-C is a lawyer assigned to Confidante). All participants work in the United States.

7.3. Participants

We interviewed a total of 15 participants, including 8 U.S. lawyers and 7 U.S. journalists (3 of them students). These participants are summarized in Figure 5. Participants ranged in experience in their field from less than a year to 38 years. The lawyers work in diverse practice areas, including criminal defense, immigration, employment, business, estate planning, health, and family law. Some are in solo practice, some in small practices, and some in large law firms. The journalists cover a variety of issues as well, including politics and government, the environment, health, the economy, sports, and culture for different news organizations (one radio, one print, and two online).

Experimental Condition Assignment. We assigned participants to one of two experimental conditions: Confidante or Mailvelope. The study involved a different fictional scenario for lawyers and for journalists, but required participants to complete the same set of tasks. Figure 5 shows how participants were assigned to conditions. Following accepted practices in qualitative methods [5], we stopped assigning the Mailvelope condition after we felt that we had saturated on its themes (i.e., we were not learning anything new). We assigned participants such that degree of prior PGP experience was evenly split across both conditions. No participants reported having previously used Mailvelope.

8. User Study Results

We describe our results in two parts: first (Section 8.1), we focus specifically on evaluating Confidante. Second (Section 8.2), we discuss our broader findings about the security and operational needs of journalists and lawyers. In Section 9, we then step back and reflect on lessons and recommendations from these findings.

Linear Model Coefficients

Variable	Estimate	Std. Error	p-value
(Intercept)	2.11	1.08	0.06
2nd Email	-1.17	1.53	0.45
Mailvelope	4.92	1.65	<0.01**
2nd Email*Mailvelope	-2.77	2.34	0.25

Significance Codes: *** p<0.001 ** p<0.01 * p<0.05 . p<0.1

Figure 6. Regression results for time spent on key management. The variables correspond to one level of a study condition: which of the two emails is being sent, the tool used, and interaction effects respectively. We find that Confidante users spend significantly less time on key management.

8.1. Results: Evaluating Confidante

We begin with an evaluation of Confidante itself, including Keybase's role, based on the encrypted email task component of the study and participants' subsequent responses about their impressions. We find that participants who used Confidante completed the email encryption task more quickly (Section 8.1.1) and made fewer errors (Section 8.1.2). We explore their security and usability perceptions of Confidante in Sections 8.1.3 and 8.1.4, finding that participants generally find Confidante as easy to use as ordinary email but identifying remaining security and usability challenges.

8.1.1. Timing

To quantitatively estimate the usability of each tool, we measured how long it took for participants to accomplish the task. Overall, we found that participants assigned to Confidante spent less time (μ=17 min) than those assigned to Mailvelope (μ=32 min), (t_9=-2.45, p<0.04, 95% CI [1.18, 29.43] min faster with Confidante).

We hypothesized that Confidante users completed the task faster in part because they spent less time on key management. We thus measured time spent on key management, twice for each participant: when sending the initial email to the provided contact, and when sending the first email after being redirected to another contact. For Confidante, we counted time spent typing in the Keybase username into the compose window, interacting with the Keybase autocomplete UI, and viewing Keybase profiles. For Mailvelope, we counted time spent importing keys, choosing recipients to encrypt for, and sharing public keys.

We performed a linear mixed effects analysis of the relationship between the tool and time spent on key management (see Figure 6). We treat participant as a random effect, and the tool and email as fixed effects. We also investigated interactions between tool and email, to see if one tool was more prone to learning effects than the other.

We found that tool had a main effect on time spent on key management, particularly that *participants who used Confidante spent less time on key management* ($p < 0.01$, 95% CI [1.51, 8.32] min faster with Confidante). We found no significant learning effect between the first and the second emails sent ($p = 0.45$), no significant interactions between the first/second email and tool ($p = 0.25$).

8.1.2. Errors

Participants made a variety of errors while using both Confidante and Mailvelope. We classify these errors:

- Some of these errors have been made **impossible** *by design* in Mailvelope or Confidante.
- Some of these errors are **irreversible** (e.g., they involve sending email), while some merely cause frustration and delay.
- Some of these errors are security **critical**: they may violate confidentiality, authenticity, etc.

We list the errors made in Figure 7. We found that Confidante renders many of the irreversible and critical errors which were common among Mailvelope users impossible, significantly reducing the chance that users will make serious security and privacy mistakes. These errors were mostly related to key management and confidentiality, such as leaking plaintexts to Gmail, and encrypting emails to the wrong key. As we discuss further in Section 9, we thus find a positive outcome to our experiment of leveraging Keybase to automate key management in Confidante.

On the other hand, the complexity of Confidante's relationships with Gmail and Keybase was confusing for some users. Participants often mixed up the roles of Confidante and Keybase, and they often hesitated about giving Confidante OAuth permissions to access Gmail (required for Confidante to act as a mail client). Two participants also misunderstood Confidante as a closed ecosystem (e.g., like WhatsApp) rather than a layer atop email.

