

# FEC E-filing Study

Modernizing the E-filing experience and infrastructure.

Prepared by 18F | February 2, 2017



# Table of contents

<b>Table of contents</b>	<b>2</b>
List of figures	5
List of tables	5
<b>FEC E-Filing study: Overview</b>	<b>6</b>
1. Make changes in a human centered, open, and agile way	6
2. Help filers report correctly	6
3. Help those who help the filers	7
4. Ensure future filing reliability	7
<b>The e-filing process</b>	<b>8</b>
Filing is just one element of the complex campaign fundraising processes	9
Filing software requires filers to manually integrate data from a variety of physical and digital sources	10
How do committees file?	
<b>Recommendation 1: Proceed in an open, agile, and human centered way</b>	<b>14</b>
Human centered design	14
Agile, iterative software development	14
User testing	15
Formalize feedback channels from FEC staff	15
Open source	15
<b>Recommendation 2: Help filers do the right thing</b>	<b>17</b>
Improve the general usability of FECFile	17
Improve data validation within FECFile	20
Support filers' workflow	34
<b>Recommendation 3: Help those who help the filers</b>	<b>35</b>
Support the developers both inside and outside FEC with open source software, documented in plain language	35
Support personal connections within the filing community	37
<b>Recommendation 4: Ensure future filing reliability</b>	<b>40</b>
Moving to cloud infrastructure	40
Changing underlying form design and processes to increase data quality	45
Paper filings processing opportunities	49
Improve data formats and structures to manage increased volume	54
<b>How to move forward: software development workflow &amp; technical road map</b>	<b>59</b>
Evolving software development workflow	59

Technical Road Map	61
Ongoing tasks and processes	64
Recommended tech stack	65
<b>Research methods</b>	<b>67</b>
Research approach	67
Phase 1: Contextual interviews	67
Phase 2: Prototyping & testing	68
Phase 3: Survey	68
Overall Study Limitations	69

## List of figures

- Figure 1. The flow of campaign finance information throughout the efilg system.
- Figure 2. Number of transactions on a typical report by type of filing software typically used.
- Figure 3. Example of in-context help and reference material from a prototype developed as part of the efilg study.
- Figure 4. Example of inline validation from a prototype developed as part of the efilg study.
- Figure 5. Example of current contact blocks and suggested contact rows for comparison of contact attributes.
- Figure 6. Responses in answer to the question “if a web-based version of FECFile were available would you use it?”
- Figure 7. Highest priority potential capabilities of a web-based FECFile among survey respondents.
- Figure 8. Screencast showing prototype of revisions to import function
- Figure 9. Image of the filer community feature from a prototype developed as part of the efilg study.
- Figure 10. Volume of transactions by receipts, disbursements, independent expenditures, and cumulative total
- Figure 11. Number of transaction records filed per day.
- Figure 12. Number of transaction records processed by hour of day.
- Figure 13. Diagram of the current information management process.
- Figure 14. Example of form fields that serve dual purposes on the paper form.
- Figure 15. Process diagram showing electronic data flow
- Figure 16. Process diagram showing current paper process
- Figure 17. process diagram showing revised/improved paper process
- Figure 18 & 19. Comparison of report images with and without QR codes.

## List of tables

- Table 1. Issues in that resulted in RFAs in the first quarter of 2016
- Table 2. How different groups in our survey feel about using a web-based version of FECFile.
- Table 3. Projected costs of paper automation contract

# FEC E-Filing study: Overview

The objective of this study was to understand how the FEC can modernize the electronic filing (e-filing) system to get better quality data, be responsive to user needs, and handle increasing volumes of data.

Campaign finance data that is reported by political committees, candidates, and groups is critical to the transparency component of the FEC's mission. Over the past two decades, as records have evolved from paper to microfilm to databases and APIs, report filing has moved from a paper-based process to a process that depends heavily on software. The FEC engaged 18F to conduct this study of e-filing in order to look forward and prepare for users' changing needs.

We interviewed a wide range of people who interact with the e-filing process and then focused on top priorities for filers, data users, and FEC staff to create and test prototypes aimed at improving the filing experience. From this research, four main recommendations emerged:

## 1. Make changes in a human centered, open, and agile way

The FEC has been successful in revamping fec.gov using an agile, human-centered approach. Our primary recommendation for improving the e-filing system and experience is to continue and extend this approach. This will help ensure any changes to the e-filing experience and infrastructure are based on evidence that they will improve user experience. An open source, API-driven approach will also help the FEC expand their use of industry-leading software development practices and make it easy for commercial campaign finance software to incorporate these improvements. Incremental changes with frequent user testing can help ensure that the FEC is continuing to best serve the wide range of people who depend on its data.

## 2. Help filers report correctly

We found that filers often want to report correctly, but struggle to do so. Even data power users like journalists pointed to filer confusion as a cause of data errors. Filers spoke of a desire to file correctly the first time so they would not have to amend reports, receive requests for additional information, or handle calls from the press.

One of the most powerful ways to make the filing process more intuitive is to make the FECFile software more usable, and we recommend improvements to FECFile in this report. Helping filers complete forms correctly is critical because it improves their experience and improves the experience for other users of the data both inside and outside the FEC.

### 3. Help those who help the filers

FEC staff help improve data quality and the filer experience by helping people file correctly. Some filers also take advantage of a robust vendor community that integrates with FEC tools to help filers. Vendors often offer filing as part of a broader suite of products and services that make running a campaign easier, so making it easier for vendors to integrate with FEC tools is an effective way to improve the experience for filers, vendors and data users. In our research, supporting those who help filers complete reports successfully emerged as an important means of ensuring accurate data.

### 4. Ensure future filing reliability

Notably, while our interviews with filers and vendors revealed ways that the FECFile interface and software could be enhanced, these interviewees articulated few concerns with the overall e-filing infrastructure's reliability (up-time). Through our analysis and interviews with stakeholders close to the process, we learned that this is not due to the fundamentals of the technical architecture, but rather heroics on the part of the team maintaining the legacy system. The FEC has, through evolving practice, timely fixes, and sustained attention, kept the system working under substantially increased load. In coming years, the FEC will need to support rapidly increasing volumes of data while maintaining performance and reliability. To do this, the FEC needs to expand existing work to streamline data flow and build for scalability in the long-term.

Some of this work will be tied to development processes. For example, increasing testing and automation can make the development process smoother and more resilient. In other cases, changes to the filing process — such as making data-gathering less error-prone — can help reduce load on the system. Likewise, mitigating the impact of large paper filings, which are more costly and less accurate than electronic filings, would yield data quality dividends as well as improve maintainability.

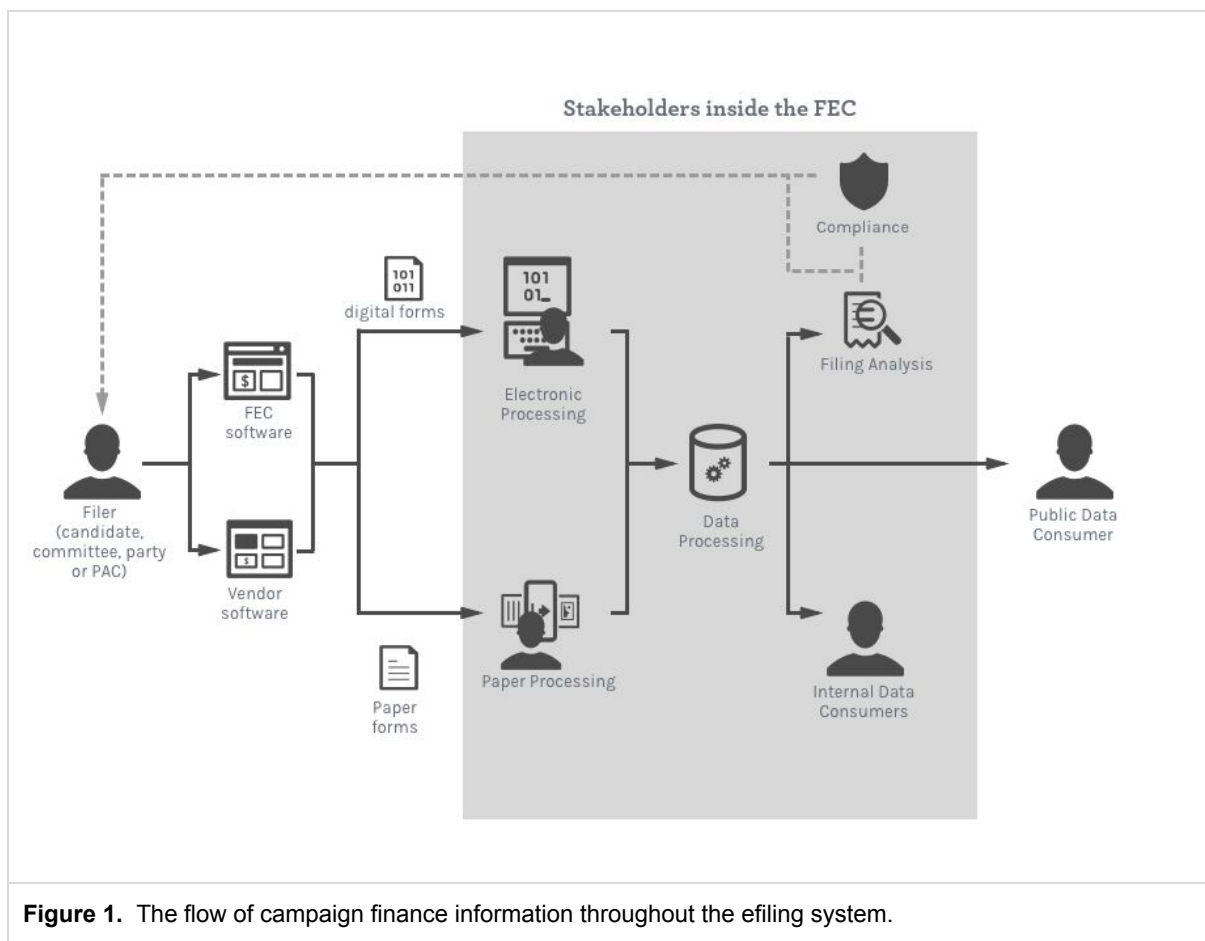
Infrastructure changes can also support exponential growth; creating a cloud-based system will allow the system to scale up or down to adapt to spikes in traffic and use.

In some cases, (such as for larger structural changes) technology and interface design plays a smaller role and higher-level agency and legislative approvals will be necessary to fully address the problems.

# The e-filing process

In order to understand the e-filing process, we interviewed and observed dozens of people, including those who file reports to the FEC, people in various job roles at the FEC, and vendors of commercial campaign finance software. While these stakeholders that comprise the e-filing ecosystem may have different missions within their own organizations, they all play critical roles in helping the FEC achieve its mission of publicly disclosing campaign finance activity.

Our understanding of the way campaign finance information flows through the e-filing ecosystem is depicted in Figure 1.



**Figure 1.** The flow of campaign finance information throughout the e-filing system.

- Committees, candidates, parties, and PACs accept contributions and spend money to fund political committees.
- Each committee or committee's designee tracks the flow of funds and then reports this information to the FEC using paper forms, vendor software, or the FEC's filing software, FECFile.

- The data in these reports passes through the Electronic Filing Office (EFO) and then moves on to the Reports Analysis Division (RAD), where algorithms and analysts review the report and check for errors and discrepancies. The data is also augmented with helpful codes, like committee IDs and type descriptions.
- At the same time that the report is being reviewed, it is made publicly available so data consumers (including journalists and the general public) can access it through the FEC's website, API, and bulk data offerings.

FEC has been a pioneer in open data, but some data available on the FEC website remains difficult for journalists, researchers, and others to use. **During this study, we learned that many data quality issues can be traced back to confusing elements in the filers' experience when they are preparing information for the FEC.** When filers make mistakes or misunderstand form fields, their errors have repercussions through the entire data ecosystem: data entry errors create a greater processing burden on the EFO and heavier workload for FEC personnel, both of which can lead to delays in data becoming publicly available. Ultimately, errors impact the timeliness and accuracy of data available to the public.

## Filing is just one element of the complex campaign fundraising processes

Participants in the campaign finance process (candidates, committees, parties, and PACs) have a complex set of activities to manage: courting voters, working with candidates, recruiting donors, raising and collecting money, using it to pay for campaign related expenses, and balancing their checkbooks. Filing activities follow timelines and processes separate from day-to-day campaign activities: filing deadlines are defined in the [Federal Election Campaign Act](#) (FECA), and depend on the type of committee. Filings may be monthly, quarterly or semi-annual depending on committee type and whether it is an election year. Committees involved in an election must also file pre and post election reports and often must report certain financial activity within 24 or 48 hours of occurrence. [Read more about filing requirements.](#) Additionally, the people managing the campaign on a daily basis are often not the people who file reports.

The two interviewees quoted below are compliance professionals who file on behalf of multiple clients. Here they describe conflicts between the rhythms of their clients' day-to-day work and FEC reporting rules and schedules.

*"It is all scheduled and triggered by schedules. And the 24 and 48 hour [reporting] schedule... [We] set up a process to track the spending. [I have] some clients who have their own spreadsheets. Some have their own software. [I have] to explain to clients: 'This is different.' It's all scheduled dates, and 24 and 48 [hour reports] which*

*are triggered by spending. I'm back end compliance, we set up a process to track the spending. Some clients have spreadsheets, some have sophisticated systems. Every client is different—no one size fits all process.”*

*“I have one unit that has like 15 different payrolls. Some of them are weekly, some of them are bi-weekly, some of them are semi-monthly. And just the reconciliation process is the most challenging aspect of it, ‘cause you’ve got to go through and when you’re doing multiple payrolls for periods that don’t end on the end of the month—the last one ended on the 28th. We started in the middle—we started on the 20th of October, and then the period closed on November 28th—so when you go to reconcile with the bank, you don’t have a definitive cutoff date. You have to get interim statements from the bank, and not all the payrolls may have cleared.”*

## **Filing software requires filers to manually integrate data from a variety of physical and digital sources**

Many of the filers we talked to use financial and accounting software applications (such as Quickbooks), payroll application exports, bank statements, membership databases, and physical check registers to help them complete their filings. Many rely on Microsoft Excel as an intermediary or central exchange to tie the tools together.

Some committees use commercial campaign finance management software that is designed to support more campaign processes in addition to reporting capabilities. Because the FECFile software is designed to support reporting activity alone, FECFile users must bridge the gap between the process and data format required by FECFile and their existing campaign management workflows. This often requires significant overhead tasks both before and after they enter data into FECFile, and makes reporting more onerous.

*"I download through ActBlue, where most of my transactions come from. Then I have a [Microsoft Excel] file where everyone that's given me cash or check... I take the form, cut and paste everything. I do everything in the format of Act Blue; I have to reformat everything for FEC."*

## **How do committees file?**

Committees and their designees file their reports to the FEC in one of a few ways:



## **Filers using FEC provided software [59% of entities filing electronically]**

There are two principal ways filers can use FEC provided software to file: using FECFile, a desktop based application for filing most forms, and for some filers, using specific forms that have been developed as web applications.<sup>1</sup> The vast majority of filers use the former, FECFile, to complete at least some of their filings.

FECFile is free software designed to support political committees in reporting their campaign finance activity to the FEC. Committees that use FECFile often have relatively small budgets; several people we talked to described committees that use FECFile as "grass roots," and postulated that these committees choose FECFile because they cannot afford commercial software.

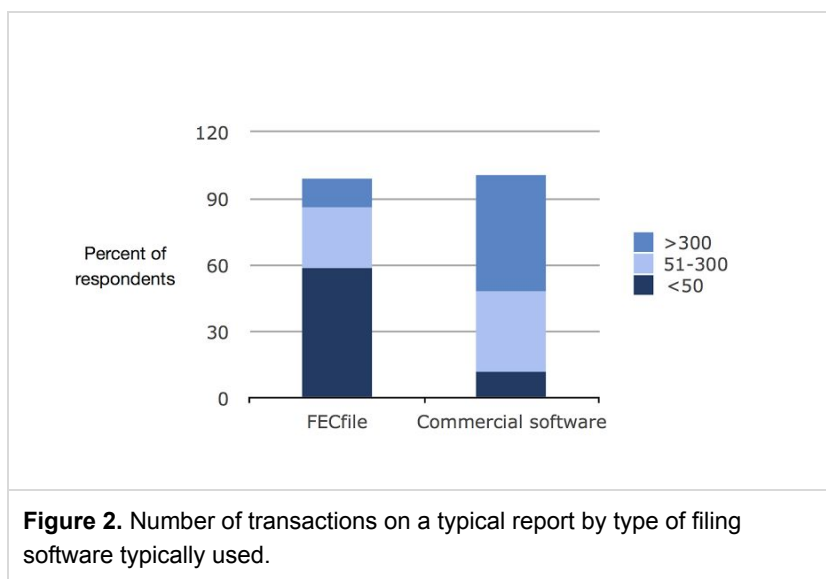
Most filers who typically report small numbers of transactions use FECFile (92% of the filers in a survey we conducted). One way to interpret this number is that it is the small "mom and pop" committees that help the FEC live up to its ideal that "anyone can run for office" and who really rely on FECFile being easy to use.

