1095B Form Technical Discussion Notes

**February 2, 2022**

**Present**: Joshua Faulkner, Jordan White, Ben Morris, Nadya Primak, Jared Cooke, George Tidwell, Tracey Mulrooney, Ashley Castillo, Tarsha Tremble

**POCs**

* George Tidwell: Analytics & Informatics member services
* Joshua Faulkner: Technical lead for \_\_?

**DISCUSSION OVERVIEW**

**Print Vendor -> DB process**

Print vendor (GPO) creates PDF copies of them, Enrollment System picks them up and brings them in and stores them in binary format – that way there aren’t 100 million PDF files sitting out there because they are deleted as soon as they are imported into the database

Could find some way to just take them instead of deleting them (we could store them ourselves) but then you have to store a LOT of PDF’s (tons of space would be needed)

Getting the PDFs from the repository (print vendor has 10 days to submit) just read the database and convert it from binary to PDF

**PDF Data / fetching by Veteran ID**

When they authenticate you can match the veteran by ICN

The PDF’s that generate don’t put full SSN (just last 4)

SFTP protocol would send us the data and we could create the PDFs from there (delimited text file) – first name, last name, ICN, DOB, address, coverage period (covered under ACA mandate)

He can put the PDFs data into S3 right now (super simple) plenty of scheduled scripts

**Data in Enrollment System**

A whole separate project would have to be completed in order to store in the DB if veterans are opting out of receiving form in the mail (for now we would have to still mail the form regardless)

The opt out would be handled by and update to the VA profile

The current data in the Enrollment System would not be representative of what was mailed to the IRS **BUT** you could use the data that is sent over to the print vendor)

**IRS Rules / Process**

IRS has their own SOAP service (similar to e-filing taxes) every year bulk reporting goes to IRS – packaged up and they send acknowledgment that they got it

You have to make an attempt at a paper mailing for minimum essential coverage

IRS opens environment in middle of January, to get IRS’s data it’s going to be delayed even more than waiting on print vendor to produce a PDF – delay of a month and a half or more

They log the XML of what’s sent to the IRS but using that instead of the letter data it will still be delayed because of how long it takes for IRS to open their environment– data is identical of print vendor and data sent to IRS just in different formats

Have to get the ICN and query based on that unique veteran ID

**Populating forms**

Programmatically you can do it in Java ( populating the forms ) -getting a fillable version of 1095B – generating them on the fly versus storing them

**Snapshotting PDFS (pulling data from date of IRS submission)**

To pull the snapshotted PDF’s (wouldn’t be available till later, would require new feature of E&E have service updated with new operation for document access)

Scanning through audits would take too long (they quit using it because its too slow)

Add another field in that file for the vendor to ignore the record, add field update (for print vendor to generate a form and skip the printing)

**Testing**

This all exists in development + pre prod, it still exists for testing (manually send them the file and generate proofs/samples to get approved by business)

Dev + SQA is fake data

Few S3 buckets for different environments – flat file if there is a version of that which is easy to share in advance to start to get a handle on (send him an email he will give us one from dev)

Format of the file is unchanged ^

ICN padded with 6 zeros, you can get ICN with that (for querying based on veteran ID)

the data fields ( whatever is needed to populate the PDF)

**Corrections**

Checkbox on 1095B that says “corrected” – we would get a new data file if veterans make a correction

Seeing an ICN another time means you’re overwriting an existing version (it would be another copy) – multiples instead of overwrite

User corrects coverage period, next batch script would run and send to vendor, then goes to IRS immediately (batch file runs all year long) – have a process looking for new files that would be running all the time

**Optimal Solution (future?)**

Document service has other PDF’s but not 1095B – repository for each application which you can share – makes it easier to hook up to API and pull the documents (would be a simple integration but 1095B’s aren’t there)

There are no existing API’s to grab the PDF (E&E Service) Community Care API on VA.gov

**MEETING TRANSCRIPT**

Josh:

* Can store copies of them locally, there are many millions and they are PDFs (not lightweight)
* Or you can connect to our repo to get them
* Currently they are created when we send the request to mail out forms.
* We pick them up from the vendor and store them in DB in binary format and then delete them as soon as they are imported into the DB.
* Going forward we need to find some way to let us take them instead of deleting them, we store them ourselves. Can't get existing ones cause they've been deleted. Or connect to repo, and get them on demand when they are stored.