One critical error possible in Confidante that we did not observe was sending mail encrypted to the wrong key. Confidante users must still select the correct Keybase username for their recipient unless the email is a reply, and errors could result in encrypting to the wrong key. Though none of our participants made this mistake, the likelihood of such an error may be higher in a real-world context where a user has many contacts, and two Keybase users may have similar usernames (maliciously or coincidentally). Indeed, 8/15 (6 lawyers, 2 journalists) participants said they would like the mapping from email to Keybase username automated, including asking for an "encrypted contact list" (L1-C). Thus, identifying the right recipient remains a challenge with Keybase, even after all other key management operations are automated away. We discuss this challenge further in Section 9.

8.1.3. Security Perception

We evaluate how participants perceived Confidante's security properties, highlighting errors in their perception as well as legitimate concerns they raise.

Awareness of Encryption. Most participants (13/15) understood that the tool they used encrypted email such that eavesdroppers could not read it. Sometimes, however, nuances or implications were lost. For example, one participant who considered third-party mail providers an adversary did not recognize that Confidante helped defend against that adversary, asking:

Error	Occurrences Mailvelope	Confidante	Irreversible	Critical
Key management				
Difficulty importing public keys	6/6	3/9		
Forgetting to share own public key	6/6	Impossible	✓	
Difficulty exporting own key	5/6	Impossible		
Exposing private key	0/6	Impossible	✓	✓
Confidentiality				
Leaking plaintext to Gmail	5/6	Impossible	✓	✓
Sending mail unencrypted	3/6	Impossible	✓	✓
Sending mail encrypted to wrong key	1/6	0/9	✓	✓
Mental Model/UI				
Not encrypting mail to self	2/6	Impossible	✓	
Encrypting mail *only* to self	3/6	Impossible	✓	
Mixed up tools/sites	3/6	6/9		
Invited existing users to Confidante	Impossible	2/9	✓	
Unaware that messages were encrypted	0/6	2/9		
Tried to encrypt via Keybase instead of tool	1/6	2/9		
Worried by or denied Google OAuth dialog	Impossible	4/9		
Mixed up public keys vs Keybase usernames	Impossible	2/9		

Figure 7. *Irreversible* errors cannot be undone, though they may be harmless. *Critical* errors may have security or privacy consequences, e.g. loss of confidentiality, authenticity, etc. For Confidante, "Difficulty importing public keys" refers to confusion over looking up Keybase usernames.

> So does [Confidante] take Gmail's ability to eavesdrop out of it, or do you still have to allow Gmail to eavesdrop? (L4-C)

Two users of Confidante were initially unaware that their emails were encrypted at all. These participants had not seen the ciphertext of the email, which could be made visible by clicking the "Show Encrypted" option, or viewing the message in Gmail. At the end of the study, we pointed out this feature to them, and they expressed greater certainty that the mail was encrypted. This result suggests that testing UI design for communicating security properties is important, and echoes findings from Ruoti et al. [22, 25] (but not confirmed in other studies [2, 8, 23]) that manual encryption has some benefits over automated encryption. We discuss further in Section 9.

Concerns from Technical Participants. For technical users, the fact that Confidante was so easy to use and cryptographic operations were automatic and hidden sometimes lead to suspicion. For example:

> It almost feels like... because this is so easy, and I know it's PGP, it really feels like there must be something wrong... [PGP is] a rite of passage. (L5-C)

Participants with GPG and security backgrounds had other technical concerns about Confidante, e.g., uploading private keys to Keybase and the lack of frequent passphrase prompt before signing or decrypting mail. These concerns could be eliminated at the expense of usability, an option that may be appropriate for these users. More generally, one technical user said he would only feel comfortable using Confidante once it passed an independent security audit.

Concerns about Encryption. While technical participants tended to focus their concerns on more nuanced issues, several non-technical participants were worried that encryption itself might be easily broken, due to encryption that is "not very good" (L1-C) or by strong adversaries:

> I would have to see some sort of narrative about why [encryption] works... that there's nobody at the NSA who is able to decrypt this... The fear that Mr. Snowden et al. put into us about what government capability really is, about what corporate capability is. (J1-MV)

Others felt vaguely uncertain about what security guarantees encryption provides in the first place:

> [Despite encryption,] I would still be concerned that there's some kind of software that would be able to divulge the details of that message while it's being sent. (J3-C)

Concerns about Drawing Suspicion. Some participants (4/15) expressed a worry that sending encrypted email or even installing encrypted email tools makes the user more of a target.

> [A] lot of [sources]... would say "Well, what's the purpose of this? Why are you making me do this? Is this actually going make it more likely for this to raise a red flag with my employer?" (J6-C)

This concern may prevent even users who need to protect their communications from using tools like email encryption. Wider adoption of these tools even for non-sensitive communications can help reduce this concern.

Concerns about Metadata. 2/8 lawyers and 1/7 journalists proactively observed that the tool they used didn't encrypt metadata, and only two participants explicitly mentioned metadata as an asset. For others, metadata may not have been important, or they may have been unaware of its risks.

Concerns about Contact Authenticity. Some participants were concerned about the authenticity of contacts. For example, L4-C mentioned a personal policy of only using contact information provided directly by a client (rather than

searching online). That participant indeed clicked through from Confidante's autocomplete dialog to inspect a recipient's Keybase account. More generally, as discussed above, 6 participants expressed a concern about the authenticity of a recipient's Keybase account or about selecting the wrong Keybase user from the autocomplete dialog.