The majority of FECFile users in our survey (59%) identified themselves as treasurers for candidates, committees, parties or PACs, and 100% of self-identified treasurers said their reports typically include fewer than 50 transactions. Only 17% percent of FECFile users in our survey identified themselves as compliance professionals. Because most FECFile users are not compliance professionals and typically work for small campaigns, they may be learning the rules of campaign finance reporting at the same time they are trying to learn the ins and outs of electronic filing and the idiosyncrasies of filing software.

Over half (58%) of FECFile users in our survey estimated that their reports typically include fewer than 50 transactions, 27% said their reports typically included 51-300 transactions, and only 13% reported that their reports typically include more than 300 transactions. This is in contrast to the respondents who typically use vendor software who file higher volume reports (Figure 2).

---

<sup>1</sup> The FEC has [online web versions of some forms](#) (Forms 1, 2, 5, 6, 9, 24, and 99), however, transactions filed via webforms make up <1% of the transaction total. In addition, many filers that use these web forms will also need to file schedules that are not available as webforms to remain in compliance, and will thus fall into one of the three categories of paper, FECfile, or vendor software filers.



RAD analysts in the Authorized and Party/Non-party branches estimated that 50-60% of calls come from novice filers, and observed that novices and infrequent callers take up a much larger percentage of the actual time they spend on the phone. We asked RAD analysts in the Compliance Branch to characterize the people they get calls from:

*"...[I] spend more time with the people who are novices just learning about filing. Most people are novices. "*

### Commercial software filers [41% of entities filing electronically]

Some political committees can purchase tools to help manage campaign finance tasks, including fundraising, list management, and reporting to the FEC. This software typically costs hundreds of dollars per month and allows committees to complete many campaign finance tasks within a single system. These filers often report their information to the FEC using the same system they use to process and store their financial data.

### Paper filers

Filers who raise or spend \$50,000 or less in a calendar year may choose to file on paper. While these make up the largest number of paper filers, the relatively low transaction amount limits the burden they impose on the agency. Filers who raise or spend more than \$50,000 during an election cycle must file electronically, with one exception: Senate filers.

By transaction volume, Senate filers generate the most paper because they are bound by Senate rules to first file with the Secretary of the Senate, who currently only accepts paper filings. The Commission's 2016 Legislative Recommendations identified that e-filing for Senate committees could save at least \$876,000 per year and make these vital records available to the public within minutes. Discontinuing larger paper filings was also one of the major recommendations of the an E-Filing Study that the FEC commissioned in 2013.

# Recommendation 1: Proceed in an open, agile, and human centered way

The FEC has been successful in revamping fec.gov using an agile, human-centered approach. Our primary recommendation for improving the e-filing system and experience is to continue and extend this approach. This will help ensure any changes to the e-filing experience and infrastructure are based on evidence that they will improve user experience. An open source, API-driven approach will also help the FEC expand their use of industry-leading software development practices and make it easy for commercial campaign finance software to incorporate these improvements. Incremental changes with frequent user testing can help ensure that the FEC is continuing to best serve the wide range of people who depend on its data.

## Human centered design

Human centered design (HCD) is a methodology that incorporates feedback from the people for whom you are designing (users) throughout an iterative design process. The goal of human centered design is to end up with a solution that is tailored to meet users' needs, with little wasted effort or risk of unintended, negative consequences. Critical to achieving this goal is working in small iterative cycles where design teams (1) talk with and observe real users to understand their needs, context, and challenges, (2) come up with design concepts that might address these challenges before making or building out one or more of these concepts, and then (3) testing them with real users.

A human centered design lens can be applied to the design of almost anything from technology to policy, but it is increasingly associated with the design of software and information systems. While HCD is still a relatively new field, it is the leading edge in commercial software design and development, and it is now a part of many formal education programs in computer science and engineering, interaction design, and information science.

The FEC has already adopted a human centered and agile (described below) approach to the design of its new website, which has been received very positively by the campaign finance community. **We recommend that the FEC continue this work by expanding its human centered design approach to include modernizing the e-filing system.**

## Agile, iterative software development

Agile methodology insists on building large systems incrementally by making small pieces of working software. Using an agile approach, teams make progress incrementally, checking and

course correcting as necessary. Combined with human centered design, new concepts are researched, prototyped, tested, and implemented or changed in response to user feedback

The Beta FEC project has instituted the Scaled Agile Framework (SAFe) planning technique of Program Increment (PI) planning to accommodate frequently changing user requirements and to coordinate and communicate goals and progress.<sup>2</sup> **We recommend that the FEC leverage the skills and experience within the organization and on the Beta FEC team to apply an agile approach to any changes to the e-filing system.**

## User testing

While user testing is fundamental to both human centered design and agile software development, we want to highlight specific user groups that should be involved in testing; **any changes to the e-filing system must done with input from filers and data power users, such as data journalists and academics**, who would be impacted by the change. The FEC is regularly incorporating user feedback into the design of the new website and should carry this practice over to efforts to modernize e-filing.

## Formalize feedback channels from FEC staff

In the course of our research, we were impressed by the dedication and expertise of the FEC staff. We learned of several informal processes by which internal FEC staff share ideas on how to improve e-filing. For example, analysts in the Reports Analysis Division (RAD) have created their own centralized list of requested changes to the e-filing system. While RAD participates in a final quality check on new versions of FECFile, the FEC's free filing utility, this may not be in time for their feedback to be incorporated. **The FEC could leverage internal expertise by further involving staff, particularly analysts, in the formative evaluation of design changes.**

In addition, mentioned in the [data validation section of the report](#), it would be beneficial to mine the RFAI data on a quarterly basis to see if there are opportunities to add data validation or make other improvements to enhance e-filing.

## Open source

The FEC has a chance to lead the government community (state and federal) in the creation of open source forms, the lack of which is a problem that is mirrored across the government. Building the e-filing system and related utilities using open-source software could help build trust among users and vendors. This would make it easier for commercial software vendors to

---

<sup>2</sup> Business process will be discussed more in-depth in the upcoming 18F cloud study.

integrate with FEC software and make it clear what changes are made to the modules the vendors use. We include more on this recommendation in the [technical approach](#) section.

## Recommendation 2: Help filers do the right thing

Every filer we heard from wants accurate and complete reports, but we found that proper reporting can be a struggle, given the existing processes and tools. Many filers have established complex workflows to support their reporting efforts, and several had stories illustrating considerable fear about doing the wrong thing.

In this section, we recommend ways the FEC can empower filers to file the way they want to file—accurately and completely. These recommendations cover FECFile’s general usability, data validation in FECFile, and support for filers’ workflow and community.

### Improve the general usability of FECFile

There are many existing user interface patterns and best practices for form design that can help make FECFile more usable. In many of our sessions with filers, we found that the current design and presentation of many form fields causes confusion, which results in filers reporting incorrect or incomplete data. The FEC has made recent usability improvements to some aspects of e-filing with web-based versions of some forms, and we recommend that the FEC prioritize improvements to FECFile going forward.

#### A. Provide clear labels, suggestions, and hints

FECFile does not make it easy for users to identify which input fields are required, recommended, or not required. The resulting confusion generates a lot of calls to RAD, particularly around deadlines.

In some cases, filers are guessing as to whether they have input the right information or selected the right option from a dropdown.

*“There was a drop down, but no place to write in text. We would have expenses that didn’t quite match what was in the dropdown. This led to questions like “we are trying to do this as accurately as possible, but if this category is not right, is the blame on us...?” I wanted to make sure I crossed Ts and dotted Is... oftentimes it didn’t seem clear how to list certain expenses from that dropdown menu.”*

All form fields should be clearly labeled and accompanied by hints about desired data and necessary format. See figure 3 below for an example of a in-context help prototype that tested well during our research. In addition, each form field should have only a single meaning. You can find more of our recommendations on form field specifications in recommendation 4: Ensure future filing reliability. Finally, the user interface should clearly indicate which form fields are optional and which are required.

Breadcrumbs » Breadcrumbs

Friends of Bo | October Monthly Report Due in: 3 Days

### Enter Transaction | Contributions

AMOUNT

TRANSACTION DATE  ELECTION

☒ INDIVIDUAL ☐ COMMITTEE ☐ OTHER

LAST NAME  FIRST NAME  SUFFIX

ADDRESS

ADDRESS LINE 2

ZIP

Full Name (Last, First, Middle Initial)  
A. Dolby, Thomas

Mailing Address  
32 Wallaby Way, Suite 3032

City Sydney State WA Zip Code 20007

FEC ID number of contributing federal political committee C

Name of Employer Megahyper 2000 inc Occupation Analyst

Receipt For: ☐ Primary ☒ General ☐ Other (specify)

Election Cycle-to-Date

Date of Receipt 10 01 2016

Amount of Each Receipt this Period

☐ Memo Item

**Figure 3.** Example of in-context help and reference material from a prototype developed as part of the eFiling study [animated demo available at <https://beta.fec.gov/about/reportse-filing-study-2016/>].

## B. Remove redundant and unnecessary fields

When a filer creates a report, they enter information FECFile could use to customize subsequent forms and remove fields that are not needed for the filer to meet their specific reporting requirement. For example, when a filer opens the committee file, they must enter the committee ID before moving on to select the type of report they wish to work on. The types of forms a committee might be required to complete varies by the type of committee, but this is not reflected in the software, which currently presents all committees, regardless of committee type, with the complete list of all possible forms. This list could be tailored to show the filer only the types of reports that are relevant to their type of committee.



The FEC is already doing this in another area with their online web forms, which have been received favorably by the filing public and commercial software vendors. For example, the [web-based version of Form 1](#) guides new committees through the registration process by asking them clear questions and then progressively tailoring subsequent options and form fields based on their answers. We recommend that FECFile be enhanced to take into account previously entered information.

In another example, a committee's election cycle could be auto-populated based on what the system knows about the candidate, election, and state (which the filer entered when they created their report).

The system could also help users and prevent errors by allowing them to set certain fields as the default for a given work session. For example, when a filer creates a report, they enter the type of election and election year for the report, yet they are asked to enter this information again for each individual transaction. There are some cases in which a filer might want to modify this, but in most cases this information will be the same for all the transactions in the report.

*"When you're doing these reports, once you start one, you're only doing it for that one [report]. Seems like it should automatically come up default to be the same one [election and election year] every time, and then you could change it if you needed to—It'd be a lot easier."*

### **C. Provide triggered and dynamic help**

In addition to clearer form field display and labelling, FECFile should link to explanatory text about filing requirements. We heard from many filers that they often have questions throughout the filing process, including questions about campaign finance law, what information they need to report, and how to use FECFile. During research sessions, both experienced and novice FECFile users interrupted their reporting process to search for information about how to complete form fields. Providing contextual information within the FECFile interface will reduce the need for filers to switch back and forth between FECFile and external reference material.

We prototyped one idea for how to provide better context-specific information within FECFile by linking users from a specific form field to the glossary definition of data that belongs in that field.

*"I think this is really helpful... having the definition right there, in the glossary, is extremely helpful. In the old software, there was nowhere to get that guidance—I would go to Google. But it was hard to match up, having the keywords line up and know what's there to help you is great."*

## **D. Show users where they are in the process**

Providing users with a way to track their progress through a complex process, such as e-filing, is a best practice in user experience design. Many users in our research cited time constraints as a major concern during the filing process. Providing clearer feedback about where people are in the process of submitting their reports could help them plan for the time they need and allow sufficient time for reviewing and correcting their information before deadlines.

In addition to notification of where they are in their reporting process, filers need clearer feedback that the system is processing their information, particularly when they are uploading their reports. Uploading can take a variable amount of time depending on the FEC's server load, and without feedback that the upload is in process, filers sometimes quit or restart the upload again and again before the upload is complete.

## **Improve data validation within FECFile**

For the data consumers who ultimately use FEC data, data quality is a primary concern. A [2014 study](#) conducted by the [Sunlight Foundation](#) and the [Center for Responsive Politics](#) found widespread underreporting of required fields:

*In the last six months, campaigns, PACs, individuals, parties and other organizations have, on average, submitted almost 5,000 blank required fields per month to the FEC, according to the FEC's submission validation tool, which is distributed to vendors of filing software. And the most common errors are some of the most important to the voting public: 82% of required fields left blank have to do with identifying either the donor or recipient of a contribution.*

The most common error is not providing basic, essential information like candidate information. Data validation could easily alert filers of these types of errors. Currently, data validation does not happen until the start of the submission process.

Giving users real-time feedback about their information can help reduce errors. We recommend improving how FECFile handles data validation by providing on-the-spot field validation for critical fields, promoting contact record validation, and making the process for resolving potential duplicate contacts easier.

Currently, FECFile performs minimal data validation: filers can submit forms through FECFile with required data missing or incomplete. The validation requirements were created at a time when online systems were new and people were uncomfortable with the concept of electronic filing. At that time, the FEC sought to minimize the impact of the change from paper to electronic

reporting on filers, so the system's design was based on existing paper forms, and the data requirements were flexible with few but the most egregious data errors being rejected. This system has since been outgrown by complex data needs of the Agency and those of the public.

Data users both at the FEC and outside have increasingly gone to automated data processing and review systems that do not function optimally with the loosely structured data provided by the current system.

When required data is missing, analysts in the Reports Analysis Division (RAD) may send a Request for Additional Information (RFAI).

As part of this study, we obtained a sample of the top issues in that resulted in filers receiving RFAs in 2016 (Table 1). RFAs are sent when an FEC Campaign Finance Analyst needs additional clarification or identifies an error, omission or possibly prohibited activity.

**Table 1.** Top issues that resulted in RFAs in 2016

Occurrences	Description
367	Incorrect Column B Figures
330	Receipt of Contribution from Persons/Individuals in Excess of \$2,700 and/or Qualified Multi-Candidate Committee in Excess of \$5,000
321	Column B Totals Incorrect
229	Cash Discrepancy
169	Inadequate Purpose
169	Failure to File 48-Hour Notice
160	Line 8 Incorrect (Column A/B Discrepancy)
158	Report Not Signed by Designated Treasurer
144	Receipt of Contributions from Unregistered Committees or Organizations
134	Contributions to Candidate After Election -- No Debt Designation
124	(IN) Category(s) of Financial Activity on Incorrect Line
123	Name of PCC Does Not Include Candidate's Name
113	Inadequate/Incomplete Contributor Information - Citing Emp/Occ (Best Efforts)
105	Beginning Cash Discrepancy
102	Contributions/Transfers Made Missing Information
91	Payments for Salary/Wages Allocated on Schedule H4
91	Omission of Candidate Information
89	Multicandidate Committee Contributes > \$5,000 to Candidate
83	Reattribution/Redesignation Requested or Refund to be Issued for Contributions from Individuals
82	Receipt of Primary/General Contributions after Primary/General

Responding to RFAs is a stressful and time-consuming process for the filer, because it often means they will need to amend their filing. Even small errors can compound, causing a cascade of amendments that take additional time and stress to untangle.

*“Any error you ignore grows and becomes worse. It’s just something—you have to be constantly looking at it.”*

*“Sometimes I’ll spend hours just looking at it because as soon as I push that button, if I made a mistake, I’ll have to make an amendment. It’s not like ‘if you made a mistake and you can fix it.’ You’re going to get a letter.”*

In addition to costing filers time and stress, incomplete data also results in RAD analysts spending time on correspondence and phone conversations to help filers disclose information correctly, which may include guiding filers through fixing data issues. Better data validation could save analysts significant time. Analysts’ time is scarce; each RAD analyst is assigned to support between 250 and 400 Committees, and any time saved spent helping filers fix common errors could free up time to review reports or provide deeper support.

#### **A. Help filers identify and resolve errors by providing real-time validation for form field entries**

More feedback about data quality during the form completion process would improve the filer experience and would more evenly distribute the burden of data quality control among filers and the FEC. Transitioning control over data quality to filers upon data entry was also a key recommendation in the 2013 e-filing study.

The basic framework for tighter data validation is already in place; many of these errors already trigger warnings, but allow users to submit their reports anyway.

In-line validation should be implemented iteratively through a human-centered design and evaluation process. The risk of stricter data validation is that it can cause upfront frustration for filers. Right now filers can enter incorrect information without pausing, but in-line validation might require filers to address errors before moving on. Careful implementation can balance this risk with the downstream effects in a user-centric way, but more research is needed to know how validation should be implemented.

FECCheck already runs as part of the submission process for FECFile users, but FECCheck before submission is optional for vendors. To mitigate the effects of omissions and errors in the short term, the module that sends information from vendors to the FEC, FECLoad could be modified to run validation. The submission process could still allow the validation module to be called independently but also invoke validation as a part of the client-side upload process before submission.

