vCon: an Open Standard for Conversation Data

Summary

In December 2021 Thomas Howe, CTO of STROLID, introduced vCon.

We have iCal that enables anyone to store and exchange calendaring and scheduling information such as events, to-dos, journal entries, and free/busy information. And vCards so anyone can store and exchange electronic business cards, name and address information, phone numbers, e-mail addresses, URLs (Universal Resource Locator), logos, photographs, and audio clips.

Yet despite all the talk about conversations in the programmable communications industry, conversation data remains trapped in silos. vCon (virtual Conversation, like vCard), is a new open standard for sharing conversation data: transcript, video, audio, participants, metadata like timestamps and location, tamper protections, certifications, etc. vCon has the potential to create an ecosystem of innovators focused on creating new conversation intelligence tools, in addition to established platform providers.

This whitepaper provides an introduction to vCon, and an invitation to join the creation of an open standard whose aim is to greatly improve the programmable communications industry. Check out the vCon github repository and get involved, it's open to everyone.

Sharing conversation data is a significant problem faced by programmable communications developers today. vCon makes working with conversations easier, hence the addressable market of developers expands ten thousand fold across web and enterprise developers. This happened one decade ago when telecoms became easy to use with simple web-centric APIs. This will happen again for conversations thanks to vCon.

An innovation ecosystem of specialists in conversation intelligence will flourish solving specific business pain points across operations, compliance, privacy, security, ethics, etc. Being able to access customer data often trapped within communication platforms. Think of vCon as 'robot food', enabling conversation data to be presented in a common format and more easily cleaned for training of machine learning. ASR and conversation AI solutions do not meet the needs of some businesses with respect to accuracy, vCon will help our industry close the gap with respect to the hype.

Through the creation of a common unit of exchange and openness, industries have been transformed. Global trade in the 50 years from 1967 to 2017 went from 22% to roughly 60% of global GDP. SMS in 5 years grew 3000%, and created multi-tens of billion dollar industries across CPaaS and A2P SMS. Open banking in the UK became ubiquitous in 4 years.

vCon will do the same. Conversations currently trapped in silos of communication platforms or simply stored in the company's data lake will become common units, where innovators will create new value, solve business problems, and deliver valuable insights because their business lives/dies on delivering that. vCon is the next leap forward in the programmable communications industry.

Table of Contents

Introduction to vCon by Thomas Howe

We have standards for all kinds of business data elements. vCards for contacts; iCal for appointments; XLS (Excel Spreadsheet) for worksheets; etc. Yet, none for conversations. We can forward a single static document, but we still can't forward a conversation. We have standard tools for viewing and editing documents, but even if we could forward a conversation, how would the recipient know the context or how to open it?

We could send a transcript, but this is a derivative of the conversation, and it doesn't provide enough detail about the people, organizations or context related thereto. Surely it doesn't identify the participants authoritatively. We can forward an audio file, but how would we deal with that as database developers? We can't search an MP3 file with a SQL (Structured Query Language) statement. Given their obvious importance and ubiquity, conversations and the businesses that rely on them, deserve better.

In reality, conversations are an aggregate of many types of and instances of data, that data being unified in the fact that it's an accurate record of what one or more humans actually communicated with each other (through all the ways they can, from voice to video to text) in the real world, at a particular place and time. We can record every aspect of a conversation, yet we have no standard to join those pieces together in a single data structure to represent a conversation. Yet it is the conversation that ultimately matters the most to business personnel, engineers and scientists. Today our capture of the human conversation is all words, with little context or punctuation; just long run-on sentences.

The time has come to deal with conversations as first class citizens; a natural aggregation of data. CX (Customer eXperience) efforts, AI (Artificial Intelligence) as applied to conversations and even enterprise pricing structures all naturally use a conversation as the basic unit of measurement. We have a once in a decade opportunity to capture the real and actual voice of the customer, arguably the root of all real customer data. If data is important, then the quality of the data is arguably more important. Literally no other segment of the business world has the opportunity to deliver the most real, most accurate and most reliable customer data then those of us who design with it. vCon can become the common language we use to represent them.