8.1.4. Usability Perception

We examine how usable participants found the tool they tried. To mitigate participant response bias, we do not report on generic positive feedback but rather focus on *specific* responses participants gave about how and why they would (or would not) use a tool.

Easy to Use. Participants expressed some concrete positive feedback about Confidante. This feedback often (in 4 of 9 Confidante experiments) took the form of likening it to a normal, unencrypted email experience, or of comparing it favorably with other PGP-based tools. For example, L5-C called it "the easiest PGP experience I've ever had." The relative simplicity of Confidante also encouraged that participant that their clients might be able to use it:

> I could see, in a way that you never could with PGP before, being able to send a one-page instructional thing on how to set this up, and trust that they could actually do it themselves. [With Thunderbird,] you would send that off and then get on the phone with them for an hour and a half and walk them through various things. (L5-C)

Usability Challenges. Participants explicitly noted some usability challenges with Confidante (in addition to the errors/confusion we observed during the task, discussed above). Most frequently, participants found it frustrating that Confidante required that they enter both an email address and a Keybase username when sending an email — 4 of 9 expressed a desire for more automated email-to-Keybase mappings, as discussed above. Usability challenges also manifested in some participants' confusion between Confidante, Gmail, and Keybase.

Integrated vs. Standalone Client. Not all participants liked the idea of a standalone client, asking explicitly for integration with Gmail (2 of 9 who used Confidante). However, direct integration with Gmail was not universally preferred: in both experimental conditions, some participants preferred an integrated client and some preferred a standalone client (contrary to some prior findings [2, 22]). For example, one participant liked the standalone client because it removed the need to make a decision about encryption:

> One of the things that's great about [Confidante] is that I didn't have to make a decision about whether or not I should encrypt this. Having something that's easy to have always be encrypted, where you don't have to decide if this is [*gesturing air-quotes*] "worth the hassle." (L4-C)

Selective Use. Most participants (10/15) said they would use encrypted email for only their most sensitive communications, not for all emails, suggesting that despite improvements in usability, people still have reasons to not encrypt their emails, usability or otherwise. This finding echoes results from prior work, which identified social and other issues hindering the adoption of encrypted email [12].

One-Time Startup Cost. Despite these issues (and despite issues with Mailvelope), some participants (3/15) expressed that one-time startup costs of teaching a communication partner to use a tool were acceptable if the benefits provided by the tool are worthwhile. One journalist recalled the early days of email:

> There was a time when people... couldn't even understand the architecture of an email address. I would spend so much time with people on the phone, like "No, no spaces. No, no punctuation at all. And then you have '@', you know the 'a' with the funny curl around it." ...People would just be like "this is terrible," our brains would be burning, our neural circuits would be on fire. So that's all [key exchange] is, it's just a new layer... I remember people thinking this was a grotesque inconvenience to have to email people, but it was not long before all of that evaporated. (J1-MV)

8.1.5. Reactions to Keybase

To use Confidante, users must sign up for Keybase. Since Keybase is separate from Confidante, we separately assess participants' reactions to it. Overall, participants were largely ambivalent about Keybase. Many said they would sign up, but the risk of participant response bias makes this result inconclusive. Instead, we highlight strong positive or negative reactions towards Keybase.

Positive Reactions. A few participants had strong positive reactions to Keybase. For example, one lawyer said:

> If something like this caught on, I could see putting my Keybase on my business card, or putting it in the signature line of my email. (L5-C)

One journalist was excited about it because she felt that "it says something... about the reporter's style" (J1-MV), because it shows that they take source protection seriously. That same journalist was also enthusiastic about linking her social media accounts with Keybase, saying she would "pimp my LinkedIn, pimp my Twitter" (J1-MV).

Negative Reactions. Some participants had negative reactions to Keybase. Two lawyers with prior PGP experience worried about storing (passphrase-protected) private keys on Keybase. One also said that he wouldn't use Keybase because he wanted one key pair per client to mitigate the risks of a single private key being compromised. Other participants were unenthusiastic or reluctant about using Keybase, sometimes citing a lack of social media accounts to link with it, or an unwillingness to link existing accounts. As with other tools, participants worried that Keybase is an extra barrier to bootstrapping communications.

8.2. Results: Security Concerns and Operational Constraints of Laywers and Journalists

By conducting our user study with participants from specific user populations, our findings include not only the evaluation of Confidante's design choices, but also shed light on the security needs and operational constraints of these users groups more generally. These findings have implications for researchers and tool designers (Section 9).

8.2.1. Email & Security Practices

Each lawyer and journalist we talked to uses email on a daily basis. Threads with existing clients are often initiated by lawyers to push new information regarding cases to clients. Exchanges with potential clients are always initiated by clients, due to rules against marketing and solicitation of legal services. Journalists usually send the first email to sources to ask for information.

All participants said that they regularly read and respond to emails from their phones. We conclude that encrypted email solutions must work on mobile devices to be practical for day-to-day use. As a standalone client, Confidante can meet this requirement; integrated tools (i.e., browser extensions) like Mailvelope do not.

Existing Security Related Practices. Many participants said that they took steps to secure their computers, phones, and communications, such as password-protecting devices and accounts, and using good passwords and secure WiFi. These practices often related to the security of devices.