### *Prototype testing of in-line validation*

We prototyped one idea for in-line date validation in which filers received a warning when they entered a transaction date outside of the reporting period. Filers who tested this prototype largely missed this warning at the field level; filers, particularly power users, focus on the source of the information they are entering into FECFile (their spreadsheet or other account document), and they seldom look at the FECFile UI. This implies that while relying on real time warnings and validation may help users avoid misreporting, those alerts should find ways to signal the users as close in time to when the potential error happened, as soon as their focus returns to the FECFile UI. For example, using an audio indicator to indicate that an attempted submission of a transaction failed would draw the user's attention back to the problem field shortly after it was entered (Figure 4).

The screenshot shows a web interface for 'Friends of Bo | Welcome Back!'. The main section is titled 'Enter Transaction | Contributions'. It contains three input fields: 'AMOUNT' with the value '300', 'TRANSACTION DATE' with the value '09/09/2016', and 'ELECTION' with a dropdown menu showing 'primary '16'. Below these fields is a 'CONTACT TYPE' label. To the right of the input fields is a vertical list of names: 'Name', 'YGLE!', 'THOM', 'THOM', and 'THOM'. The 'TRANSACTION DATE' field is highlighted with a red border, indicating an error or warning.

**Figure 4.** Example of in-line validation tested as part of a prototype developed for the efilings study [animated demo available at <https://beta.fec.gov/about/reportse-filing-study-2016/>].

### **B. Help filers fix errors and omissions by providing clear warning and error messages**

It is also equally important that the validation runs quickly and error messaging makes the problems clearly understood by the users so they can easily address data problems. Having more lightweight data checks through the process can also avoid filers having to deal with a large number of errors right before the submission. Without in-line data validation, the validator finds an error or warning just prior to submission. Because filers receive errors and warnings far

from when they entered the data, they have trouble recalling and navigating back to where they need to make corrections.

Handling error and warning messages is further complicated by the fact that the messages that FECCheck returns are difficult to interpret and do not guide filers in how to handle the issue. For example, a user might get a warning that says “Leading Blanks {e.g. " TEXT"} not allowed,” which indicates that there are added spaces, or “blanks” that precede text they entered at some point earlier in their reporting. Since this could be referring to any number of form fields, figuring out first how, and then where to delete the added space can be like looking for a needle in a haystack.

Several interviewees mentioned error messages as a cause of frustration and confusion:

*“When I was trying to submit and it wasn't going through it would be helpful instead of an error message, it told me where to go to fix it. In general more thorough information help.”*

*“The error messages are so confusing to people. Treasurer signature date? That's the date you close the report, it doesn't say anything about it in the software. The validation message should say “close your report”*

Additionally, using more accessible language was one of the 2013 e-filing study recommendations.

Unclear error messages also create problems for vendors who use the FECCheck module, because these same confusing messages are returned via the API. This creates overhead for software vendors who either need to translate the messages in their own software or field calls from confused customers, as in the quote below.

*“On the validator? [My] biggest complaint is explaining what the messages meant. We've had to translate that on our end [to be] sure I understood it. The validator isn't wrong, it's just cumbersome” - Commercial software vendor*

Some analysts in the Reports Analysis Division keep track of the errors and warning messages that frequently confuse filers. The FEC could leverage RAD expertise in creating improved warning and error message content.

### **C. Validate contact records as part of the submission process and make the reconciliation process easier**

We recommend helping filers maintain clean contact lists by making contact validation part of the submission process, using validation logic that can catch more duplicate contacts (even

when there are small differences in text), and improving the process for resolving duplicate contacts.

Filers often make small re-keying errors, which can result in duplicate contact records. These duplicates can lead to undetected excessive contributions or incorrect aggregate figures.

FECFile has a feature that checks for and helps resolve duplicate contact records when filers import their data, but we observed users struggling to move through workflow for checking and resolving duplicates. This used to be an automatically triggered feature, but the FEC was asked to turn it off because filers found the existing feature difficult to use. With this automatically triggered feature disabled, filers who enter data manually are never prompted to check for duplicate records or clean up their contact list. Some filers who chose manual data entry may not know that this feature exists at all.

We recommend that the FEC prioritize the iterative, human-centered design and development of FECFile's capability for checking and resolving duplicates to make this feature more usable, and then make the new and improved feature a part of the submission process.

#### *Prototype testing of finding and resolving duplicate contacts*

We explored how filers might react to these changes by testing a prototype submission process that checked for duplicate contacts.

Filers reacted positively to the idea of making contact validation a part of the submission process, but several commented that they would need a way to bypass the process of resolving duplicate contacts if they were short on time.

*"I like that this [merging duplicate contacts screen] comes up. I can very easily see multiple donors with similar info."*

*"Seeing these every month may be annoying, but if you are doing the same report, there probably aren't that many. It may not be so bad to have because it will pick up fat fingering, and that is not a bad thing. Annoying, maybe, but I bet it catches more stuff that you'd be glossing over in reviewing the report."*

Filers appreciated when the prototype identified duplicates using "fuzzier" validation logic:

*"This is a weird thing with FECFile: if it [the contact record] wasn't pretty much exactly the same, it wouldn't pick up on the fact that those might be the same. I like that this is grabbing more things. Sometimes we get biz address and personal address and we can't tell that it is the same person."*

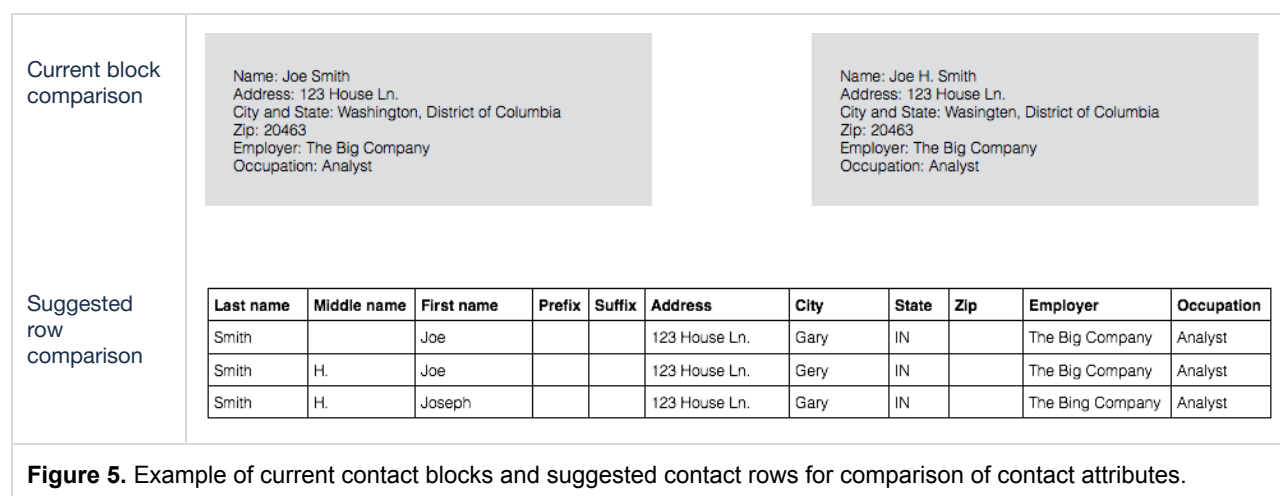
Once potential duplicate contacts are found, filers need to be guided on how to resolve them. We tested several different designs for how filers might resolve duplicate contacts. The current system handles this by guiding users through merging duplicate contacts into a single record. We kept this same basic concept, but tested different UI designs for how this concept is communicated and visually represented.

We learned that the language and visual representation of what happens when you merge duplicate contacts needs to be crafted with care. In both our initial observations of filers using FECFile and in our prototype testing, users gave clear indications that they were confused as to what merging would do to the contact records in their database.

*“Is this giving you the option to merge? Is it combining the transactions? What I don’t want to do is leave the record in there, that leaves you susceptible. You obviously can’t delete the info. Now how you get there, I don’t see. It’s not clear to me how the contributions and transactions are getting moved.”*

Without a clear understanding of what merging contact records will do to their database, filers will avoid resolving duplicates.

The design of the merge feature needs to help filers quickly identify discrepancies between potential duplicate contacts by enabling them to compare like form fields directly adjacent to one another. We learned in our interviews and observations with filers that they determine whether a set of contacts are duplicates by comparing specific contact attributes rather than the contact records as a whole. Arranging potential duplicate contacts in rows with columns for the different attributes (name, address, occupation, employer, etc.) would support this comparison far better than the current design that places the complete contact records side by side (**Figure 5**).





## Support filers' workflow

### A. Offering a web-based version of FECFile will allow filers to work the way they want to work

FECFile is currently a desktop application, which means it comes with limitations that constrain many filers' workflow. FECFile is currently only supported on PC, but many filers want to be able to work on reports using multiple computers, including both Windows and Mac operating systems. Filers who primarily use Macs must find secondary computers, use a windows emulator, or seek out PC users who can file on their behalf. Filers also expressed a desire for cloud-based storage, access to FECFile away from the office, and the ability to collaborate with colleagues in other places to complete reporting tasks.

*"He asked me to do the filing because he had a Mac and he couldn't use FECFile."  
(interviewee has a PC)*

*"The one big drawback is that we use Macs in our office, and you can't download it [the software] on a Mac."*

*"I travel extensively and sometimes must complete the report from locations other than my office and computers other than my regular workstation. As a result, I must remember to download the data file to a USB drive and take with me to use. Then I must remember where my most recent data is stored. A web based system would eliminate this problem."*

*"[a web-based system] would absolutely be helpful! There are two of us that work on the FEC filings and, currently, I have to wait until I can use the other person's laptop (where the FECFile software is stored) in order to enter the data needed for a filing."*

*"[a web-based system] would be an outstanding upgrade. Many of my colleagues (and I) are Mac users and we've had to keep a PC around the office for the sole purpose of filing our monthly reports."*

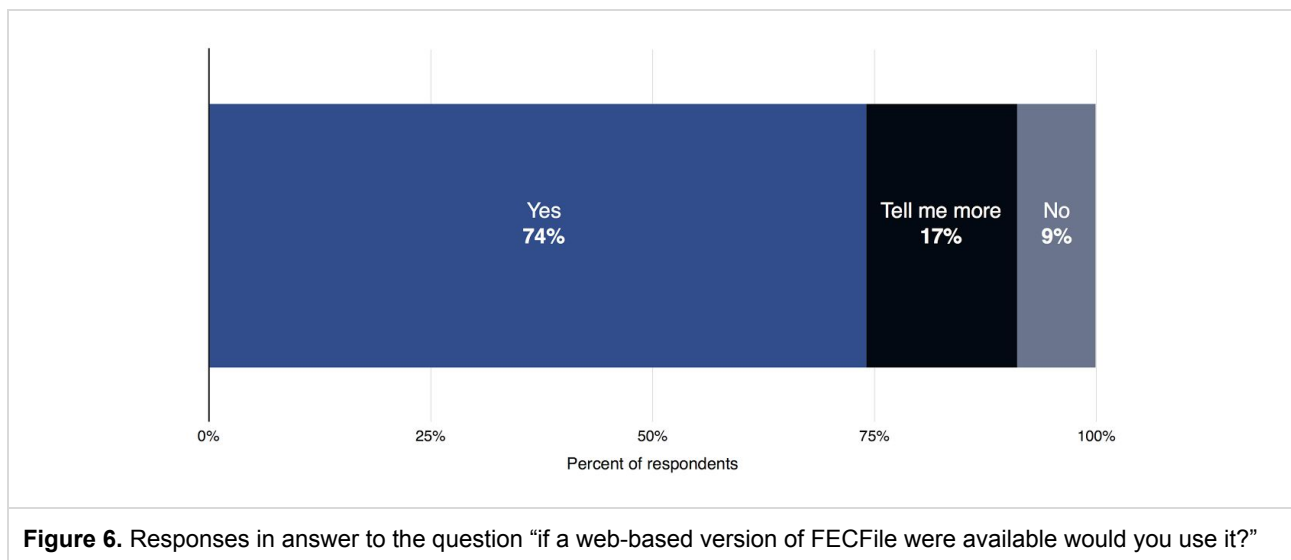
*"I live in constant terror of a computer crashed around filing time, even to the point of being afraid to make updates on my computer in the event that I may lose filing information that I have already entered. A cloud-based storage system would provide so much relief from those concerns."*

One way to deliver these capabilities is to move the software, which is stored locally on users' computers, to a web application, and to store data on 3rd party servers or FEC servers, as it is entered into FECFile.

We were concerned that filers who opposed a web-based version of FECFile might be reluctant to be interviewed but wanted to give them an opportunity to share concerns anonymously. We conducted an anonymous web-based [survey](#) by sending the link to 8,906 unique email addresses. We received 533 responses. **You can read more about our methods in the methods section.** Our survey included 9 questions, seven of which were quantitative and two of which were qualitative.

### *Quantitative survey findings*

The majority of respondents (74%) said they would use a web-based version of FECFile if it were available (**Figure 6**). Just 9% of respondents said they would not use it, and about 20% said they would need more information (**Table 2**).



In general, the people who said they would use a web-based FECFile are already FECFile users, report small numbers of transactions, and are less likely to be compliance professionals. Of the people who said they primarily use FECfile, 84% said they would use the web-based FECFile if it were available. Those who said they would not use a web-based FECFile are more likely to use vendor software, be compliance professionals, and work for multiple committees, parties or PACs. This second group also submits reports with higher volumes of transactions.

**Table 2.** How different groups in our survey feel about using a web-based version of FECFile.

If a web-based version of FECFile were available, would you use it?	Yes 392 (74%)	No 49 (9%)	Tell me more 92 (17%)	Total N = 533*
<b>Role in filing</b>				
<i>Treasurer for a candidate, committee, party or PAC</i>	219 (56%)	22 (45%)	48 (52%)	289
<i>Consultant or compliance professional</i>	87 (22%)	19 (39%)	23 (25%)	129
<i>Candidate</i>	12 (3%)	0 (0%)	2 (2%)	14
<i>Other**</i>	74 (19%)	8 (16%)	19 (21%)	101
<i>Unanswered</i>	2 (<1%)	--	--	
<b>File on behalf of</b>				
<i><u>Single</u> candidate committee, party or PAC</i>	271 (69%)	28 (57%)	36 (39%)	335
<i><u>Multiple</u> candidates committees, parties or PACs</i>	109 (28%)	19 (39%)	52 (57%)	180
<i>Other†</i>	9 (2%)	2 (4%)	3 (3%)	14
<i>Unanswered</i>	3 (1<%)	--	1 (1%)	4
<b>Type of software typically used</b>				
<i>FECFile, the FEC's free software</i>	327 (83%)	23 (47%)	52 (57%)	402
<i>Vendor software</i>	57 (15%)	26 (53%)	38 (41%)	121
<i>N/A - I don't actually use the software or forms myself--someone else does it for me.</i>	3 (<1%)	--	2 (2%)	5
<i>Unanswered</i>	5 (1%)	--	--	5
<b>Volume of transactions reports typically include</b>				
<i>&lt;50</i>	207 (53%)	12 (25%)	36 (39%)	255
<i>51-300</i>	113 (29%)	19 (39%)	23 (25%)	155
<i>&gt;300</i>	65 (17%)	18 (37%)	33 (36%)	116
<i>I don't know</i>	5 (1%)	--	--	7
<i>Unanswered</i>	2 (<1%)	--	--	
<p>*Numbers may not total 100% due to rounding.</p> <p>**Unanswered (n = 2); Other (n = 99): Prevalent answers were assistant to a treasurer, PAC administrator, bookkeeper/accountant, and campaign manager.</p> <p>†Unanswered (n = 4); Other (n = 14): PAC, assistant, analyst, person responsible for ensuring reports are filed but not the person who actually completes them</p>				

### *Qualitative survey findings*

We asked filers to tell us more about why they would, or would not, use a web-based FECFile and what questions they would have about it. We also asked participants to tell us their questions about a possible web-based version of FECFile; many responses were not questions but comments about people's hopes and desires for the new system.

**Of the 392 respondents who said they would use a web-based FECFile**, 80 responded with substantive comments about why they would make that choice. 81% of respondents in this group expressed enthusiasm, and many mentioned specific capabilities that they are excited to see. The most common hope was that a web-based FECFile would let filers work remotely, followed by the hope that it would allow them to use a Mac computer. Many expressed excitement about the potential for collaboration on reports, cloud storage, or better access to the data. Several included stories about the current system crashing, resulting in lost data — along with hopes that a web-based system would prevent these losses in the future. A few respondents (6 people) had concerns with the security, speed, or performance of a web-based system, and one person (a current user of vendor software) expressed discomfort with the FEC having access to data before it is filed.