In retrospect, a major impediment to the use of conversations as data, to feed software and machine learning, is that we did not get the resolution right. As far as data goes, our field worries about packets, SMS messages (Short Messaging Service), SIP invites (Session Initiation Protocol), CDRs (Call Detail Records), emails, WebRTC, web chat. Yet, at a business level, the proper unit of measurement for real time communications, the thing that is the least common denominator of a human expression, is the conversation and context between the parties, human or artificial. It was easier in the good old days of POTS (Plain Old Telephone Service). Conversations started with a ring, and ended with the handsets hitting the princess phones. With smartphones, and their ability to generate and consume every kind of real time

data, we need an industrial strength answer to what a conversation really is, that is the real value of vCon

Thus begins our need to define the data structure of a conversation, to define what conversations are, but more so to define what a particular conversation is, with as much precision as possible, including the identities of those involved.

What is a vCon?

A single real-time conversation
A vCon is a file format that contains and describes a single conversation between two or more people, or between a business and a person (e.g. customer or employee), as a result of real time communications, such as phone or video calls, and non real time options such as email, SMS, chat sessions. Even though the conversation is singular, the modes of communication may be multiple: a chat session that escalates into a voice conversation can be contained in a single vCon.

A definition
Conversations are inherently hard to define: when did it start? Who was involved? vCons are the definition. What constitutes a conversation, the participant and the media, is defined by the person or application creating the vCon.

A vCon contains information in four categories:
- **Meta Data** - including the parties involved and their identities
- **Dialog** - the captured media of the conversation in it's original form (e.g. text, audio, video)
- **Analysis** - transcriptions, text to speech, summaries, sentiment, translation, etc.
- **Conversational Documents** - documents discussed or presented during the conversation (e.g. NDA, contracts, presentations, research data)

A computer file format. A vCon is a computer file format for storing the details of a real-time conversation. At a minimum a vCon contains a start time, stop time, and duration of the conversation. vCons are represented in JSON (JavaScript Object Notation) format.

Shareable. Like any other file type, vCons can be attached to emails and shared, stored in databases and file systems, and analyzed for content.

Lossless: A vCon may contain the raw and processed transcripts of the conversation or links to those files; pointers to audio, video, text, documents or URLs shared, automatic speech recognition text, sentiment data, engagement data, current speaker, actions points, questions, summaries, all appropriately timestamped.

Metadata for the conversation. Conversations have a topic, purpose, and goal. The identities used by the participants of the conversation such as phone number, email, corporate ID, SSN, name and address, customer ID; location of the participants during the conversation (this could impact PII rules and where the vCon can be processed); devices used by the participants during the conversation such as PSTN voice line (would only have access to the audio content), iPhone, desk-top phone; <location could be derived from the phone number, a business would only want to pay once then attach to the conversation>

Other examples on the value of metadata for improving CX include reading/typing styles e.g. words per interaction. Adapting the conversation style to each individual customer, fast versus slow readers, focused versus unfocused readers. The full transcript with timestamps recorded in the vCon can be used to create metrics to optimize each customers' experience.

Tamper protections. For example JOSE (JSON Object Signing and Encryption) JWS (JSON Web Signature) RFC 7515 https://datatracker.ietf.org/doc/rfc7515/.

Certifications. Here are a couple of scenarios:
- A carrier knows both participants of a conversation were in Germany as the call was made over their network in Germany. Hence, they can certify the conversation was made in Germany. Personal data is constitutionally protected in Germany. This means individuals have the power to decide when and to what extent personal information is published.
- The conversation could be processed to confirm no PII was shared in this conversation. A business would only want to pay once for that service and then attach the certification to the conversation.

vCon Creator specific data. Guidelines for vCon creator specific data.

Assumptions on the creation of vCon

A vCon could have a call recording attached to it, but it could as easily have not, or only have a partial recording / transcript in compliance with PCI-DSS or GDPR.. For example, with conference calls that include 3rd parties, if recording is enabled a vCon would be created, and when recording is stopped, the vCon stops. It's possible only segments of the conversation are recorded, for example, for compliance purposes only one person on the conference call needs to be recorded. This could lead to a vCon with one or more recordings attached, labeled with a start time for each recorded segment.

vCons are not automatically created unless all parties are notified, e.g. when a customer has a conversation with a contact center they are informed that the conversation may be recorded for training purposes, or all employees know all their work conversations are recorded automatically for compliance as it's in their contract.

A vCon is simply a file, and it is subject to the creators' IT policies. A policy could be 'in confidence' conversations are automatically encrypted and stored in a password protected folder with limited access based on IT policies. Or a vCon between 2 or more organizations is stored within the shared workspace of those companies, and can not be shared outside those companies. vCons do not contain confidentiality marks.