Tracey:

* Is there an option to connect directly to enrollment center data so we can create the PDFs ourselves?

Josh:

* Yes, the problem is, you will be building a form based on current data on their record, it wouldn't be what they are going to mail to the IRS.

Tracey:

* Can we use the same data as the print vendor?

Josh:

* That is possible. They are just text files with variables in there that we put in the fields. That would also be point forward. We don't retain those forever.

Tracey:

* We want Veterans to be able to opt in going forward..?

Josh:

* It wont work that way, that will be handled on VA profile (opt in email, text, etc.), won't be soon but it is already planned
* Communications work with a vendor by a scheduled job that creates these files and transfers them to the server, doing that same thing for you is super simple, we can allow you to build your own PDFs.

Ben:

* Is it a flat file?

Josh:

* Yes, flat file, with some variables.
  + DOB, ICN, coverage period, name, address, etc.
  + Once they authenticate we can match them by ICN
  + Coverage period is important (what month theory covered under ACA mandate) That is all you need for form.

Nadya:

* We are talking about getting data that is sent to the print vendor. In a blank format flat file?

Josh:

* yes.

Ben:

* What gets sent to the IRS?

Josh:

* The IRS has their own service, SOAP service, similar to how you file taxes. Every year bulk reporting goes to the IRS, we package them up, send them through soap service.
* They don't care about the forms, just care about the coverage period. It is on us to send out the forms.

Ben:

* For future state, would it make sense to key in on data sent to the IRS, if there was an opt out process, the file sent to the print vendor would be a subset of the IRS?

Josh:

* Can't do that, the way the law is written, you must make an attempt to have a paper mail in.

Tarsha:

* The law states, if we give them the options, and we record their response, then we can do it (opt-in). But we have to document their opt-in or opt-out status. In case of an audit.

Josh:

* We tried this in 2015

Tarsha:

* Yes but the law has changed a bit.

Josh:

* Question for another time: How can we capture that and document it in the enrollment system?

Josh:

* Include them in the file, add a field for the vendor to ignore it, because the IRS uses the same file, we must report to the IRS exactly what we send to the veteran.

Ben:

* Another similar option: get the data in the right state, to do a query live out of the enrollment system, if it's possible to pull the data as reported to the IRS. Not sure what data structure is like?

Josh:

* The IRS does not open the environment until mid January, if you use IRS data it will be delayed. We generate all of them in early December, In January we pick up people enrolled in December, and send those out. Month and a half or more delay (takes another month to send them out)
* If someone gets form in December and wants another copy, they would have to wait another month and a half until it has been reported to the IRS.

Ben:

* Data structure in enrollment system allows for querying for a given person to see the state of the data. If there was an error, is it possible to grab the data as of when it was sent to the IRS?

Josh:

* We log the tables we send to the IRS. It isn't sent at the same time the letters are sent. If you use that rather than letter data, it will be delayed by 2 months.

Ben:

* Print data pull & IRS data pull come from the same place?

Josh:

* Exactly the same, print happens first because the deadline is the end of January. We start in December with everyone who has been covered for the year. Send out the rest in January to make sure everyone gets their letters printed and mailed.

Nadya:

* Confirming the fastest option: getting text delimited data and populating PDF ourselves. Other options have a delay?

Josh:

* Yes, or get PDFs from our repo, but then you're waiting on the vendor to create them. They have 10 days to do so once we send them the print order.

Nadya:

* Do we have to create a way to query the PDF to get a specific veteran?