Of the respondents who would expect to use a web-based FECFile, 66 responded with substantive questions. 20% of comments evinced enthusiasm and impatience for a web-based FECFile, and 17% expressed desire for the new system to be more usable. About 11% of comments included questions about data security. About 10% of comments requested specific tweaks to the FECFile user experience. An equal number of comments (6%) asked that the new system make importing data easier or exporting data possible. A few respondents asked about accessing backed up data, whether the new system would still be free, whether users could use the old FECFile, and what would happen if the system crashes. Respondents also commented on how the new system might handle access control, after hours support, and improved instructions.

**Of the 49 respondents who said they would not use a web-based FECFile**, 22 responded with substantive comments about why they would make that choice. More than half (59%) of this group commented that they use commercial software instead of FECFile, so they do not care if it is web based. Five respondents would not be interested in a web-based FECFile because they only work on reporting when they are in their office. Three respondents commented that they prefer desktop software to web-based services because they believe they are superior in performance and reliability. In two cases, participants shared concerns about the FEC having their data on FEC servers before they would be ready to share it.

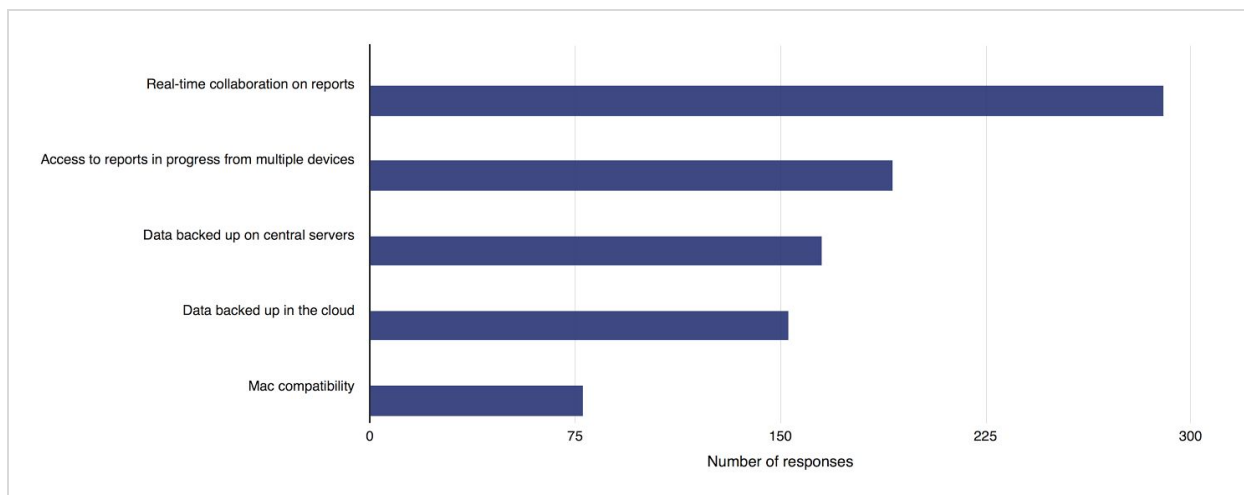
Only six participants who said they would not use a web-based FECFile responded with substantive questions. Two expressed concern that a web-based FECFile would infringe on

commercial campaign finance software vendor's market share. Other comments included concerns about the new system's data security, quality, and performance.

**Of the 92 respondents who said they would need to know more before deciding to use a web-based FECFile**, 56 responded with substantive comments. About one-third of the comments from this group of respondents said that they would need to know more about how data is secured and accessed before they could decide whether they would use a web-based FECFile. We intentionally provided very little information about how the system might be implemented, because we wanted to learn what concerns people had that the FEC would need to address in the system's design. 12% of comments raised questions about how the new system would integrate with their current way of doing things, and how they would access their data stored in the current FECFile. An equal number of comments also expressed questions about whether they could continue to be able to use the desktop version of FECFile. A handful of respondents echoed comments made by the "yes" group regarding capabilities they hope to see; specifically, an easy-to-use import function, the ability to set access controls, and an overall better user experience. Three respondents expressed fear that making FECFile web-based would negatively impact the speed and reliability of the system; others felt more comfortable with a web-based system that could back up their data to the cloud.

Thirty one of the respondents who would need more information before deciding to use FECFile responded to our survey with substantive comments about their questions. Two themes appeared in 23% of comments each: concerns about data security and a desire for the new system to make it easier for filers to import their data. About 13% of comments questioned whether the new system would come with improved help and/or support. A small number had questions about whether they could choose how to file reports going forward. Remaining questions were about exporting data from FECFile and how data would be transferred from FECFile to the new system.

Finally, we asked users to tell us which of the following capabilities they were the most excited about: access to reports in progress from multiple devices; data backed up on central servers; Mac compatibility; data backed up in the cloud; and real time collaboration on reports. Their answers are shown in **Figure 7**.



**Figure 7.** Highest priority potential capabilities of a web-based FECFile among survey respondents.

## B. Simplify and promote importing to reduce transcription errors

Most information in FEC filings comes from other tools or processes: financial and accounting software applications (such as Quickbooks), payroll application exports, bank statements, membership databases, or physical check registers. Filers using FECFile have two options for inputting this financial data: they can transform and import it or manually enter the data by hand.

Most users we spoke to expressed a desire to import their data but many did not know it was possible. (Those filers who do use the import feature are often more experienced.) This lack of awareness may be a result of the current FECFile user interface, which does not make it clear that importing data is an option.

For those who do know about the feature, preparing data for import requires substantial pre-processing to get it into a format FECFile will accept. Even experienced filers may avoid using it for small numbers of transactions. Even when imported, some data includes a manual component: data with parent-child relationships such as memo entries must be entered manually. For this reason, many experienced filers said they choose to enter all disbursements manually, even when it means inputting a relatively high volume of transactions one by one.

We designed a prototype to test our hypothesis that improving awareness and usability of the import feature would increase adoption of data importing, which in turn reduces the creation of duplicate records and manual data entry errors.

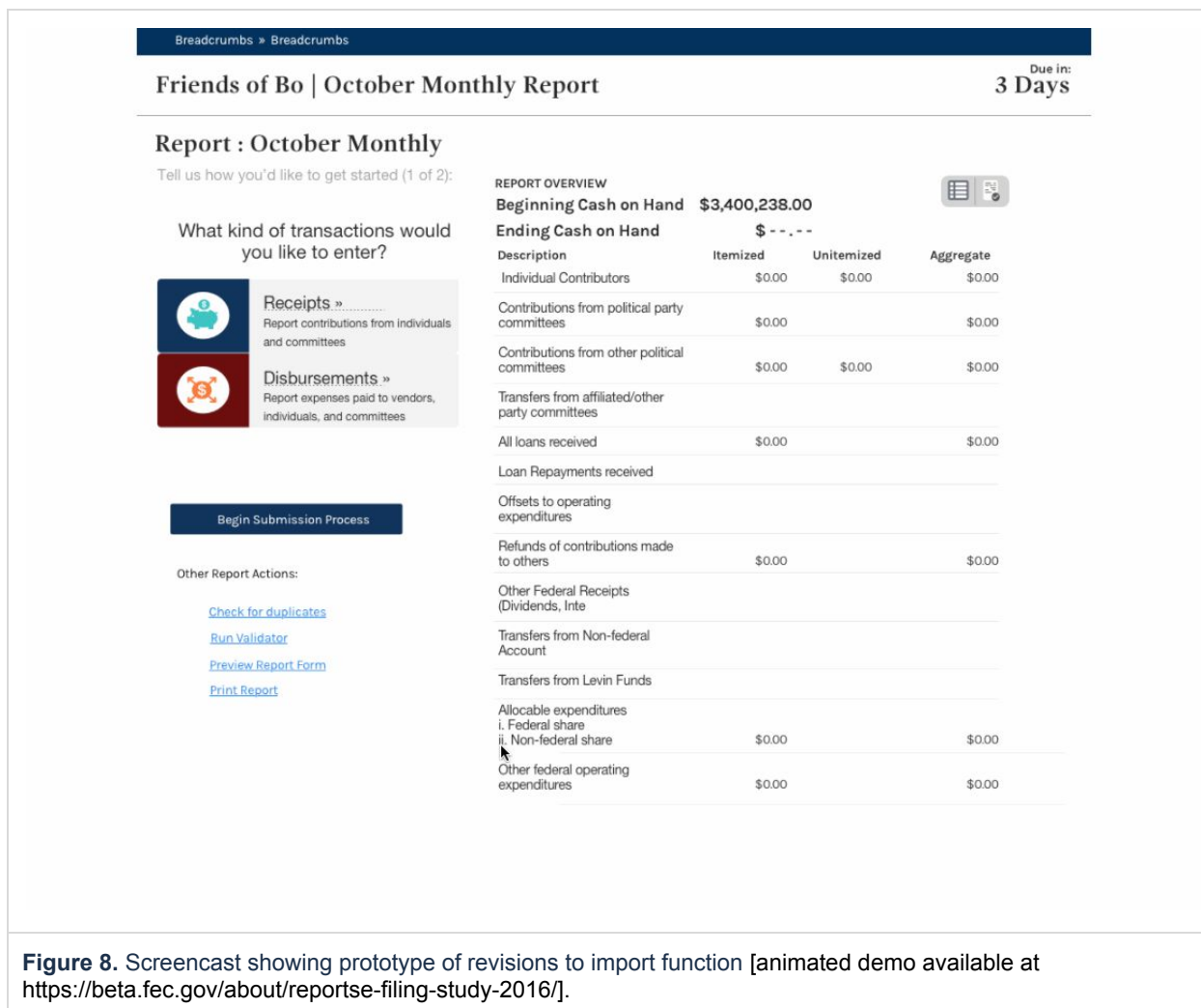
For our prototype, we required users to choose how to enter their data: on starting the process, users were greeted by a decision to enter manually or to import. After this, we observed the users for errors (wrong clicks) and reactions in attempting to proceed through the import flow. These tests revealed that by placing the choice of import or manual entry directly in line with the

data entry flow, users were able to initiate the import process, and that once in the flow, users were able to understand what was happening and complete the import with little difficulty: how to choose a file to import, how to choose which columns in their spreadsheet map to data elements in FECFile, initiate the import, preview the results, and confirm the success of the import. Many users specifically called out their preference toward the idea of mapping columns.

*“A lot of our current work is formatting the Excel file. If we were able to select from the import file that we have, that would be really great. I would like to be able to select the columns that I need from the spreadsheet.”*

*“That’s nice... So I guess if those [referring to a import file column] weren’t names I could say ‘oh no this is the employer or whatever.’ I like that very much. So if someone didn’t have the spreadsheet in the right order, it would make an attempt to understand what they’ve got. Now they give you a dummy sheet but they talk about it in such a convoluted way, it takes a phone call to RAD to understand.”*

*“My friend referred me. He opened a PAC for the Bernie Sanders campaign. He was selling t-shirts he was making to raise money for the campaign. I didn’t know importing was an option. That probably would have saved me a lot of time.”*



**Figure 8.** Screenshot showing prototype of revisions to import function [animated demo available at <https://beta.fec.gov/about/reportse-filing-study-2016/>].

### C. Provide an export capability

FECFile does not currently offer a way for users to export data into a spreadsheet form, which hinders the way they can manipulate data and their ability to transition between different systems and software. Adding this capability would help many users groups including filers who use FECfile and commercial filing software, the FEC Audit Division, and commercial software vendors.

Filers can sometimes lose their reports (in the case of a natural disaster or where their computer is lost or damaged) and without the ability to export from FECFile, they have to “rebuild” their entire database. Were the data exportable, filers could simply import it back into FECfile, which would prevent rekeying errors and save them time. In addition, filers could use an export feature to sort and manipulate their data in a spreadsheet, which would help them find and resolve errors and amend reports.



The need to manipulate report data in spreadsheet form is also shared by the FEC's Audit Division, and auditors have asked for an export capability in the past.

An export capability would help vendors easily transition FECFile users to their software by enabling them to download data from FECFile and upload it to the vendor's system, rather than rekeying it.

Finally, adding an export capability is a milestone on the way toward transitioning users from FECFile to a web-based system, as the FEC will need to have a plan in place for getting data out of FECFile and into the new system.

## Recommendation 3: Help those who help the filers

FEC staff help improve data quality and the filer experience by helping people file correctly. Some filers also take advantage of a robust vendor community that integrates with FEC tools to help filers. Vendors often offer filing as part of a broader suite of products and services that make running a campaign easier, so making it easier for vendors to integrate with FEC tools is an effective way to improve the experience for filers, vendors and data users. In our research, supporting those who help filers complete reports successfully emerged as an important means of ensuring accurate data.

**Support the developers both inside and outside FEC with open source software, documented in plain language**

### Open Source

Building the the e-filing system and related utilities using open source software could help build trust among users and vendors. Open source software would make it easier for commercial software vendors to integrate with FEC software and make it clear what changes are made to the modules the vendors use.

[The 18F Open Source Policy](#) identifies the main benefits of open source to be:

- Flexibility: you can customize existing libraries faster than creating your own software from the ground up.
- Community involvement: enabling continuous and broad peer review.
- Cost savings: increased flexibility and reuse can reduce development time, and shared packages can defray the costs of maintenance.
- Reusability: "By coding in FOSS (free open source software), we help populate a larger commons that cities, states, businesses, and individuals can participate in. This creates real economic value by lowering the burden of replicating similar work or by allowing the private sector to build off of and create new businesses around code..."

Additionally, there are cost savings because open source software doesn't have licensing fees. Database licenses can cost over \$500,000 a year.

## Plain Language

One of the value adds of openFEC, the API that supports [beta.fec.gov](https://beta.fec.gov), is that it presents the variables as words rather than clipped abbreviations that require experts to decipher. For instance, there are hundreds of data columns. While some abbreviations are intuitive, there are many cases where it is not -- for example, needing to know that the meaning of a particular variable name such as 'ind\_uni\_con' actually means individual unitemized contributions. This issue is compounded by the fact that the abbreviations are not consistent across database tables and views. Having full words as column and variable names makes the code easier to read. When the code is easier to read, it is easier to maintain.

In addition to clear, consistent naming, the codebase could benefit from more inline documentation. Documentation should describe the function of code and also the reason that the function is necessary. For example, if code is changed to fit a rare but important scenario, the condition that the code was written for should be added to in in-line documentation.

Having code with readable variable names and in-line documentation, along with some instructive, narrative documentation about how to start up and maintain the codebase, will make maintenance and onboarding developers to the FEC team more approachable.

Aside from the internal benefits, readable, consistent and well documented code creates output that is easier to read and data that is easier to use by individuals outside the organization. During October and November of 2016, the FEC website with the API documentation had 23,500 page views, each representing an individual attempting to understand aspects of the data.

## Support personal connections within the filing community

Filers often reach out to others for help in the filing process. Many seek help from analysts in the Reports Analysis Division, but we also heard stories of filers reaching out to others in the filing community or compliance professionals in their networks.

## Build on trusted relationships between filers and their RAD analysts

RAD is divided into four branches, two of which are responsible for reviewing reports. Each committee is assigned an analyst within the branch that serves their type of committee. We heard from a number of filers and RAD analysts that they develop trusted relationships over

time. RAD analysts who get to know the committees and the individuals who file on their behalf are able to help them more quickly and easily.

During business hours, filers can call the Reports Analysis Division and talk with their analyst to resolve an issue. However, [many reports come in late in the day on the day the report is due](#), and while RAD and EFO extend hours of support from 5:30 pm to 8:00 pm on major filing deadlines, filers sometimes still need support when RAD analysts are not available. We heard of instances where filers' inability to contact RAD resulted in not being able to file their reports, or receiving a fine—not to mention causing the filer stress.

*“Yes, first time I filed, it was late. I knew I had a deadline, but I left it till last minute. The day before it was due I went to submit, but it came up with all these error messages. And I didn't know how to fix it so it would go through. I called the FEC that night, but it was closed. So I got a late notice about a \$60 fine because it was late... It was pretty stressful. It was hard to find out info about what was going on. I really thought I had done it correctly. I was reading the guidelines but it still wasn't going through”*

In the example above, the filer looked for help only within the FEC, but several compliance professionals we talked to outside of the FEC told us they informally help one another and less experienced filers navigate the complexities of filing their reports.

*“There are always instances that come up—scenarios you haven't encountered before that, in our office, we kick around amongst the other reps and see if perhaps anyone else has encountered it and how they did it... I'd be hesitant just to take the advice kind of blindly. There are not a million of us that do this... So there are the usual suspects, and we all kind of bounce things off each other.”*

*“[Name of compliance firm], they've been doing it a long time and are very well respected, and we're all friends... so I'll shoot them an email and make sure we are all thinking about it the same way.”*

In response to a question about where to go for after hours help: *“I deal with a lot of different law firms, so I can call... I could always pick up the phone and call and say ‘I've got this, what do you think?’ For the average person that doesn't have those resources, shy of searching through advisory opinions on the website or reaching out to somebody else in the field, I guess you kind of do it the best you can.”*

We also heard from filers who work on reports as a side job during evenings and weekends when they cannot access the some types of FEC support.