In terms of secure messaging, most participants had little experience. Three lawyers had significant prior experience with GPG, while one had tried using GPG once. One (L5-C) has used Ricochet and Signal with some clients. One journalist (J1-MV) has used Signal in the past with sources. Several journalists said they often switch from email to phone calls for sensitive communications, echoing findings from prior work on journalists [16]. Thus, the motivation to secure communications is clearly there.

8.2.2. Operational Constraints

While lawyers and journalists both have an interest in protecting their email communications through encryption, we find that they have other professional obligations and constraints that may limit the ways they can use encrypted email tools or the features they require of those tools.

Time is Money. Some (2/9) lawyers voiced concerns that encrypting email takes longer. One participant (L8-MV) said the extra time could reduce the hours of billable work they could accomplish in a day, hurting their firm's income. Another (L7-MV) noted that clients care about time delays since they pay hourly rates for legal services.

Onboarding Burden and Scaring People Off. Some lawyers worried that asking clients to encrypt email could make it more difficult to establish relationships with new clients. One participant described:

> Having a conversation with a client, especially your first few conversations, can be really tenuous things... it's easy to scare them off. Having a conversation about surveillance and encryption, as one of the first things that you do, is a really really precarious thing. (L5-C)

Some journalists had similar concerns about scaring off sources by asking them to use encryption, echoing earlier findings [16]. However, other journalists thought encryption could *benefit* relationships. For example:

> I don't see that it's an obstacle at all. In fact it's finally a way for journalists to show that they care about the consequences of the work on individuals. (J1-MV)

Both lawyers and journalists said clients or sources often dictate communication methods (again echoing prior findings on journalists [16]). For example, several lawyers mentioned using encrypted document depot services like Clio in response to requests from specific clients. A particularly technically savvy lawyer said that they offered (but did not require) encrypted chat tools such as Signal and Ricochet to their clients.

A difference we observed (and which we discuss further in Section 9) is that lawyers typically do not initiate conversations with new clients (due to rules about solicitation), while journalists may contact new sources. And because clients need the lawyer's services, while sources may benefit less from speaking with a journalist, lawyers may feel more comfortable making requests of clients. An example that involved *not* using encryption despite a client's request:

> I've had a handful of clients ask about [encryption]. ... [One client] at the beginning of the relationship sort of suggested it, and I said look, you know, I'd prefer just not to do it. And so we ended up not doing it. (L8-MV)

Searchability. Four out of 9 lawyers explicitly mentioned the need to search past email, including encrypted email, particularly when legally obligated to produce emails during the discovery process of legal proceedings. Without this feature, encrypted mail may impede the legal process.

Sending Documents. Five out of 9 lawyers mentioned often sending attachments containing sensitive data such as personally identifying information and financial records. One lawyer said sending large collections of documents by delivering hard drives was common at their firm (L1-C). Thus, email encryption tools for these users should support usable attachment encryption as well.

8.2.3. Participant Threat Models

Perceived Adversaries, Threats, and Vulnerabilities. Understanding the threat models of users is important to designing tools that protect against relevant threats. Figures 8 and 9 list adversaries, threats, and vulnerabilities mentioned by participants. Participants identified many different threats and adversaries, but there was no consensus on any particular concern. The range of concerns include some that

Lawyers	Journalists	Adversary Mentioned
4/8	0/7	3rd Party Service Provider
0/8	1/7	Competing News Organizations
3/8	5/7	Government/Law Enforcement
1/8	0/7	Industrial Espionage
1/8	0/7	Judge
5/8	2/7	Random Hacker
0/8	5/7	Source's Employer

Figure 8. Adversaries identified by participants.

Lawyers	Journalists	Threat/Vulnerability Mentioned
4/8	2/7	Account Hacking
3/8	0/7	Client/Source Indiscretions/Errors
1/8	0/7	Discovery Requests
2/8	3/7	Encryption may be Broken
3/8	2/7	Government Surveillance
2/8	0/7	Insecure Internal Networks
1/8	0/7	Malware/Ransomware
2/8	1/7	Physical Intrusion
3/8	1/7	Plaintext Held by Mail Provider
3/8	0/7	Spoofing Email
1/8	1/7	Subpoena/Legal Request

Figure 9. Threats identified by participants.

security experts might not share, such as a lack of trust in encryption primitives (as discussed above), or concern over threats likely unrelated to email.

Nine of 15 participants expressed strong concerns about the security of endpoints (e.g., with malware and physical intrusion), while only 4 were concerned with mail provider access to plaintext, 5 with government surveillance, and 6 with account compromise.

Few participants mentioned a concern over email content authenticity (as might be addressed by digital signatures in addition to encryption). Only 3/8 lawyers and 0/7 journalists mentioned that emails could be spoofed. One lawyer said:

> [I]f someone picks up my client's phone and sends me an email, we trust that it's actually them... I definitely have received emails from clients where it's pretty clear based on what I know of their writing style... from previous encounters... that that's not... the way that they write, so their girlfriend wrote it or something. (L4-C)

Our participants were not security experts, and so their self-reported threat models may not be complete or accurate. Nevertheless, the threats these users *perceive* may influence their tool choices, and understanding these perceptions can help guide UI design or other interventions.

