

Beratung

Challenges of Audio eDiscovery

21. April 2016, von Irène Wilson



Introduction

Have you ever heard of audio eDiscovery? Last year Swiss FTS were given an opportunity to enter this fascinating field. While projects in this field have traditional eDiscovery considerations, such as the quantity of data and privacy restrictions, we encountered some interesting new challenges. For example, how do you search audio data? Can you access obscure or proprietary file types? If you start delving into the realm of audio eDiscovery you will be blown away by the breadth and depth of information and articles available. Let me introduce you to this new dimension of fluid and evolving possibilities...

What audio data?

While audio files are usually overlooked in traditional eDiscovery as “Mr. Bad Guy’s favourite music playlist”, it happens to be of the utmost importance in some specific areas. For example, companies offering trading services are bound by law to record traders’ calls. A lot of hotlines and support centers record calls as well, and voicemail and private records can also be relevant in some cases.

While audio data exists in the eDiscovery world, it often hides in the shadow of traditional email

review. The field is not entirely alien though, as it shares a tricky characteristic with email review: the variety of formats. This goes much further than just the file format, and involves finer nuances such as codecs and bit rates. You will need to know the recording system your clients use, as well as the software version and settings in order to perform a full analysis.

Not only can the audio itself be formatted in many different ways, but the metadata related to it can be stored in various locations, such as within the file name, on a distinct spreadsheet or in a database. This will need to be parsed or extracted, and mapped to the relevant audio tracks.

The identification of the protagonists within a conversation is much more complex than with emails. Do you actually know whose line you are listening to? While the custodian is sometimes clearly defined, it's not uncommon for the audio feed to be tied to a specific source rather than a person, such as a specific desk or phone. In these situations, custodian mapping is not straight forward and will require some additional work. This cumbersome task can be supported by speaker identification systems, which allow you to identify a speaker's speech pattern with just 5 minutes of their recorded speech, and expand the recognition to your whole data set.

To ensure the completeness of the captured data, you will also have to consider some additional points. Are outgoing and incoming calls recorded the same way? How are redirected calls taken into account by the recording system, and how does this impact audio metadata? This last aspect could lead to missing data if overlooked.

Managing your audio eDiscovery projects

This already sounds like a thrilling challenge, doesn't it? I know the techies are already lost in their thoughts, assessing the different pitfalls possible when preparing this type of data for their clients to review. That's great, but remember not to lose track of some down-to-earth realities. Let's be organized and provide our client with some project management plans, and consider issues such as the expected project time line. Several factors need to be clarified before this question can be answered, even vaguely.