*“There are times when I have to spend my entire weekend, 14+ hours each day, entering data to meet a filing deadline, and if there's some problem getting to the*

*website one of those days, I'm going to end up getting fined because I have a "real" job during the week."*

Filers could leverage the knowledge within their community to answer their questions without turning to the FEC for help, but trust is critical to information sharing within this community.

We tested one idea on how the FEC might foster these relationships and provide help after hours via a prototype. We prototyped a “community” feature in FECFile that would let users live chat with others online—both other filers and folks at the FEC—and that their conversations could be viewed by others using FECFile. (Figure 9)

The screenshot displays a web interface for the FECFile prototype. The main section is titled "Enter Transaction | Contributions" and contains a form for entering a transaction. The form includes fields for "AMOUNT", "TRANSACTION DATE" (set to 10-01-2016), and "ELECTION" (set to General '16). Below these are radio buttons for "INDIVIDUAL" (selected), "COMMITTEE", and "OTHER". The form also has fields for "LAST NAME" (Dolby), "FIRST NAME" (Thomas), and "SUFFIX", followed by "ADDRESS", "ADDRESS LINE 2", and "ZIP". To the right of the main form is a sidebar titled "Community" with a "Glossary" tab. The "Community" tab is active, showing "Recent conversations". The conversations include a post from Bfranklin asking about credit card transactions, a post from TJeffer asking about memo entries, and a post from JPolk asking about reporting a loan. Each post has a "see full thread >>" link.

Breadcrumbs » Breadcrumbs

## Friends of Bo | October Monthly Report

### Enter Transaction | Contributions

AMOUNT

TRANSACTION DATE: 10-01-2016

ELECTION: General '16

☒ INDIVIDUAL ☐ COMMITTEE ☐ OTHER

LAST NAME: Dolby FIRST NAME: Thomas SUFFIX: [dropdown]

ADDRESS: 32 Wallaby Way

ADDRESS LINE 2: Suite 3032

ZIP: [dropdown]

Full Name (Last, First, Middle Initial)

A. Mailing Address

City: [dropdown] State: [dropdown] Zip Code: [dropdown]

FEC ID number of contributing federal political committee: C

Name of Employer: [dropdown] Occupation: [dropdown]

Receipt For: ☐ Primary ☐ General ☐ Other (specify) [dropdown]

Election Cycle-to-Date: [dropdown]

#### Community :: Recent conversations

**Bfranklin:** We don't pay our credit card transactions all at once. How do I record this in memo transactions?

**TJeffer:** We report those though memo entries; the memo entries can exceed the amount of money in the initial disbursement. The memo entries are a way to let the FEC know what the vendor is for the information like credit card disbursements. [see full thread >>](#)

**JPolk:** My committee received a loan for \$300, how do I report that?

**JJett:** Loans are recorded as debt. Report debts and obligations continuously if they exceed \$500 or have been outstanding 60 days or more. 11 CFR 104.3(d) and 104.11.. [see full thread >>](#)

Participant's responses to this feature were positive, with a few caveats. While the filers who tested this prototype were enthusiastic about the ability to connect with others inside of FEC, many were not comfortable with letting just anyone respond to their questions. Participants said they would not trust answers from unknown respondents and some felt that only experts within the FEC should be allowed to respond to questions in FECFile. This is evidence to the filing community's desire for increased access to RAD and others at FEC.

The FEC is already working to meet this need with the Correspondence Tracking System, which will allow filers to identify and contact their RAD analyst via email using a contact form on the website. The system will also help ensure that filers get consistent answers. RAD keeps a library of responses to common questions and the tool will allow the Analysts to easily respond with approved language.

Our research points to a clear need for support that filers can trust. The FEC's customer Correspondence Tracking System will enable filers to submit questions at any hour, but their need would further be met if the this system were staffed by RAD analysts during peak filing times outside of normal working hours.

### **Build trust among members of the filer community**

Through our prototyping activity, we learned that filers are enthusiastic about learning from others in the filing community, but that they need to be able to trust the source of the information. Through prototyping, we learned that filers believe communication in an FEC forum (such as within FECFile or the FEC website) should be between filers and the FEC. And although we also learned that participants had concerns about the trustworthiness of filing advice from people outside the FEC, opportunities to enhance information sharing among filers out in the community do exist.

Research on trust in computer mediated communication shows that the design of an information and communication system can influence users' community trust by providing visibility into community members' characteristics and behavior. If filer's verifiable characteristics, such as that of being a board certified attorney, were visible to others in the filing community via a Google group or listserv, it might promote the development of trusted relationships among filers.

Making filers and analysts' behavior more public could also aid in the development of trust within the community. Filers who tested our prototype were excited about the possibility of being able to see and search other filers' conversations with RAD as a way to learn from others experience before calling RAD themselves. Making live chat transcripts available to filers would make their behavior more visible to others in the community and thereby enable community members to develop knowledge-based trust based on community members past behaviors. Additionally, filers might be able to resolve their questions by reading about how similar issues were handled without picking up the phone, and could therefore reduce the burden of phone calls on their RAD analyst. However, a building block in the trusted relationships that filers develop with their analyst is the fact that analysts keep filers' information confidential. It is possible that filers would not want to publicize conversations with their analyst, so whether transcripts are shared should be optional. The planned Correspondence Tracking System will achieve this by giving analysts the ability to suggest that an answer they gave be published as a "solution" that other filers can view.

Finally, by fostering trusted relationships among filers who have different levels of experience, the FEC could alleviate some of the burden on RAD. Experienced filers are proud of the workarounds they have created that help them file their reports efficiently, and in several cases they talked about how they are happy to share this knowledge with others in the filing community. The FEC could help filers like the one who could not get help after hours by helping them connect with others in the community.

## Recommendation 4: Ensure future filing reliability

While our study revealed ways that the FECFile interface and software could be enhanced, filers and vendors articulated few concerns with the overall e-filing infrastructure's reliability. This is praiseworthy: the FEC has, through evolving practice, timely fixes and sustained attention, kept the system working under substantially increased load. In coming years, the FEC will need to support increasing volumes of data while maintaining performance and reliability. To do this, the FEC needs to expand existing work to streamline data flow and build for scalability in the long-term.

### Moving to cloud infrastructure

#### Invest in cloud infrastructure that can scale to meet demand

Cloud infrastructure refers to hardware that is not owned or maintained by the FEC and can add or subtract resources with minimal effort. While it does require a concerted effort to retool systems and processes for the cloud, the benefits outweigh the costs.

First, the FEC has a real need for scalable infrastructure. The FEC's traffic, and e-filing in particular is bursty but predictable. The flow of data is driven by filing deadlines that are known in advance. A sizable number of filers wait to get as close to the midnight deadline as possible (see figure 12 below). So, instead of buying and continuously maintaining software at the peak capacity, the FEC can scale down its infrastructure and pay less the majority of the time when traffic and processing requirements are low.

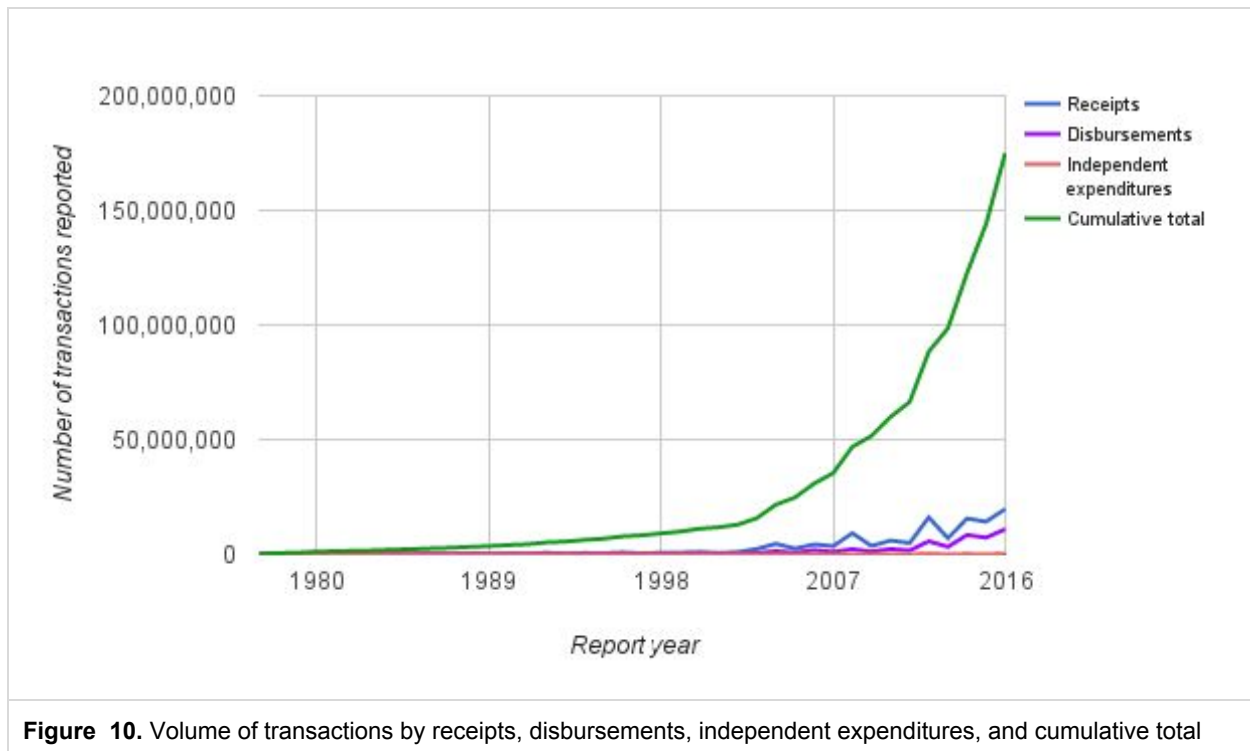
Perhaps more important than cost savings, cloud infrastructure would give the FEC the flexibility to scale up computation and memory capabilities during crunch time. The ability to scale infrastructure as needed is especially helpful if the growth of data outpaces the projected workflow. In the case of e-filing, even the most accurate projections of future load can be made instantly irrelevant by legal rules outside of the FEC's purview. Legislation and court cases can change or redefine the filing process at anytime in a way that can never be predicted in a five year plan that determines the number of servers to buy.

In the chart below, you can see the growth in receipts, disbursements, and independent expenditures since the seventies. These transactions are the most numerous records by transaction type. There were over 30 million records produced in 2016.

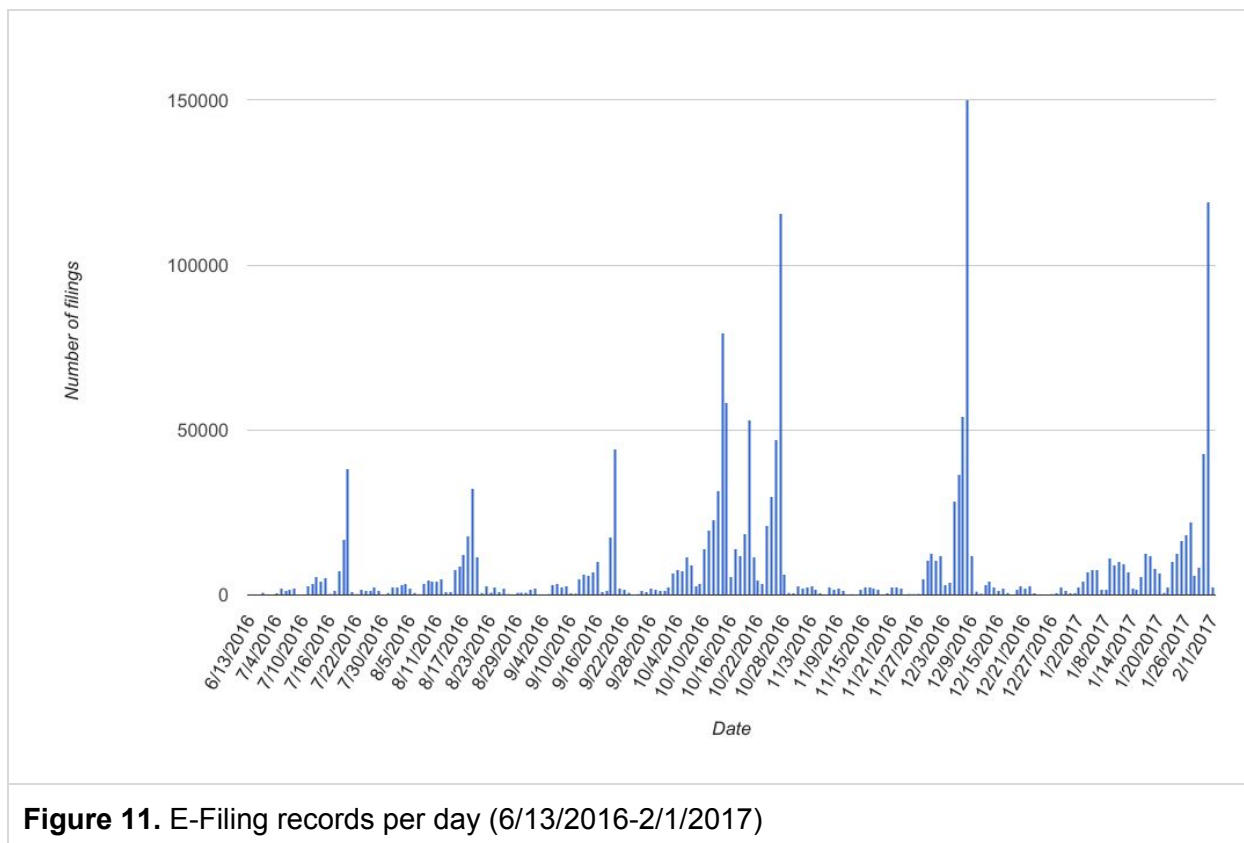
Figure 10 shows how the growth in transactions causes even larger growth in the cumulative amount of data that needs to be cared for. There are about 175 million of these records now,



and the data is growing exponentially. These records are being kept, totaled, and analyzed in multiple places. In terms of storage, the EFO office estimates that it has about 7 TB worth of PDF and image files and a database of about 350 GB.

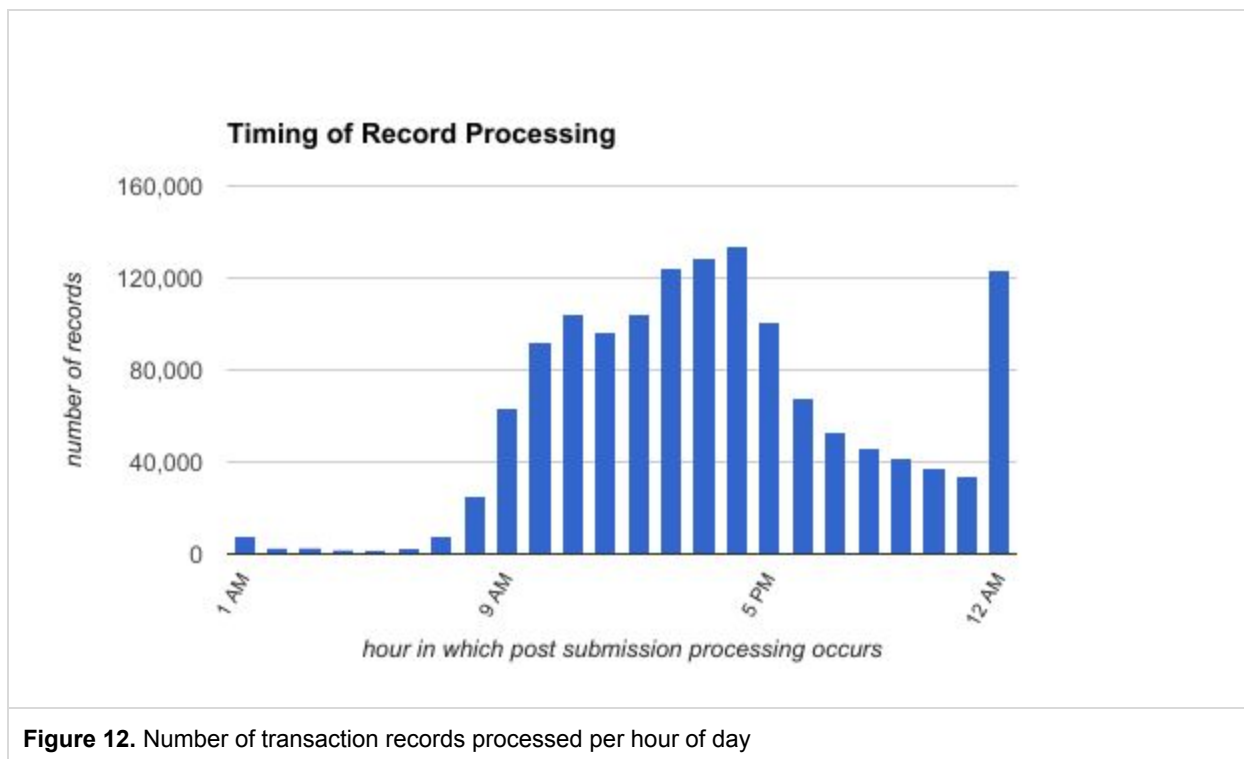


Another striking attribute of the traffic is that it comes with tremendous peaks and valleys, which you can see below in the figure showing E-filing records per day (**Figure 11**). Traffic peaks on filing deadlines and the rest of the time it is low. When you buy and maintain your own hardware, you have to purchase the computational and storage capacity for peak times, even though you don't need all of the capacity most of the time.