Assets. Unlike perceived threat models, which may be inaccurate or incomplete, users may be the best judges of their own assets, i.e., what they consider important and worth protecting. Figure 10 shows the list of assets which we coded from participant responses. Lawyers and journalists voiced concerns about different types of assets. Lawyers often framed their interest in encryption relative to the ethical standards and legal requirements which govern their behavior and obligations, such as attorney-client privilege (6/8), work product (2/8), malpractice/liability (2/8), and

Lawyers	Journalists	Asset Mentioned
6/8	–	**Attorney-Client Privilege**
4/8	3/7	Client/Source's Data/Information
3/8	1/7	Embarrassing Content
2/8	1/7	Liability
1/8	1/7	Metadata
2/8	0/7	Personally-Identifiable Information
1/8	**6/7**	**Protecting Clients/Sources**
1/8	3/7	Relationship with Client/Source
0/8	1/7	Reputation of Self/Employer
2/8	0/7	Work Product

Figure 10. Assets identified by participants. Bold rows were reported most often by one group.

their obligation to protect clients' trade secrets and other intellectual property (4/8).

Several lawyers explained that they are required to take "reasonable steps" to protect communications subject to attorney-client privilege. Each had different interpretations of what precautions constituted "reasonable steps" for email. Some (e.g., L2-MV) expressed worries that storing plaintext emails on third party email provider servers "eroded" the attorney-client privilege of those messages, since they were available to a third party.

Some lawyers noted that there are limits to their responsibility. For example, they would not be liable if clients exposed privileged information through poor security practices or intentionally shared privileged information.

> As long as I'm taking reasonable steps to protect the confidentiality where I can, I'm liability-wise okay... It's [the client's] right to waive, ...so if they want to throw my emails around to anyone they want, ...that's their business. (L4-C)

By contrast, journalists tended to frame their concerns in terms of protecting sources from the consequences of divulging information. One journalist said, "I really don't believe in doing stories in which sources get hurt" (J1-MV). Of course, journalists are also motivated by practical concerns; another journalist said:

> If we can't maintain confidentiality... then sources will dry up. This has already started to happen in the Obama administration... it's hard to develop sources, it's harder than ever... It's an existential threat to investigative reporting[.] (J2-C)

These differences between journalists and lawyers suggest that different tools with different security properties may be appropriate, as we discuss further below.

9. Discussion: Lessons and Recommendations

9.1. Reflecting on Design Decisions

Automated Key Management. Confidante's automated key management saved time and reduced confusion for participants. While Mailvelope participants made frequent key management errors (e.g., forgetting to share their own public key, impossible in Confidante), Confidante participants completed the study task more quickly, spending less time on

key management (Section 8.1.1), and with fewer errors (Section 8.1.2). This result echoes prior work [8, 11, 12, 25, 26] and suggests that automated key management is critical for usability. Indeed, four (of 9) participants who used Confidante made comments such as: "It's no different to use than just using Gmail directly" (L4-C). *In other words, with automatic key management, the act of using encrypted email became essentially the same as sending ordinary email.* Confidante combines this ease of use with convenient security, since keys can be verified by Keybase social media proofs instead of out-of-band communication.

Automatic Encryption. In addition to streamlining key management, Confidante automates encryption and decryption. By contrast, in Mailvelope, users compose messages in a separate dialog and transfer the ciphertext into Gmail's compose dialog. Though Confidante's transparent encryption contributes to it feeling "just like regular email," it also had several problematic consequences.

For expert users, the transparency sometimes resulted in a lack of trust (recall L5-C from Section 8.1.3: "because this is so easy... there must be something wrong"). For non-experts, the transparency contributed to — or did not help users overcome — confusion about how the tool worked. For example, two Confidante users seemed stumped when asked to send an encrypted email to the second character in the role-playing scenario: they did not realize the first message they sent had already been encrypted. Confidante has a button that shows the ciphertext of messages, but no participants noticed that button without prompting.

These results highlight a remaining challenge: balancing automation and trust. Prior work explored this tradeoff with differing results [2, 8, 23, 25]; our findings suggest that a tradeoff may indeed exist. Encrypted email must balance transparency — approaching the feeling of a "normal," unencrypted email experience — with communicating "under-the-hood" functionality and security properties to users. Balancing these enables usability *and* trust while preventing users from making mistakes due to incorrect mental models. Addressing this challenge (e.g., through careful UI design) may require a deep understanding of the security needs and mental models of target user groups.

Standalone Email Client. We built Confidante as a separate app, providing a view of only the encrypted messages in a user's inbox. Conversely, Mailvelope is a browser extension that integrates directly with Gmail. Prior work [22] found that integrated solutions were preferred, but in our case, we did not find conclusive preferences in either direction. Most participants did not have strong opinions one way or the other, instead voicing various pros and cons. Benefits of a standalone client included the inability to make certain mistakes (e.g., accidentally sending a plaintext email, which 3/6 of our Mailvelope participants did) and mobile compatibility. Section 8.1.4 included a quote from L4-C, who liked the standalone client because it removed the decision about whether encrypting any given email was "worth the hassle."

However, two participants who used Confidante explicitly wanted Gmail integration, and the fact that most participants said that they would send encrypted emails only under certain circumstances suggests that an integrated solution may be more appropriate (at the possible risk of mistakes).