Josh:

* Get their ICM, then query based on that.
* Fastest way: no physical printing and mailing, just generating PDFs only.

Nadya:

* Is there an existing system the VA uses to populate PDFS?

Josh:

* We use PDF\_fill to fill manually. Programmatically you can do it in Java. Just need to get a fillable version of 1095B, fill it in either .Net, or Java

Nadya:

* Batch job? We can create our own?

Josh:

* Yes, but now you're storing them unless you can figure out how to generate them on the fly when they are requested. Not sure how slow or fast that is, we are talking about 10 million PDFs, (very large)

Nadya:

* And is that why you store it in binary?

Josh:

* We store in oracle and use advanced compression. Or else they would be 10x bigger.
* If you could take the file, generate it when someone requests it, wouldn't be too bad because there aren't too many variables.

Ben:

* Curious, one path is the same batch job that sends the flat file to GPO to the printer, have that just add a second destination (for us, an S3 location) that seems relatively straight forward, is that easy to get that process set up?

Josh:

* S3 you could do it now, just need an access key, lots of scripts that write to S3 daily.

Ben:

* Another option, pulling binary files from the DB, are there other read APIs in vets-api that access the enrollment system?

Josh:

* Yes, but they don't connect directly to the DB. E&E service (enrollment & eligibility service), community care API that VA.gov uses. Need to know exactly what is sent to the IRS.

Nadya:

* Does that mean a vet could call and update their data?

Josh:

* For the correction process: Say a veteran reported as uncovered but that isn't right. You can go in and fill in what it should be, it will go to the print vendor and get reprinted and re-sent.

Ben:

* E&E service API does not include an interface to access stored snapshot PDFs?

Josh:

* No

Ben:

* Pulling snapshotted PDFs wouldn't be available until later (after batch job) and would require a new feature of the E&E service?

Josh:

* We would have to add a new operation to the service. That won't happen quickly because we have too much in the backlog.

Ben:

* To confirm, in E&E service we can get enrollment data but only the most current? can't get it as it was at print date?

Josh:

* We have a history operation but it is very slow. It has to scan through all the audits and piece it together over time. Too slow.

Ben:

* If we did get data from a batch job in January, if a veteran is mailed their form and sees that a correction needs to be made, then they go through whatever means to update the data. Then, how does that flow through to the IRS? Batch again?
* Flow: User corrects form -> form goes to IRS -> form goes to vendor -> form goes to us at same time.

Ben:

* We would potentially have to create a new batch file?

Josh:

* Runs all year long, bulk in Dec and Jan, rest of year is light, just re-mails or corrections.

Ben:

* Do we need a process running all the time to look for new files?

Josh:

* Yes, all the time

Ben:

* Daily to the vendor?

Josh:

* Twice a week to the vendor. Also get files that are corrected from last year or before. Someone calls and says two years ago you reported wrong, now we have to correct a form from 2019 for example.

Ben:

* Next year, if opt-in (document consent) gets worked out, this could be done if there becomes a discrepancy between the whole set of data that goes to the printer?

Josh:

* I would add another field in that file for the vendor to ignore the record.

Ben:

* Is there test or sample data for this?

Josh:

* Yes, this all exists in pre-prod, we test it and manually send the file in lower environments, they generate proofs, then it runs for real. For this pattern I would add an S3 copy script so we can get the file.

Ben:

* Is there test or sample data we can get now?

Josh:

* Yes, email me.

Ben:

* If it is an update we can..?

Josh:

* No overwrites, you would see multiple forms for each correction.

Ben:

* Batch file has all the data fields?

Josh:

* Yes, it has everything needed to populate the PDF.

Ben:

* If this follows patterns of other cases of making data available to the vets-api is there any other easy references to look at?

Josh:

* All stored in DAS, these 1095Bs are not included in there. It is for DD214’s and things like that, these are only stored locally.

Ben:

* If there are things that are a good reference pattern?

Josh:

* DAS has a document service, basically a repo for each application to allow them to be shared with other applications.