**Figure 11.** E-Filing records per day (6/13/2016-2/1/2017)

In Figure 12 below you can see when e-filing summary reports are processed. They create a normal distribution of records for most of the day, peaking at 3:00pm before a spike in traffic around midnight.



As inputs grow, increasing on-demand and parallel processing can maintain or improve the timeliness of making valuable data available to the public.

Moving to the cloud is also a way to unify existing infrastructure. Currently, data is siloed in different databases that cannot easily access each other's information. **Figure 13**, below, shows the current flow of data from the filers through the servers and databases where it is maintained. While this structure makes sense for private data such as payroll, it makes less sense for less sensitive information.



Further, the FEC's could leverage existing cloud infrastructure that supports Beta FEC to expand e-filing infrastructure. The FEC's API and the new FEC website are built in the cloud. These webapps leverage Amazon Web Services and Cloud.gov's platform as a service (PaSS).

General cloud migration issues and logistics will be discussed further in the upcoming 18F cloud study.

## **Changing underlying form design and processes to increase data quality**

Like many regulatory agencies, the FEC's data collection has its roots in the paper forms that have traditionally transmitted that data. While paper forms can be helpful models for understanding the data, only thinking of the data as it appears on a paper form can hide the data structures that the forms create. That structure can impact the data intake, presentation and how easy or hard the data is to use. As the volume of data grows, data users can't read all the paper filings and are more reliant on searching, filtering and linking data to find answers.

Shifting towards a data-centric model can take many iterations, but will ultimately enable a much more flexible and sustainable system. In a data-centric model, forms aren't the end product, they are simply tools for collecting specific data. While there may always be a paper representation of the data, the guiding principal when crafting what is primarily an electronic data collection system should be to focus on the best way to automatically collect and maintain the data rather than on the manual human approach that the paper forms represent. This concept also came up in the 2013 e-filing study, recommending a transaction focused, rather than form focused approach, and the FEC already has a working foundation for this approach in the current paper processing system.

### **A. Improve the process by which form changes are designed and implemented**

We recommend that any form changes be based on evidence that the change will have a positive impact on filer experience and data quality. Evaluation by real users can deliver evidence of a change's impact on filer comprehension and data accuracy, among other things, before the change is implemented. A form change is not only an opportunity for a new look on a piece of paper—it is also a chance to better ask for the needed information. Usability testing can also help the forms committee make more informed decisions about information architecture and wording. For example, if there were a few legally correct options for instructions on how to fill in a line item, testing with filers could help uncover which labels were more likely to solicit the correct response.

The FEC faces challenges to changing forms because it doesn't always have control or warning over when new legislation or judicial decisions require collecting new or different data.

Some of these changes are easy to implement, while others need regulatory changes or legislation to be enacted.

- **Candidate ID for independent expenditures** One of the most consistent requests from data journalists was to add candidate IDs for independent expenditures. This would allow for FEC data users to attribute expenditures to the candidate that the group is supporting or opposing. Currently, the same name can be represented multiple ways: nicknames are common, various honoraria may prefix the name, or a suffix can be attached. On top of that, human error makes it is easy to misspell names. Concurrent members of Congress can even have the same name, such as Rep. Mike Rogers (MI) and Rep. Mike Rogers (AL). These inconsistencies makes the name hard to use in an automated way. Having an ID is a way to programmatically identify candidates in a reliable way and allow for timely data aggregation and reporting. This could be implemented through a candidate lookup field that is populated from the FEC's API, with consistently spelled candidate names that can provide the ID automatically, taking that burden off of the filer in input and off of data users in analysis.
- **Candidate ID when applicable on leadership PAC form** Leadership PACs are not legally bound to the candidate, but are often used as a tool leveraged by candidates to gain influence with other members and gain important committee positions. Because of their importance, Beta FEC website usability testing showed many users wanted leadership PACs information linked to the candidate. The way to add and link that information in a structured way is to add the leadership sponsor's candidate ID, when applicable. This change could be added to the Statement of Organization on the Form 1, or the IDs could be added retroactively in the coding process: from 2012 to 2016, there were fewer than 700 of these PACs created, and not all of them have candidates to code.
- **A place for a Vice President's name** - For presidential candidates, their running mate is an item required to be collected by regulation, yet it does not have a unique form field. This leads to people entering both names into one candidate name field on Form 2. To fix this, the FEC can add a space on the Form 2 for this field, or create a separate post-convention form for this purpose.

## B. Amendments

Amendments are currently processed in a way that often results in the re-submission of entire forms. This can make it hard to identify and track which information is being amended.

The current system processes paper amendments through a series of inferences. These can include what type of filing the amendment is, or when the amendment came in. In some cases, there is not a way to properly track amendments. When evaluating the chain of amendment documents, 18% of reports lack information that enables data users to tell if the document has been amended or if it is amending another document. Ninety eight percent (98%) of the documents with unknown amendment status (amendment or new filing), are paper amendments.

Further improvements can be made by scoping e-filing amendments to include only the information that needs to be amended, rather than replacing the data wholesale. In the current system, one or two small corrections on a large report can result in thousands or even a million new records. A report might be amended multiple times, multiplying the number of records to be maintained with each submission. This creates unnecessarily large data sets, which require more processing and maintenance, without adding much new information. The FEC's 2013 e-filing study estimated allowing users to file amendments that include only the transactions that changed would reduce filing volume by about 70%. Although it would require a regulation change, scoping e-filing amendments only to include the information that needs to be amended, would enable filers to identify the transaction that needs to be changed or deleted, and then pass along only the information to be changed or deleted for processing.

Additionally, and in the short term, adding form IDs to amendments that link them to the document that the amendment is intended to replace would make the amendment chain more clear to data users. This would enable filers and data users to use the data from the amendment and the existing filing to see what the current data looks like as a whole.

### **C. Child Transactions: memo entries and earmarked contributions**

Memo entries and earmarking represent two special case transaction types which, under the current reporting system, are sources of confusion for filers and data consumers.

An earmarked contribution is a contribution the contributor gives to a second party with instructions to give the contribution to a clearly identified candidate or candidate committee. Disclosing these types of transactions requires up to four itemized transactions: the original receipt to the earmarking committee, the disbursement from the earmarking committee to the intended recipient, a receipt to the intended recipient, and a receipt memo entry to the intended recipient. The memo aspect disclosing the earmarking committee of the receipt reported by the intended recipient committee can be part of the receipt transaction that discloses information about the contributor. Representing that transaction twice; once as a receipt as it is coming in to the first committee, and once as a receipt coming to the final committee, would be an

improvement. The “memo” aspect of the transaction can be part of each transaction. (See an example earmark receipt in this page of a [report](#).) This would require a change to regulation.

There are additional cases where the itemized data that the same transaction is reported twice—once to show that it received the money, and then a separate line to give context to that transaction. The duplicate transaction is marked as a “memo”, much as earmarks are. These transactions should also be unified as a single transaction with more contextual data attached. There are also cases, such as a disbursements with credit card payments, where there can be multiple child transactions.

The memo approach also adds additional overhead for people filing the reports.

*"I get the data in import. Because they are consistent in their names, I import their disbursements. I spend a bunch of times doing memo entry, no way to import parent child, those are done by hand because there is no way to import child transactions. I was told that was going to be fixed."*

Aside from creating overhead for filers, this process creates data that is more difficult for new data users to understand, as half of the transactions needs to be discarded to create an accurate total. Additionally, not all transactions with memos are “memoed items”, and additional text searches and filtering is needed to calculate totals correctly. To see an example of how this is handled on the current website and API, see the [code here](#).

Duplicative records also causes problems for data users because filers can create data that makes sense in their context, but makes the data harder to use. In the duplicative record, filers often write “see above” when dealing with the purpose in these duplicative entries. People using this data run into problems because the order or filtering of a particular query often produces results that don't preserve the original ordering. Data users then have to look up the previous record relative to that record as it would appear in a data entry system or paper representation. It also makes automated analysis harder when the data you are interested might live in a different record.

To make earmarking more clear, the “memo” concept, with its duplicate entries, should be replaced with one transaction per schedule that includes the donor and conduit information. This would be clearer to filers and data users. Additional user research with prototype testing should be conducted to determine how to make earmarks and contextual information more user friendly.

#### **D. One field, one purpose**

For some fields in FECFile and on the paper forms that underlie them, a single field with an identical label can be meant for very different data elements. For example, the “name” field



(below) represents different things, depending on the context: when the row represents a conduit committee, “name” refers to the committee name, but when the row represents an individual, “name” refers to the contributor’s name.

**Figure 14** shows how the fields that serve dual purposes look on the paper form:

Full Name of Individual (Last, First, Middle Initial) or Full Organization Name

**A.**

Mailing Address

City State Zip Code

FEC ID number of contributing federal political committee. C

Name of Employer (for Individual) Occupation (for Individual)

Receipt For:

☐ Primary ☐ General

☐ Other (specify) ▼

Aggregate Year-to-Date ▼

Date of Receipt

MM / DD / YYYY

Amount of Each Receipt this Period

☐ Memo Item

**Figure 14.** Example of form fields that serve dual purposes on a paper form.

While the data elements meant for the “name” field might seem similar, the meaning is actually quite different when one names an organization and the other names a person. This pattern of field reuse is employed in several places, including committee name and candidate name, committee ID and candidate ID, and the memo field (where the same form field can contain either a specific attribute or free form notes). When data are stored this way, data users cannot find and use information without creating complex rules or filters—those less technically capable may be excluded from using these data.

We recommend that each form field should only capture one, tightly scoped type of information.

## Paper filings processing opportunities

In 2016, electronic filing made up 97.8% of all reports filed (by page volume). The remainder of the filings were filed on paper, and while this represent a small percentage of the total, it still represents a tremendous volume and processing burden on the FEC; as of October 30, 2016, it represented over 535,000 pages of transactions for the year. The vast majority of this volume (91% in 2016) originate from campaign committees involving senators. FECA dictates that these filers must file their reports with the Secretary of the Senate, who is then required to remit these filings to the FEC within two working days of receipt ([52 U.S.C. §30102 :G2](#))<sup>3</sup>. The scope of this study did not involve the Senate or its filers; few of the details of the Senate’s process are clear, save the parts that intersect with the FEC.

<sup>3</sup> 52 USC 30102 : g.2: Organization of political committees,  
[http://uscode.house.gov/view.xhtml?req=\(title:52%20section:30102%20edition:prelim\)](http://uscode.house.gov/view.xhtml?req=(title:52%20section:30102%20edition:prelim))

1. While the Secretary of the Senate receives the filings as paper, it is likely that the vast majority of these filers are using either FECFile or vendor provided software to complete the filings, and inspection of a sampling of the forms indicates that they were not handwritten.

*“All the senate people [are] not out there printing paper and filling it in by hand. There is no reason they couldn’t e-file if there was a system in place to support it.”*

*- Stakeholder familiar with the process*

2. The senate reports arrive at the FEC as scans of paper pdfs, through a secure FTP process.
3. From here, the TIFs are converted to PDFs and stored on a webserver and are then “indexed” for basic metadata to associate them with the committees and make them otherwise findable.
4. At this point, document images enter the paper automation process, where they are sent to a contractor who “shreds” the document, a process that divides the document into individual fields for machine interpretation through Optical Character Recognition (OCR) or field-level transcription by contract data entry clerks.
5. The output of this process is an e-file (.fec) manifestation of the paper filing that is then coded and processed into the master record, as the files produced by the standard e-filing process are.

The current paper automation process described above replaced one that did not make use of OCR, but rather relied on contract data entry clerks who worked on the whole document, rather than by “shredded” fields. This process upgrade has yielded substantial improvements to the speed of processing—the previous process took closer to a month before the transaction level data was publicly retrievable. The new process is much faster, but still creates many data quality issues—OCR and contract data clerks dealing with “shredded” documents seem more likely to misinterpret an individual field than those dealing with an entire filing. That said, the process continues to evolve, and recently, the forms were optimized to allow OCR software to better capture the data on the paper.

While improvements to the paper process continue and the overall percentage of paper filings volume remains minor, problems continue to affect them, including timely availability of publicly accessible data, substantially increased costs borne by the FEC, and the introduction of substantial errors affecting all members of the ecosystem.

### **Delays in the data’s availability to the public**

The paper process is substantially more time-intensive than the electronic filing process, which is nearly instantaneous. Paper files are remitted from the office of the Secretary of Senate, often around five days after receipt (regulations stipulate files should be remitted within two working

days <sup>4</sup>to the FEC); from here the scanned pdfs are pushed to the server the morning after they arrive at the FEC, followed by a process of “Indexing” whereby the committee information and summary finances are hand-entered; this can happen as soon as the next morning but must (and does) be made publicly available within two business days of receipt. After this process, the scanned pdf images are placed in the paper automation process, and take an additional five days, or sometimes longer, before the transaction level data is made available to the FEC, at which point the process mirrors that of the electronic filing process.

### Increased costs borne by the FEC

On top of the delay in making this information public, the costs to processing paper are substantial. In 2016, direct contracting costs for processing the paper was \$711,000, up from \$681,000 in 2015 and projected to rise at least 3% per year, totaling \$3.7 million over the next five fiscal years (**Table 3**). This number almost certainly substantially underestimates the costs involved, as it does not account for the additional staff time needed to track down and reprocess errors, index and code paper forms, and otherwise troubleshoot a substantially more complicated process.

**Table 3.** Projected costs of paper automation contract

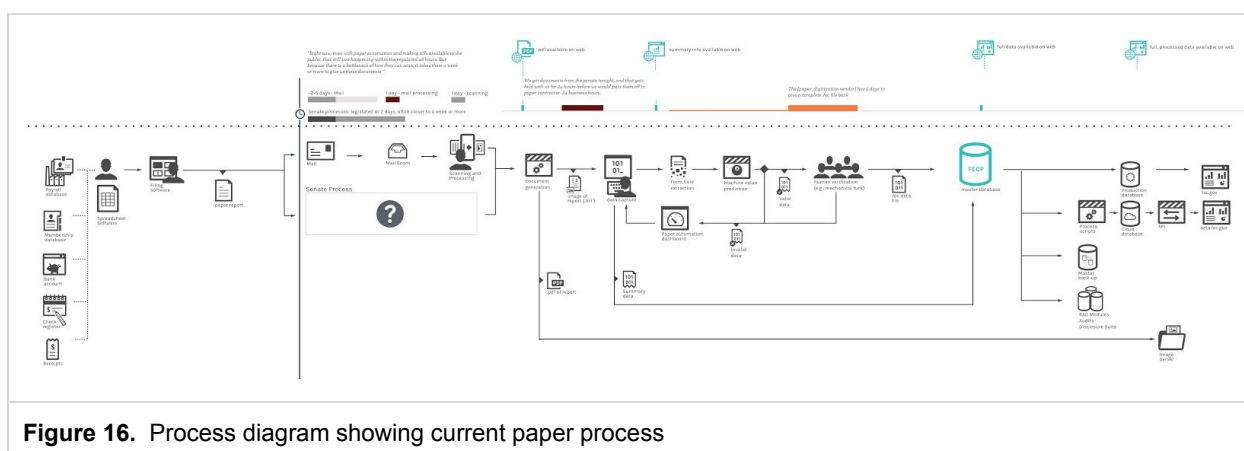
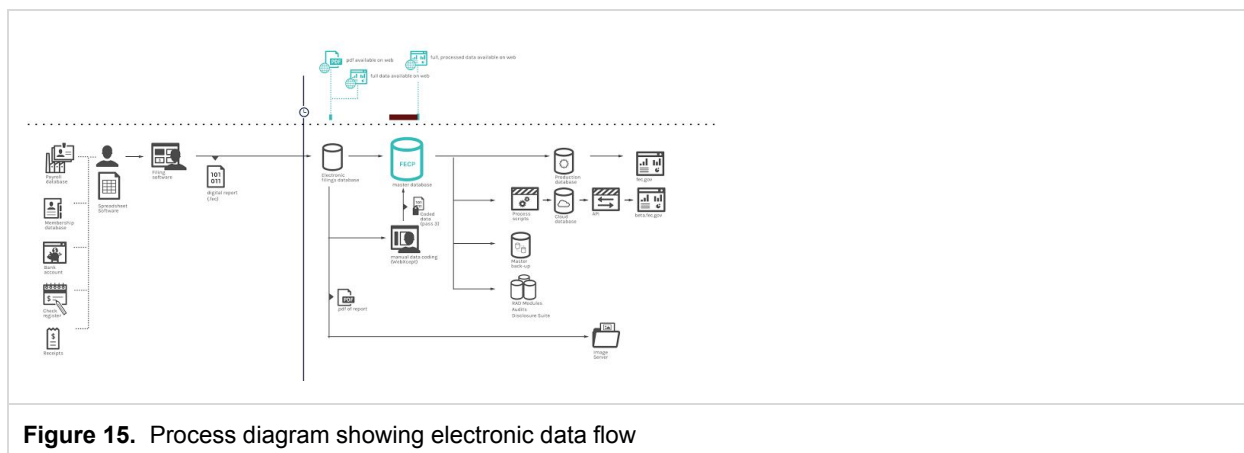
Fiscal Year	Projected cost of paper automation contract
Fiscal Year 2016 - 2017	\$711,580.60
Fiscal Year 2017 - 2018	\$732,928.02
Fiscal Year 2018 - 2019	\$754,915.36
Fiscal Year 2019 - 2020	\$777,563.31
Fiscal Year 2020 - 2021	\$800,889.67

### Increased errors affecting all

On top of being considerably more costly, paper filing also is a significant source of errors. The technology behind OCR software’s ability to recognize and encode a printed character is far from perfect, and the and our conversations across the agency have indicated that the current process could still be improved.