9.2. Keybase: Opportunities and Challenges

Overall, our experiment integrating Keybase as the key management portion of Confidante is a success: participants spent less time on key management and generally found the Confidante user experience similar to regular email.

Keybase innovates on key management: unlike traditional key servers, it enables usable verification of public keys without out-of-band steps. Their solution is only recently possible, with the rise of social media and the frequent use of real names and relatively trusted identities online. Services like Keybase raise the ceiling of usability for encrypted email. *Independent of any other design decisions, we recommend that encrypted email systems incorporate a key management solution like Keybase.* For example, if Mailvelope used Keybase for key management, many of the errors made by our participants would be eliminated.

However, some challenges with Keybase remain:

Challenge: Private Key Management. Storing passphrase-protected private keys on Keybase allows Confidante to simplify private key management and easily support multiple devices. However, this design choice puts private keys at risk if passphrases are compromised, e.g. by weak password choices, reuse, or phishing. Participants often confused their Gmail and Keybase accounts, making it easy to imagine people entering their passphrases in the wrong place. Security-conscious participants expressed wariness about storing their private keys on Keybase.

Keybase does not *require* users to upload private keys, and Confidante could support local private keys instead. Such a feature would push some key management burden back to users. Designers should determine whether this tradeoff is worth it for their target user populations. Alternate multi-device key management approaches may also be feasible.

Challenge: Identifying the Correct Recipient. Though Keybase aids identify verification by allowing users to identify each other via social media links, the risk remains that a user may select the incorrect Keybase username to whom to encrypt. Though no participants using Confidante actually made this mistake, several were concerned that they might do so in the future and asked for an email-to-Keybase mapping. As described in Section 2, securely mapping email addresses to Keybase usernames is challenging, since email addresses cannot be used to post public proofs as on other social media accounts. However, our study results suggest that such a mapping is important for usability and security. A possible solution that can be implemented within Confidante is the trust-on-first-use model, in which email addresses are associated with Keybase accounts in the user's contact list. Other solutions may be possible through collaborations between Keybase and email providers — for example, if

email providers allowed users to post public Keybase proofs associated with their email accounts.

Challenge: Social Linking. Keybase does not require social linking, but without it key verification is no better than with key servers. Some users will be uncomfortable doing so, and it may be inappropriate for others. For example, one participant said it might look suspicious for journalistic sources to have a Keybase account, and not all users know each others' online identities. Developing other techniques for key trust remains valuable and necessary.

9.3. Designing for Specific User Groups

Beyond lessons we learned about Confidante's specific design choices, our study with targeted user groups sheds light on the broader contexts in which these users work and communicate. An implication of the differences we find between journalists and lawyers is that no single design is likely to suffice for all users. Instead, we argue that tools may be more successful — both for security and usability — if they are tailored to the needs of particular groups.

Below, we highlight important differences we identified among the journalists and lawyers we interviewed.

Searchability and Access by IT Staff. Searching email is a standard feature, and enterprise IT staff typically need access to employee emails. Several lawyers at large law firms told us that the requirement that their IT staff be able to search their emails went beyond HR concerns: it is required for their legal obligation to fulfill discovery requests. Supporting search and IT access may be possible through the use of searchable encryption schemes (e.g., [3, 27]) and/or shared organizational keys, though these features naturally involve design tradeoffs. *If these types of operational constraints are not included in a tool's design requirements, the tool may be irrelevant or unusable for groups that need it.*

Legal Protections May Suffice. Lawyers in our study often did not need to protect against the strongest adversaries or threats. Rather, their goal was to take reasonable steps that would protect attorney-client privilege from the perspective of the U.S. legal system [9]. One lawyer explained:

> Attorney-client privilege is one of the sacrosanct things we deal with. If you are careless about letting a third party view or hear a privileged conversation, it will defeat the privilege. ... For example, if I have a document that's a privileged document, if somebody breaks into my office and looks at it, that doesn't defeat the privilege. But if I leave it out where somebody walking by can see it, that could. So you'd have to take reasonable precautions. (L1-C)

It is in this context that U.S. legal organizations are considering encryption as a best practice (e.g., when storing emails on third-party servers like Gmail) [1]. To serve this purpose, email encryption must not meet the theoretically highest possible security standards to count as a "reasonable" precaution. In other words, *some groups may be able to make security tradeoffs others cannot, replacing technical protections with legal ones*. In this example, encryption may act in a supporting role to the legal defense provided by attorney-client privilege.

Lack of Metadata Protection in Encrypted Email. An important weakness of encrypted email is that metadata about communication patterns is not protected.

Journalists noted that sources must sometimes hide even the fact that they talk with journalists, even if the content of the communication is protected. In other words, for some groups, metadata may be *more* important than content. For these groups, encrypted email may be unhelpful or worse, if it provides a false sense of security.

On the other hand, lawyers said they must sometimes divulge communications metadata in a "privilege log" even when the contents are protected by attorney-client privilege. That is, U.S. law may not permit metadata to be hidden, even if technical protocols can protected it. Similarly, in the U.S., the fact that a person contacts a lawyer with whom they have a preexisting relationship cannot be used against them in court, reducing the importance of metadata protection.