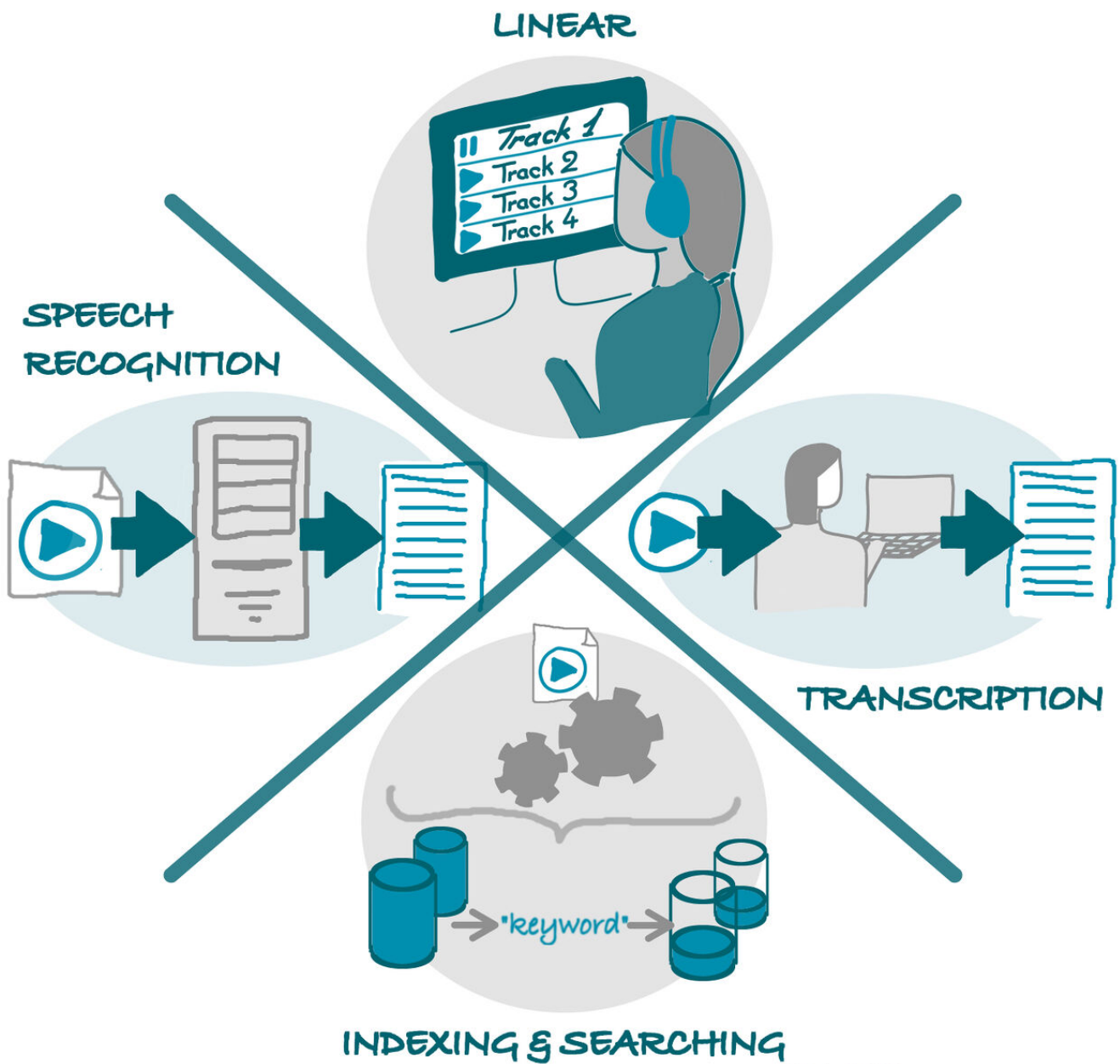
First the scope, in terms of date range and custodians/lines, need to be defined. Some additional considerations to reduce the data set include limiting the data investigated to working hours/days (some systems record lines on a 24/7 basis), filtering out the parts of the records that don't contain conversations, and filtering by custodians. As with email review, the amount of data impacts your projected time line. Information such as bit rates and total data volume are key factors in predicting the number of hours of audio data you have to account for.

Another factor that will impact your time line is the way the data is stored. Chances are that the files are located on backup tapes, in encrypted and compressed formats. Restoring and converting this data into a usable format are time-consuming tasks you shouldn't overlook. An additional point of interest is the audio system limitations: Jeff Schlueter mentioned in one of his articles that one major system in common use restricts exports to only 50 files at a time (i). While this aspect is out of your control, it

certainly needs to be taken into account if you want to give your client accurate time estimates.

Review strategies

Now that you have a good understanding of the number of hours of audio data targeted by your project as well as the time needed for the client to provide it to you, it's time to think about the different review strategies. In this regard, audio eDiscovery offers more variety and originality than traditional eDiscovery.



(c) Swiss FTS AG

Linear review

The first review strategy is just a simple linear review. You give the reviewers access to the audio files. They will then listen to all of them and classify the files according to their relevance. Some sources (ii) observed that reviewing 1 hour of audio content takes on average 4 hours. This information can help you assess how long the review will take, just based on the amount of audio data in scope.

This approach has the advantage of requiring little technology which limits the indirect costs related to the review. However, it might have a serious impact in terms of deadlines and review costs in big cases.

Transcriptions

Transcribing is an easily overlooked approach which involves people manually creating a text transcript of each audio file. While this takes time and involves extra costs, it has the massive advantage of bringing audio back into traditional eDiscovery. The transcripts can be handled with your usual eDiscovery methods to process, prepare, search and review. This strategy brings you back to tools you know and master, and allows you to apply a single consistent and uniform process to all your data. Furthermore, it also opens the door to analytics and computer-assisted review. It should be highlighted though, that transcripts are not exempt from mistakes, and don't capture the intent or intonation of speakers. The major drawbacks to this approach are its impacts in terms of time, costs and accuracy.

Indexing and searching

Some specific audio eDiscovery tools provide solutions for indexing and searching the data. Indexing can be applied up to 340 times faster than real time (iii) and then the quantity of data can be reduced through keyword filtering. Sounds familiar? While the basic concepts are similar to traditional searching, there are some peculiarities when compared to text data. The smallest unit in text indexing is a word or a character, whereas for audio it is a phoneme. This is the smallest part of speech, individual sounds contained in our language. As this definition highlights, phonemes are language specific. The process requires identifying the language of the document before applying the appropriate indexing model. This is highly relevant in a country like Switzerland, where 4 (or more!) languages can be expected.