*“In October [2016], over 1,400 documents were flagged for repair.”  
- Stakeholder familiar with the paper process*

<sup>4</sup> 52 USC 30102 : g.2: Organization of political committees,  
[http://uscode.house.gov/view.xhtml?req=\(title:52%20section:30102%20edition:prelim\)](http://uscode.house.gov/view.xhtml?req=(title:52%20section:30102%20edition:prelim))



## Potential Opportunities

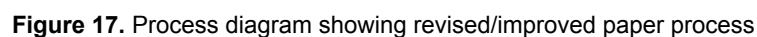
The FEC has previously made recommendations to Congress to amend the legislative requirement, and while we think that should continue, a couple of other opportunities to make headway in the short term have presented themselves.

### 1. Develop FECFile version 2 as an open source and API driven application

A large portion of the paper filing is from the Senate. FECA states that filings must be made to the Secretary of the Senate, who will forward them to the FEC. Currently, this process is done via paper, but the law doesn't state that it has to be paper, only that it has to be filed with the Secretary of Senate. While the FEC should continue to petition Congress to change the rules such that all campaigns should e-file, it is not in a position to change the way that the Secretary

## 2. Make the paper machine friendly

Translating a digital data artifact to a physical one and back is a situation that numerous other fields have wrestled with. Consider the modern airline industry: today, most passengers book online, where the transaction data begins as digital. On the date of the flight, passengers print their boarding pass, either at home or at an airport kiosk—creating a paper manifestation of their digital transaction. This is then presented to the gate attendant, who scans it in order to confirm the passenger's ability to board the plane. The gate attendant does not attempt to scan the printed characters of the boarding pass to determine the passenger's eligibility to board—instead the airline encodes and then decodes the data as a bar or QR code, as this has a much greater likelihood of being read properly by the scanning system. If a similar process were employed here, creating machine-friendly paper filings, campaign finance data could be made available much faster, perhaps even approaching the speed of purely electronic filings.



Image# 201611189037214038

# **SCHEDULE A (FEC Form 3)** **ITEMIZED RECEIPTS**

Use separate schedule(s)  
for each category of the  
Detailed Summary Page

FOR LINE NUMBER: (check only one) PAGE 5 OF 20

☒ 11a ☐ 11b ☐ 11c ☐ 11d ☐ 11e ☐ 11f ☐ 11g ☐ 11h ☐ 11i ☐ 11j ☐ 11k ☐ 11l ☐ 11m ☐ 11n ☐ 11o ☐ 11p ☐ 11q ☐ 11r ☐ 11s ☐ 11t ☐ 11u ☐ 11v ☐ 11w ☐ 11x ☐ 11y ☐ 11z

Any information copied from such Reports and Statements may not be sold or used by any person for the purpose of soliciting contributions or for commercial purposes, other than using the name and address of any political committee to solicit contributions from such committee.

NAME OF COMMITTEE (In Full)  
Shaban for Congress

Full Name (Last, First, Middle Initial)  
Springfield Republican Town Committee

Mailing Address 60 Evergreen Terrace

City Springfield State CR Zip Code 06610

FEC ID number of contributing federal political committee. C

Name of Employer Occupation

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 250.00

Amount of Each Receipt This Period 250.00

Memo Item

Full Name (Last, First, Middle Initial)  
Buckley, Jeff

Mailing Address 248 Evergreen Terrace

City Springfield State CR Zip Code 06883

FEC ID number of contributing federal political committee. C

Name of Employer Occupation  
Retired

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 300.00

Amount of Each Receipt This Period 100.00

Memo Item

Full Name (Last, First, Middle Initial)  
Johnson, Robert

Mailing Address 1 Evergreen Terrace

City Springfield State CR Zip Code 06855

FEC ID number of contributing federal political committee. C

Name of Employer Occupation  
Bicycle Records Blues Musician

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 300.00

Amount of Each Receipt This Period 200.00

Memo Item

SUBTOTAL of Receipts This Page (optional) 550.00

TOTAL This Period (last page this line number only)

FEC Schedule A (Form 3) (Revised 05/2016)

Image# 201611189037214038

# **SCHEDULE A (FEC Form 3) ITEMIZED RECEIPTS**

FOR LINE NUMBER: (Use separate schedule(s) for each category of the Detailed Summary Page)

check one: ☒ 11a ☐ 11b ☐ 11c ☐ 11d ☐ 11e ☐ 11f ☐ 11g ☐ 11h ☐ 11i ☐ 11j ☐ 11k ☐ 11l ☐ 11m ☐ 11n ☐ 11o ☐ 11p ☐ 11q ☐ 11r ☐ 11s ☐ 11t ☐ 11u ☐ 11v ☐ 11w ☐ 11x ☐ 11y ☐ 11z

Any information copied from such Reports and Statements may not be sold or used by any person for the purpose of soliciting contributions or for commercial purposes, other than using the name and address of any political committee to solicit contributions from such committee.

NAME OF COMMITTEE (In Full)  
Shaban for Congress

Full Name (Last, First, Middle Initial)  
Springfield Republican Town Committee

Mailing Address 60 Evergreen Terrace

City Springfield State CR Zip Code 06610

FEC ID number of contributing federal political committee. C

Name of Employer Occupation

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 250.00

Amount of Each Receipt This Period 250.00

Memo Item

Full Name (Last, First, Middle Initial)  
Buckley, Jeff

Mailing Address 248 Evergreen Terrace

City Springfield State CR Zip Code 06883

FEC ID number of contributing federal political committee. C

Name of Employer Occupation  
Retired

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 300.00

Amount of Each Receipt This Period 100.00

Memo Item

Full Name (Last, First, Middle Initial)  
Johnson, Robert

Mailing Address 1 Evergreen Terrace

City Springfield State CR Zip Code 06855

FEC ID number of contributing federal political committee. C

Name of Employer Occupation  
Bicycle Records Blues Musician

Receipt For: 2016  
☐ Primary ☒ General  
Other (specify) ▼

Election Cycle-to-Date ▼ 300.00

Amount of Each Receipt This Period 200.00

Memo Item

SUBTOTAL of Receipts This Page (optional) 550.00

TOTAL This Period (last page this line number only)

FEC Schedule A (Form 3) (Revised 05/2016)

Image of report

Image of report with QR code

**Figures 18 & 19.** Comparison of report images with and without QR codes. This QR code includes the committee and contributor information visible on the page, augmented **only** with the Committee ID and identifiers to link the data to the filing, schedule, and page, in a machine readable format.

### A. Streamline variables as much as possible

54

many different forms ask for the same or similar information. Data in each form, both inputs and variable names, should be as consistent as possible.

Form fields and definitions should also be consistent across platforms; paper, FECFile, and published resources should all have the same information about form fields.

As a step toward improving the quality of the data received through e-filing, FEC data experts have already begun looking into how they might address duplicative fields that cannot be standardized by improving metadata within the e-filing system. For example, FEC data experts have analyzed all transactions that can be reported on form 3 and form 3x, marking the differences and similarities as a starting point for reworking the FEC electronic filing formats and standards and specifications. Reworking filing formats, standards, and specifications in this way would also give the FEC greater flexibility with implementing reporting changes caused by court cases or new legislation.

In one instance uncovered by their analysis, FEC data experts identified at least 13 distinct itemized transaction types that are summed together on line 11(a)(i) (itemized individual contributions) which are identified only by specific words and phrases used to describe them rather than any specific unique attributes of the transactions themselves. These include straightforward contributions from individuals but also more complex transactions such as earmarked, in-kind, and partnership contributions as well Indian Tribe contributions and redesignations and attributions of individual contributions. Many other summary lines have similar issues.

Many of the characteristics of these transaction types can be traced back to the days before electronic filing when the standard of being clear on the public record meant understandable to a person looking at a printed page. Today, most reports are being filed as data and the legacy of paper forms has created eccentricities in the electronic filing system which make the data harder to use by developers and analysts. By adjusting the electronic filing formats to produce distinct transactions, these transactions would be machine readable without any special handling. Automated data handling both within FEC data processes and by outside data users could be vastly improved by reducing ambiguity which requires manual categorization to mitigate.

## **B. Unique, stable identifiers**

The FEC data model does not reflect an accurate concept of a person running for office. Each time someone runs for office, they are issued an ID specific to the state and contest in which they are running. Events—such as running for a different office or in a different state—create multiple IDs that can fragment records, and linking those records together again can be challenging. As mentioned before, variation in names (i.e. nick names), as well as multiple



people with common names (i.e. Rep. Mike Rogers - MI and Rep. Mike Rogers - AL) makes unifying the records a manual process.

In usability testing for beta.fec.gov, we have consistently seen that people who use the site would like to track data for candidates, even across campaigns. Without person-based IDs for candidates, this is not feasible. Other organizations, such as the Center for Responsive Politics, the National Institute on Money in State Politics and many more use this data to create identifiers for individual candidates and then correlate each candidate's various IDs. This is a valuable service, but could be better addressed at the root.

Changing candidate IDs also contributes to data errors. Filers might reference an older ID without realizing it changed, or because the information was saved in an auto-fill feature.

The original intent of this ID scheme was to conserve storage when it was more expensive. However, now that the FEC is handling a higher volume of data and more characteristics, there is increased risk of data collisions. (For example, much of the data from the 1970s cannot be used with data after the 1980 cycle because the IDs are not unique and cause records to be mismatched.)

Another instance of this is in filing IDs, which will have dates and information about if things were filed via paper or electronically. This is data that should be represented in the data model directly, rather than inferred by ID. The API and Beta FEC website already try to decode this information and store that the date and time it was processed and how it was processed in its respective data models. Even if the ID scheme does not change, this information, should be stored as data in data fields to make the data modes more straightforward.

Currently, characters of the ID are representing data. A better way to generate unique IDs is to use the [UUID standard](#), this would be beneficial for transactions, files and other IDs that need to be unique. These IDs are long so they are less useful for the kind of IDs people need to transcribe repeatedly like candidate ID.

For candidate and committee IDs, there are some good usability arguments to be made that starting all committee ids with a "C" can help people identify that ID. The current convention is to start a candidate ID with a "P", "S" or "H" depending on office, this helps identify it as a candidate ID but causes problems if candidates change office. So the letter could stay static even if there is a change in office, or a letter other than "C" could be chosen going forward to lessen confusion. Also, the number of candidates and committees is smaller and made much more infrequently, so an ID with fewer characters than a UUID is easier for people to report.

Going forward, encoding data in relevant data models would help ensure IDs are stable and unique.



## **C. Expand data “pre-coding” and unify the data architecture**

### **Expand data pre-coding**

At points in the process of data intake, the FEC adds metadata not explicitly on the forms to each filing that adds useful context to the data. Generally, these codes are letters or numbers that represent a category: for example, an “S” is added to indicate records that pertain to senate committees, an “I” in the data to represent incumbency. This metadata is either “pre-coded”, that is, added by the software automatically upon submission by the filer, or manually hand-coded by individuals in Reports Processing after the submissions arrive at the FEC.

We recommend that automatic pre-coding be expanded substantially. This will simultaneously reduce the burden on FEC staff to code manually and the time it takes to do so, reduce the potential for human error in coding, and provide additional useful information to campaign finance data users immediately. For example, the e-filing software could leverage the existing API to automatically apply the candidate\_id that an independent expenditure was spent for or against. The codes could then be augmented later by Reports Processing staff, as necessary.

### **Unify the data architecture**

We further recommend that the data architecture, including the metadata codes, be standardized, unified, and published for use across the agency. By doing so, these systems can more easily build off of each other's resources, simplify data center architecture, reduce chances of failed data center transport processes, make updates and maintenance easier, and avoid costly work marrying different departments metadata. For example, the project of developing [beta.fec.gov's calendar](https://beta.fec.gov/calendar) took weeks longer than estimated because of the difficulties of unifying data across departments—costs that could have been avoided if the codes were aligned across the organization.

Referring to data within the same system makes it even easier to set up triggers. Triggers allow scripts to update data if a condition changes. This is especially important because filings can be updated, no matter how far back in time they are. For example, if a campaign submits an amendment that changes its spending in the 3rd quarter of 1980, the corresponding totals for that candidate and any other aggregates that are based on those figures will have to change. While amendments that are over five years old make up less than 0.2% of amendments, the FEC master database regenerates all its calculations records nightly to ensure proper handling of these edge cases. We recommend leveraging triggers, partitioning, and materialized views to allow for processes that will only recalculate the totals for that candidate and not every candidate.

Combined, pre-coding, triggers, and unifying the data architecture allows data to be processed continuously instead of batched nightly, which ultimately creates opportunities to have more real time data for the public.

#### **D. Use standard file formats**

The FEC has made it easy for data consumers to export and analyse the data it releases by publishing campaign finance data using standard data formats (.csv, .json, and .xml). In the e-filing system, two main file types used: .dcf and .fec. The .dcf files are used to keep local database records for FECFile. The .fec files are ascii delimited text files that filers submit. Neither of these formats are commonly used for data presentation outside of the FEC, and one (the .dcf file type) is used for copyrighted audio.

During our research many groups—specifically, commercial campaign finance software vendors, data journalists, individuals working directly on the FEC system, and filers who wanted direct access to their own data—all articulated a need for more common, open data formats. Much as publishing the submitted data in open, standard formats has made its use easier, the same can be accomplished for the working data. We therefore recommend that the FEC evolve their working files towards more common, open formats.

#### **The future of .dcf and .fec files**

While the file type for .dcf files is somewhat dependent on FECFile's future form, we specifically recommend that .fec files be transitioned to be the more common .csv files. If the next generation of FECFile is a desktop application, it should transition the .dcf files to a local database like [sqlite](#) to store the data. Using a local database will speed up data validation checks and improve overall performance and is a common data framework that developers can host, and use to build their own tools. If, on the other hand, FECFile is to be a web utility (as users desire), the data would be stored in a database remotely, either encrypted with the FEC or with a commercial software vendor. In that case, data exports should be provided for all stored data types as a common format, such as .csv files.

# How to move forward: software development workflow & technical road map

While this study has already discussed the need to employ user centered and agile processes in developing enhancements to the e-filing system and provided many specific recommendations regarding the types of design and infrastructure changes we believe will help both the filer and those that help the filer, while avoiding costs, this section lays out more details of the proposed software development workflow and a rough roadmap for next steps in upgrading the infrastructure.

## Evolving software development workflow

### Test driven development

Test driven development (TDD) is an agile approach to software development that requires writing unit tests, where a unit is the smallest testable software component, prior or in parallel to writing any functional code; the code is then tested as soon as it is written. Developers write tests and code in small, iterative cycles. TDD is a best practice in modern software because the benefits of testing outweigh the time it takes to write tests. A thoughtful test suite will make sure that key processes in software do not regress as the software evolves.

One of the main benefits of integrated testing is never having the same bug twice. When you make tests for situations that you know have caused problems in the past, you will automatically check to make sure the outputs of that situation are as expected. A wide array of unit tests should check that outputs of functions behave as expected. A targeted amount of integration tests should make sure that processes are functioning correctly together. As a starting point, the FEC could ask internal software testers (such as RAD) for common scenarios that they test when they look at software problems and recreate those tests as automated tests.