Together these findings suggest that *complex social and legal factors may determine that certain tools (such as encrypted email) may be a good fit for certain groups (e.g., lawyers with established clients) but an inappropriate choice for others (e.g., journalists with anonymous sources).*

Differences in Relationships with Recipients. Lawyers and journalists differed in practices around first contact. Journalists frequently said that they may reach out to new sources, while lawyers typically do not reach out to new clients due to rules against the solicitation of legal services. Onboarding procedures and features such as the ability to send encrypted mail to people who do not yet use the tool should be designed with these differences in mind.

Both journalists and lawyers voiced concerns about scaring away sources and clients with frustrating or complicated email encryption tools, and both said the choice of which tools to use or not use tends to lie with clients and sources. However, the situation may differ significantly once a client is established with a lawyer. At this point, the relationship may be much less tenuous than with a journalistic source (who is not receiving valuable legal services from the journalist), and lawyers may be more willing to insist on the use or non-use of certain tools.

Users are Experts on Their Assets. Participants had little consensus about threats to their communications (Section 8.2.3). Computer security experts may be better able to identify and compare threats. However, we observe that users are best positioned to educate security practitioners about *what* they value (i.e., their assets) and the *systems and rules* surrounding those assets (e.g., attorney-client privilege), which may be critical to the design of systems.

In summary, we argue that security researchers cannot build effective tools in isolation. Understanding the needs and threat models of user groups is critical, as a single design or implementation is unlikely to suffice for all possible

groups. We recommend that designers of future encrypted email systems — and other security tools — begin by fostering a deep understanding of their target user population(s).

10. Additional Related Work

Section 2 discussed related work on usable email encryption in detail (e.g., [8, 11, 12, 21–26, 31]). More generally, Unger et al. [30] systematized secure messaging systems' usability, adoption, and security and privacy features along three dimensions: trust establishment, conversation security, and transport privacy. Using their criteria, Confidante optimizes for usability and adoption and provides basic security and privacy. In future work, Confidante could use alternative encryption schemes to provide features like forward secrecy, at the cost of compatibility with existing PGP clients. For example, Confidante could incorporate ideas from OTR [4] or Vanish [13]. Alternate key management and verification models also exist, such as CONIKS [18], or manual key management for expert users.

Others have studied the security and privacy practices and needs of specific user groups. McGregor et al. [16, 17] studied journalists, and we echo some of those findings in our work (e.g., the importance of protecting the relationship with sources) and find interesting differences among lawyers. Gaw et al. [12] study encrypted email adoption among activists, as discussed in Section 2.

11. Conclusion

Even user groups who regularly engage in sensitive communications have seen limited adoption of encrypted email. In this work, we present the design of Confidante, an encrypted email client that uses Keybase for automated key management. Confidante improves on traditional PGP tools by using Keybase's social media proofs to help users discover public keys and verify key ownership, but is still backwards compatible with PGP and email.

Our user study of lawyers and journalists using our Confidante prototype highlights important lessons for usable encrypted email, including the clear benefit of automated key management, the challenge of balancing automation with a feeling of security, and — via important differences that we uncover between journalists, lawyers, and generic email users — the need for security researchers to engage with and design specifically for target user groups.

Acknowledgments

We are especially grateful to our user study participants for their participation. We thank Jennifer Langston, Ed Lazowska, Emily McReynolds, and Kristin Osborne for helping us connect with local journalists and lawyers. We also thank Kelly Caine, Susan McGregor, and Emily McReynolds for early feedback on the tool design and other discussions. We thank Daniel Epstein for his advice on statistics. We thank Max Krohn and Jeremy Stribling of Keybase for quick responses to questions. We also appreciate the feedback from our anonymous reviewers, which helped improve the paper. We thank all members of the UW CSE Security & Privacy Research Lab for feedback on the study, the tool, and earlier versions of the paper, particularly Camille Cobb, Gennie Gebhart, Yoshi Kohno, Kiron Lebeck, and Lucy Simko. Kiron Lebeck suggested the name "Confidante". Finally, thanks to Chai and Tony for posing for Figures 3 and 4. This work was supported in part by the National Science Foundation under Awards CNS-1463968 and CNS-1513575.

References

[1] American Bar Association. Cloud Ethics Opinions Around the U.S. https://www.americanbar.org/groups/departments_offices/legal_technology_resources/resources/charts_fyis/cloud-ethics-chart.html.

[2] E. Atwater, C. Bocovich, U. Hengartner, E. Lank, and I. Goldberg. Leading Johnny to Water: Designing for Usability and Trust. In *11th Symposium On Usable Privacy and Security (SOUPS)*, 2015.

[3] D. Boneh and M. Franklin. Identity-based encryption from the Weil pairing. *SIAM Journal on Computing*, 32(3):586–615, 2003.

[4] N. Borisov, I. Goldberg, and E. Brewer. Off-the-record communication, or, why not to use PGP. In *ACM Workshop on Privacy in the Electronic Society*, 2004.

[5] K. Charmaz. *Constructing Grounded Theory*. SAGE Publications Ltd, second edition, 2014.

[6] J. Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37, 1960.

[7] N. Dell, V. Vaidyanathan, I. Medhi, E. Cutrell, and W. Thies. Yours is better! Participant response bias in HCI. In *ACM Conference on Human Factors in Computing Systems*, 2012.