While this approach is interesting in terms of targeting relevant data through keywords, and therefore limiting the number of hours spent to review, the IT investment it requires is not insignificant and should not be overlooked. The hardware, and especially the software, needed to apply this strategy come at a price. Different pricing models exist which usually increase with the number of languages required, although this approach is often beneficial on larger projects where linear review is no longer an effective option. For example, a project involving 5,000 hours of audio data would take approximately 5 full months for a team of 25 reviewers with a linear review strategy (iv). With an indexing-searching strategy, this data could be indexed and ready to be searched in less than a day. Even if the search terms aren't very specific and retain 20% of the data, that will still decrease the review time from 5 months to 1.

Speech recognition

The final review strategy is still in the realm of science fiction. Speech recognition, or speech-to-text, effectively automates the transcription process, using a computer to transcribe audio into text. While this is great in theory, the current state-of-the-art is not accurate enough to make it a usable solution for eDiscovery purposes. Anyone who has tried voice commands on their phone, or dictating a message in Whatsapp®, knows the error rate is still pretty high. Recordings also add extra complexity with background noise and individual accents. The problem of price with such a technology should not be forgotten either.

Searching audio

If you decide to go for a strategy which allows you to decrease the amount of data to be reviewed via searching, there are several aspects that you need to be aware of.

The more unique your search strings are, the better your results. This is not a trivial statement, as we are speaking about sounds and not words. The words “police” and “pelisse” portray very different concepts, which happen to be pronounced exactly the same way. In general, the longer your search terms are, the better. While “police” and “pelisse” are aurally ambiguous, the expressions “call the police” and “wrapped in her pelisse” are clearly distinct.



Dates, signs, numbers and acronyms are especially tricky and will require you to search for multiple variants. 153 can be pronounced “one five three”, “one fifty three” or “one hundred fifty three”.

Your search terms should also take into account that speech is less formal than writing, often including slang terms such as “quid” and “bucks” instead of the formal “pounds” and “dollars”. Regional differences are also relevant as “faucets” and “taps”, or “trunk” and “boot”, are very different sounds with the same meaning. Furthermore, the impact of different dialects should not be overlooked. Swiss German for example varies from one canton to another, complicating the issue of accounting for all of the different ways of expressing the same concept or idea. Finally, spoken language is extremely versatile. Not only does it evolve fast but significant differences can also be noticed throughout different generations.

Conclusion

As promised at the beginning of this article, audio discovery is indeed fascinating. Even though it takes place in the same context as email review, it requires a different approach and tool set, and forces you to look at your project from a different point of view. For me the most striking truth about audio eDiscovery is the importance of being prepared before entering this field. At the very least, research the subject, contact software providers and test the different tools available. If you dismiss audio data as just another file type, instead of a different field within eDiscovery, you are bound to make promises to your client that you can't keep.

Footnotes

(i) Jeff Schlueter, 2013, Dodd-Frank and Audio Discovery Requirements,

<http://www.audiodiscovery.com/dodd-frank-and-audio-discovery-requirements/>

(ii) Jeff Schlueter, unknown, AUDIO : Searching is Different than Documents, EDRM Magazine,

http://www.nexidia.com/files/resource_files/EDRM%20Audio%20Searching%20is%20Different_495.pdf

(iii) Jeff Schlueter, 2008, Did I Really Hear That?, EDRM Magazine, <http://www.nexidia.com/about-nexidia/news/did-i-really-hear-that/>

(iv) 5,000h, taking each 4h to be reviewed ends up 20,000h of review. Assuming the team works 8h a day, 20 days per months, 25 reviewers would take 5 months to complete the job.

Irène Wilson

Irene Wilson ist auf digitale Forensik und eDiscovery spezialisiert und hat im Laufe ihrer langjährigen Erfahrung für Kunden aus einer Vielzahl von Branchen in ganz Europa gearbeitet. Zu ihren Qualifikationen gehören die renommierten Master-Titel für Nuix Workstation und Nuix Discover.

Swiss FTS AG | www.swiss-fts.com | +41 43 266 78 50 | info@swiss-fts.com

<https://swiss-fts.com/blog/challenges-of-audio-ediscovery>