For the FEC, a hearty test suite could increase the confidence of vendors. Some vendors do not start using new FEC modules until they have completed their own in-house testing. This means that advancement in the modules, like the improvement to the FEC print module that enable automated processing of paper forms, are not enacted by filers until over a month later. If there is an open source test suite that proves the integrity of the e-filing modules, vendors can run the tests themselves and have confidence that the regressions they have seen before will not trouble their clients again. Open sourcing these tests could be an opportunity for vendors

and the wider community to contribute, since they share incentives to make sure the software components are functioning properly.

Adding a robust test suite makes it easier to make changes to the existing codebase or create a new codebase. Because software changes in a complex data environment, like FEC disclosure systems, can have wide-ranging impacts, tests can create more certainty that unintended consequences will not creep into the code. This can lead to faster development over time. For the Beta FEC project, we also check for things like n+1 queries that can be easy to miss and slow down queries later.

A critical part of testing is to make sure your tests run automatically. Finding bugs early leads to smoother development. One of the best ways to make sure fewer bugs make it to production is to run test as part of the deploy process and only deploy or publish those changes when tests pass.

You can read more granular information about testing in the [18F Automated Testing Playbook](#).

Moving e-filing data intake to the cloud should begin with implementing an incremental process for each process to be migrated. The first critical step is creating a test suite for core functionality in each unit. Then, the new process will be set up and it will need to pass the same tests as the process it is replacing. Then there needs to be a deeper quality inspection and the results of both processes should be compared. As bugs are uncovered, new tests should be added to the test suite so that problems are not repeated.

### **Automated workflow management**

Cloud Architecture depends on the ability of apps and databases to be reproduced on any server in the cloud system. This automated replacement of apps happens frequently and doesn't cause downtime because the system does not take the old instance of the app down until the new instance is created. That way, users get an uninterrupted experience. Because cloud systems depend on frequent, automated deploys, cloud deployments encourage consistent configuration management across projects. This can help maintain baseline security controls, requires less maintenance in terms of tracking down and patching individual systems, and encourages good practices like automated deploys contingent on tests passing.

Because the cloud model treats infrastructure like reproducible units that can be moved and recreated smoothly, there is an opportunity to automate the tedious work associated with the deployment process, minimize human error, and give people more time to more high level work.

In workshops and interviews FEC employees brought up the current workflow as a significant pain point. The EFO has made gains in managing its workflow, but the nightly processes and the data system as a whole would benefit from more automated workflow management.

Automating workflows allows for scalability and parallel processing. For example a trigger can be setup to sense high traffic, and add more app instances in response. Having multiple instances can also cut down processing times, and since you pay for what you use in a cloud pricing model, using several instances simultaneously can continue until it is no longer needed, and those instances can be shut down after the jobs are done.

## Technical Road Map

As part of this study, the 18F team was asked to make a roadmap to modernize and improve the FEC's electronic filing platform. The goal of the roadmap is to lay out a plan to build on the FEC's current strengths and identify where and how to progress.

The roadmap begins with an ordered list of tasks. While the tasks are roughly ordered, some could be done concurrently. The main technical elements of each task are listed in parentheses. Next we talk about ongoing processes that will help the FEC be successful, and we conclude with recommendations for the technical infrastructure (stack) that the FEC should use to complete recommended tasks and process changes.

### A. Add a test suite.

Test suites are a collection of test cases employed by software developers to ensure the program is functioning as expected. One of the main concerns from FEC staff in redeveloping e-filing is degrading the reliability of the system. Testing is one way to mitigate risk when making changes. Test suite considerations:

- Start by writing tests for current functionality.
- Start with writing tests for new features and bug fixes going forward.
- Add automated tests for things that typically get a manual quality acceptance test.
- Add tests for bugs that have occurred in the past, to make sure mistakes are not repeated.

### B. Improve upload speeds by creating a cloud-based upload API.

FEC should create a cloud based submission API that pushes files to Amazon static file storage. Employing cloud architecture for the submission process would allow the FEC to seamlessly add server resources as needed and maintain upload speeds even in times of high

demand. Using a static, cloud-based hosting service would also enable files to be uploaded without special processes or FEC-based server delays. One example of this process is in action is the Data Act API, an [open source project](#) that has an endpoint that uploads to Amazon S3.

This architecture could also be employed for validating filings, which would enable more parallel processing of validation: several instances could be run at once, and then extra instances could be spun down when they are not in use.

### **C. Improve data validation**

The following changes should be implemented as part of an iterative, human-centered design process in collaboration with users.

- Rewrite warning and error message text in plain language.
- Add guidance on how and where to fix the problem to warning and error message text.
- Add suggestions that help filers know how to enter the right data in the right format.
- Re-architect FECCheck to enable in-line data validation.
- Start validating the form fields currently validated by FECCheck in-line.
- Deliver warning and error messages to filers at the point of data entry.

### **D. While unifying the existing SQL databases, implement data improvements**

- Unify existing SQL databases to use postgres in cloud architecture. (The upcoming Cloud study will have more details about other cloud migration plans.)
- While making changes to the current database structure, partition data and triggers to only update changes to the information, rather than recreating the information.
- The FEC could alleviate this data issue by re-architecting the existing e-filing data schema. One short term solution is to keep the existing table and add columns for the different field types (which in turn would violate the database development best practice of [normal form](#)). A scalable solution, which will be needed in the long term, is to refactor the table into two tables.

### **E. Redesign of the print API to include machine readable data elements.**

- Add machine readable elements to FECPrint (such as QR codes) and the ability to read these to FECLoad to reduce errors to further improve processing times beyond recent improvements.

## **F. Make the filing utility platform agnostic.\***

*\*How this is achieved depends on the FEC's decision of whether to create a FECFile web app.*

- Begin with small, achievable goals.
- First, create a platform agnostic version of a small form, such as the form 2, before moving on to a larger, more commonly used form like the form 3.
- Design and test each element in collaboration with filers, RAD, and the EFO.
- Move away from the .dcf file format. This information is better stored in a database.
- Develop a web app version of FECFile. With open sourced code and API driven infrastructure, a web app will make it easy for vendors to integrate their own database and hosting solutions.

For FECFile users, the storage depends on whether the new app is a web or desktop based app. A web app could be integrated with the postgres database and the data could be encrypted for storage and unencrypted upon submission. This would be the fastest method and allow campaign members to collaborate on the same filing.

If FECFile were to remain a desktop application, it could be built with a local database (like [SQLite](#)) incorporated. This would be seamless to the users: they would not need to directly interface with the database, rather, it would function like an app on your phone. Because SQLite is a standard open database format, this would allow the software to perform faster and have the added benefit of making the data easier for developers to integrate with other tools.

Redesigning FECFile presents an opportunity to re-evaluate the .fec file format and potentially move to a more standard format such as .csv or .json, though this change would require changing the upload API as well.

The main issue with a format change away from .fec files would be recognizing the needs of departments other than e-filing. Special attention needs to be paid to make sure that Press and Public Records, and RAD's tools are still operational. Public outreach and public development of these tools will help others adapt to the changes.

## **G. Improve speeds by migrating static items to a static file hosting service such as [Amazon Web Services S3/Glacier](#).**

Cloud based hosting services provide convenient storage and delivery of static content such as images and pdfs and data files. In addition to increasing speed, moving to a managed cloud system typically reduces costs. The beta FEC architecture currently uses this architecture, and it has been shown to handle large files well.

## H. Review and adjust the plan as needed to accommodate user needs.

### Ongoing tasks and processes:

#### Create automated tests

- Testing decreases risk and can increase trust in a product. *(Here, testing is listed as the first item, but testing should continue to be part of each step of the process and integrated into the deploying and publishing code.)*

#### Document code

- Better documented code is easier to maintain and adding functionality or improving the codebase (refactoring) becomes faster and easier. Additional details:
  - Include more in-line explanations of functions.
  - The FEC would also benefit from better documentation of the data as it originates from a form, goes through processing, and is made available to the public.
  - Make sure outputs are publicly documented.

#### Open source the code base

- As components are developed or redeveloped publish the code. Consider beginning with the test suite.
- Grow FEC's GitHub organization and publish more repositories. GitHub is also a good platform for engaging the developer community that uses FEC data or interfaces with FEC systems.

#### Move to the cloud

- As the new version is built out, make [12 factor applications](#) on a scalable cloud platform. This will greatly increase the capacity and resilience of e-filing systems.
  - The FEC could easily use cloud.gov to unite the data systems, because the cloud.gov platform is already being used for the FEC's new website. Additionally, FEC staff are gaining cloud.gov skills so they will already know how to use this platform. As of January 2017, Cloud.gov is on path to receive government acquisition approvals which will help systems incorporating it to achieve many of the necessary compliance steps. Since cloud.gov is open source, and its



architecture is similar to other vendors, FEC could also have the flexibility of adapting to a different platform as a service (PAAS) down the line if desirable.

### **Use language agnostic APIs to enable components to interface with each other**

- The EFO is currently making progress towards language agnostic APIs. This progress should continue. Additionally, making the APIs restful would make it easier for internal and external developers.

### **Implement changes in an agile, human centered manner**

- It is important to make changes to the plan as user need are discovered and understood, and when unforeseen problems arise.
  - Several members of the FEC have Scaled Agile Framework ([SAFe](#)) training this is a technique that will set the project up for success. Continuing progress by continuing short sprints, with quarterly planning and regular reporting to stakeholders.

### **Continue to support the commitment and dedication that we have seen from across the FEC**

- During the study, we encountered a multitude of fantastic public servants who are dedicated to transparency and ensuring that anyone can run for office.

## **Recommended tech stack:**

### **Use existing infrastructure, automation, design, and data improvements, to build upon the strengths of the EFO and Beta FEC**

- The best way to achieve this is to create a FEC web product that leverages the technologies and code of the current Beta FEC system.
- For example, using the FEC style library would create a consistent user experience and includes features that display and explain the efilg content.
- If the FEC decides not to architect fully integrated web apps, FEC could still leverage the Beta FEC web app libraries using a service like [Electron](#).

### **Java script should be the driving technology for the new FECFile**

- Shifting some backend processes to python may be helpful in unifying the code base, but it makes sense to continue java applications as needed.

### **Data infrastructure should move to postgres for sql calculations**

- The FEC may achieve some searching and scaling gains by putting the itemized data into [Elasticsearch](#). This should be prototyped so that FEC can compare performance between Elasticsearch and the current sql offerings.

### **Moving to the cloud is integral to the FEC's modernization strategy**

- Expanding the existing cloud.gov architecture would be the easiest way to implement this change.

# Research methods

## Research approach

In July 2016, 18F and FEC staff participated in a workshop dedicated to understanding the principal goals of the e-filing study and to surface knowledge about the shape of the study. The three main goals of the workshop were to: outline the hopes and fears for the study, get an introductory understanding of the processes that shape e-filing, and, determine the key groups to talk to. The full workshop read-out is available [here](#), but some takeaways follow. Some of the major hopes of the study were to increase the overall accuracy of the filings and to understand how we might help address the burden of dealing with paper forms from the senate.

Participants expressed anxiety that the changes wouldn't be allowed, or that that it would introduce security threats or take too long. Participants also outlined some of the most important users to talk to, including FEC's RAD analysts, vendors, and of course the people who file reports to the FEC (henceforth referred to as filers). These hopes, fears, processes and research subjects gave us a strong introduction to the project to build our research plan upon.

We conducted our research in three phases. In the first phase we conducted background research, interviewed and observed stakeholders to understand the current e-filing process and areas for improvement. In phase 2 we developed design concepts that we believed had the potential to make the process better, and then we had users evaluate these concepts via usability testing. In phase 3 we conducted a survey of all e-filing users to learn more about how they would feel about specific changes the FEC might want to make.

### Phase 1: Contextual interviews

To understand how to improve the FEC's electronic filing system for the different people who interact with it, we first had to talk with those people.

The following **research questions** guided our work:

- How do **filers** currently use e-filing software? What aspects of their process are particularly cumbersome or error prone? What barriers drive them crazy? What short-cuts, workarounds, and help do they employ? What's working well?
- **E-filing analysts** (internal to the FEC): how does the form and format of e-filing negatively and positively affect the Electronic Filing Office and RAD analysts' workflow?
- How do different kinds of **vendors** interact with their filers and the FEC? How might software vendors be impacted by changes to the FEC's filing process?

- We also interviewed **FEC technologists and contractors** on the current system to ask them; What is going well? What are the pain points?

We worked with our agency partners to generate an email list of potential participants from the FEC's e-filing database. We identified additional participants via an online screener on the beta FEC website. In phase 1, we conducted semi-structured interviews with filers, commercial software vendors, analysts in the FEC's Reports Analysis Division, and other FEC stakeholders. The focus of the interviews was on participants current processes including pain points and information gaps. Interviews with filers and RAD analysts included a portion where we could observe participants on their own computer using their own systems to complete some actual filing work. In most cases, these interviews were conducted remotely using online screen sharing and audio conferencing software, which were screen and audio recorded. In some cases we conducted the interviews in-person and documented these interviews in field notes.

Following each session, we came together to talk through what we were seeing and synthesize our knowledge in the form of additional notes, whiteboarding sessions, and/or sketches. Our research team systematically reviewed field notes, videos, and other work artifacts gleaned during these sessions and interviews to identify themes. We then independently "tagged" quotes from users that are evidence to these themes. We then articulated our riskiest assumptions and developed hypotheses that we could test through prototype testing in phase 2.

## Phase 2: Prototyping & testing

In phase 2, we sought to validate the hypotheses developed in phase 1 through the creation and testing of prototypes. A prototype is basically a draft site built quickly and roughly with throw-away code that can be used to observe user interaction with examples of potential sites. It is built with a priority on quick and simple development for rapid learning and iteration, so it is not built in a way that it can serve as the foundation for future iterations of a production site. The results of prototype testing informed our recommendations and study findings.

A list screencasts demonstrating some of the features we tested in our work-in-progress prototypes are below:

- Manual [data entry flow](#)
  - Detail: [in-line field level validation on the date field](#)
  - Detail: [in-context glossary](#)
- Importing data [from a spreadsheet flow](#)
- Merging [duplicate name records flow](#)

Additionally, the code used for prototyping is available on [GitHub](#).

## Phase 3: Survey

Phase 3 comprised the distribution and analysis of a survey to help us test one of our hypotheses that proved difficult to explore via prototyping. Specifically, we needed to understand the filing community's thoughts around a web-based version of FECFile.

## **Methods**

We conducted a survey to find out how the filing community felt about the possibility of a web-based version of FECFile. Participants were invited to the survey [via an email](#) sent from the FEC's Electronic Filing Office (EFO) that linked to a web-based [survey](#) form.

Our partners at the FEC helped us compile the survey distribution list by providing us with the committee IDs for all committees who filed electronically in Q3 2016 and all of the email addresses associated with these committees. In many cases, the same email address was listed several times for the same committee or for several different committees. After we removed duplicate email addresses, our distribution list comprised 8,906 unique email addresses. We received 533 responses.

Part of our intention was to make sure participants were comfortable sharing their opinions, even if they might be perceived as unpopular by others in the filing community, so we kept the survey anonymous. We were able to distribute the survey to the entire population of committees who had recently filed electronically, and we minimized sampling error by reaching more than our target number of 361 responses to be able to make population level inferences with a 95% Confidence Interval (CI) (margin of error, 5%). We avoided coverage error by distributing the survey to all committees who filed electronically in the most recent quarter.

## **Analysis**

One researcher combed through qualitative responses tagging the data for prevalent themes. Some of the themes were mentioned by a large number of people, and some of them were mentioned less frequently. Additional analysis involved looking for themes within subsections of the population. The qualitative findings are organized by the questions we asked.

## **Limitations**

No survey is without limitations. In a few cases, we asked participants to self report estimates, so these answers could be subject to recall bias. The survey was voluntary, so there could be some self-selection bias. Specifically, the email invitation mentioned that the survey was exploring possible changes to FECFile, so we may have received a disproportionate number of responses from opinionated FECFile users. Similarly, there could be some non-response bias because filers who do not use FECfile might have been less likely to participate.

## **Overall Study Limitations**

One of our recommendations throughout this report is that the FEC adopt an agile approach to technology transformation. Working in an agile way means planning, conducting, and evaluating work in short sprints and revising plans as needed. Adopting an agile mindset means embracing the fact that plans must always change and that laying out lengthy detailed plans before the work begins can mean a lot of wasted effort. In keeping with the spirit of an agile approach, our aim here is to provide the FEC with what it needs to **begin the journey** of modernizing the current e-filing system rather than provide an exhaustive list of all that this journey will entail. We see this as a strength rather than a limitation, but recognize that this is not typical for many government reports that may be more familiar to some readers.

In keeping with this agile mindset, we worked with our FEC partners to prioritize who to talk to within the study timeframe. While we did not interview members of some stakeholder groups in this initial discovery project, we strove to make it clear throughout the report the user groups that must be involved in the iterative design, development, and evaluation of any changes that the FEC decides to make.