[8] S. Fahl, M. Harbach, T. Muders, M. Smith, and U. Sander. Helping Johnny 2.0 to Encrypt His Facebook Conversations. In *Symposium on Usable Privacy and Security*, 2012.

[9] Federal Rules of Evidence. Rule 502: Attorney-Client Privilege and Work Product; Limitations on Waiver. Legal Information Institute. https://www.law.cornell.edu/rules/fre/rule_502.

[10] J. L. Fleiss, B. Levin, and M. C. Paik. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, New York, 3 edition, 2003.

[11] S. L. Garfinkel and R. C. Miller. Johnny 2: A user test of key continuity management with S/MIME and Outlook Express. In *SOUPS*, 2005.

[12] S. Gaw, E. W. Felten, and P. Fernandez-Kelly. Secrecy, Flagging, and Paranoia: Adoption Criteria in Encrypted Email. In *CHI*, 2006.

[13] R. Geambasu, T. Kohno, A. A. Levy, and H. M. Levy. Vanish: Increasing Data Privacy with Self-Destructing Data. In *USENIX Security Symposium*, 2009.

[14] Human Rights Watch. With Liberty to Monitor All: How Large-Scale US Surveillance is Harming Journalism, Law, and American Democracy, July 2014. http://www.hrw.org/node/127364.

[15] V. Li. Lawyers slow to adopt email encryption and other forms of secure communications, ABA survey finds. ABA Journal, Oct. 2015. http://www.abajournal.com/news/article/survey_finds_-lawyers_are_slow_to_adopt_email_encryption_and_other_forms_of/.

[16] S. E. McGregor, P. Charters, T. Holliday, and F. Roesner. Investigating the computer security practices and needs of journalists. In *24th USENIX Security Symposium*, 2015.

[17] S. E. McGregor, F. Roesner, and K. Caine. Individual versus Organizational Computer Security and Privacy Concerns in Journalism. In *Privacy Enhancing Technologies Symposium*, 2016.

[18] M. S. Melara, A. Blankstein, J. Bonneau, E. W. Felten, and M. J. Freedman. CONIKS: Bringing Key Transparency to End Users. In *24th USENIX Security Symposium*, 2015.

[19] A. Mitchell, J. Holcomb, and K. Purcell. Investigative Journalists and Digital Security: Perceptions of Vulnerability and Changes in Behavior. Pew Research Center, Feb. 2015. http://www.journalism.org/files/2015/02/PJ_InvestigativeJournalists_0205152.pdf.

[20] J. Nielsen. Powers of 10: Time Scales in User Experience. Nielsen Norman Group, Oct. 2009. https://www.nngroup.com/articles/powers-of-10-time-scales-in-ux/.

[21] K. Renaud, M. Volkamer, and A. Renkema-Padmos. Why doesn't Jane protect her privacy? In *Privacy Enhancing Technologies*, pages 244–262. Springer, 2014.

[22] S. Ruoti, J. Andersen, S. Heidbrink, M. ONeil, E. Vaziripour, J. Wu, D. Zappala, and K. Seamons. "We're on the Same Page": A Usability Study of Secure Email Using Pairs of Novice Users. In *CHI*, 2015.

[23] S. Ruoti, J. Andersen, T. Hendershot, D. Zappala, and K. Seamons. Helping Johnny Understand and Avoid Mistakes: A Comparison of Automatic and Manual Encryption in Email, 2015. http://arxiv.org/pdf/1510.08435v3.pdf.

[24] S. Ruoti, J. Andersen, D. Zappala, and K. Seamons. Why Johnny Still, Still Can't Encrypt: Evaluating the Usability of a Modern PGP Client, 2016. http://arxiv.org/pdf/1510.08555v2.pdf.

[25] S. Ruoti, N. Kim, B. Burgon, T. van der Horst, and K. Seamons. Confused Johnny: When Automatic Encryption Leads to Confusion and Mistakes. In *Symposium on Usable Privacy and Security*, 2013.

[26] S. Sheng, L. Broderick, C. A. Koranda, and J. J. Hyland. Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software. In *SOUPS Posters*, 2006.

[27] D. X. Song, D. Wagner, and A. Perrig. Practical Techniques for Searches on Encrypted Data. In *IEEE Symposium on Security and Privacy*, 2000.

[28] O. W. Systems. Who can I message on Signal Desktop? Where are my messages? http://support.whispersystems.org/hc/en-us/articles/218551897-Who-can-I-message-on-Signal-Desktop-Where-are-my-messages-.

[29] R. Thomas. PGP key verification, 2002. http://www.cymru.com/gillsr/documents/pgp-key-verification.htm.

[30] N. Unger, S. Dechand, J. Bonneau, S. Fahl, H. Perl, I. Goldberg, and M. Smith. SoK: Secure Messaging. In *IEEE Symposium on Security and Privacy*, 2015.

[31] A. Whitten and J. D. Tygar. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *USENIX Security Symposium*, 1999.

[32] K. Zetter. Magistrate Orders Apple to Help FBI Hack San Bernardino Shooter's Phone. Wired, Feb. 2016. https://www.wired.com/2016/02/magistrate-orders-apple-to-help-fbi-hack-phone-of-san-bernardino-shooter/.

[33] P. R. Zimmermann. *The Official PGP User's Guide*. MIT Press, Cambridge, MA, USA, 1995.