A vCon can link to other files, e.g. the audio or video recording.

vCons are created when they are exported from communication systems like UCaaS (Unified Communication as a Service), PBX (Private Branch Exchange), trunking provider, or from APIs (Application Program Interface) like CPaaS (Communication platform as a Service) providers. Note, most enterprises continue to run on legacy PBX that do not have such programmability, though an application could be created that takes the legacy PBX CDRs and creates vCon, so the legacy PBX remains unaffected. vCon is dependent on either the enterprise's internal communications platform, or one of its communication suppliers having a programmable communications platform or adapting legacy platforms.

Conversation tools take vCons as an input to detect and obfuscate personal information such as credit card numbers. Or, artificial intelligence software uses vCons to store standard sets of test data for development, or for the addition of semantic information in post processing (this customer is angry; this conversation was about pickup trucks). The vCons can be updated with this data to avoid replication of effort, and aid business intelligence and machine learning.

vCons can be shared, and may contain links to files, e.g. the audio / video recording, for example in sharing a conversation within an enterprise where both parties can access the same audiovisual file. There will be a need to define an encapsulated vCon. A file that contains all the data within one file to aid sharing. For example, the enterprise's CPaaS provider delivers a signed vCon of a conversation held over their platform. The content and context are contained within the vCon as it is passing between security domains.

How we will Create the vCon Open Standard

We have a two pronged approach to creating the vCon standard: through the open source community, published as a library and integrated to popular platforms such as Freeswitch and Twilio, and as an IETF standard. To aid in the adoption of vCons, we are proposing it as an open standard, without any intellectual property encumbrances.

Open Source Libraries

See <link to Github repository> for the vCon specification, libraries to connect with to popular programmable communication platforms like CPaaS, UCaaS (Unified Communications as a

Service), CCaaS (Customer Communications as a Service), ASR (Automatic Speech Recognition), and conversation intelligence platforms.

Internet RFC

The Internet Engineering Task Force (IETF) is the organization that defines Internet Protocol standards (e.g. HTTP, RTP, SIP, TCP) that make the Internet work. The IETF is made up of participating individuals, not companies/sponsors, and works as a meritocracy as opposed to paid subscriber voting. IETF standards are published in documents called Requests For Comment (RFC).

To propose an RFC for discussion, typically an individual authors an Internet Draft (ID). The ID is then reviewed in the IETF and may get assigned to a Work Group (WG) and adopted as an IETF item to be iteratively reviewed, edited and refined with the hope of becoming an RFC. Alternatively, a ID can get adopted by a member of the Internet Engineering Steering Group (IESG), a group of advisory individuals elected in the IETF, and the author may continue the process of becoming an RFC as an individual submission.

We plan to author an IETF ID and take the proposal to the IETF, in the hope that we can get it adopted as an RFC. By creating the open source implementation of the vCon, we hope to gain experience and credibility of our proposal to the IETF, demonstrating merit through active working code.

vCon Use Cases

Please note most of these use cases are voice-centric, we see voice conversations as an initial opportunity. However, vCon works for conversations across any communications media, e.g. email, video, SMS, web chat, social, chat in IP messaging like WhatsApp, etc.

Think of vCon as 'robot food', enabling conversation data to be presented in a common format and more easily cleaned for training of machine learning. ASR and conversation AI solutions do not meet the needs of some businesses with respect to accuracy, vCon will help our industry close the gap with respect to the hype.

The performance of ASR varies greatly depending on the application, quality of the recording, and engine/training. ASR continues to improve, some of these applications could be a stretch for a legacy call center, however, for some scenarios they are attainable today.

Service Provider Migration

When companies merge or are acquired, they consolidate suppliers to gain economies of scale. However, migrating data from an acquired company's UCaaS / CCaaS onto the acquiring company's platform can be problematic.

Often the acquired company's customer data and historical conversations are lost, or an expensive data migration project is undertaken. Businesses discover their customer conversation data is not really their data.

With vCon such a migration becomes an export / import of the historical conversations.

This could also be applied to a personal use case where an individual wants to download all their communications from a social network.

Simplifying Working with Bots

Many programmable communication developers share the same experiences when they begin working on voice bots and transcriptions. They begin by capturing the conversation data with MS Word or Excel files. Often a combination of both to better represent conversations and enable the transcriptions to be processed / cleaned-up.

Later they build a JSON format that includes conversation data, transcriptions, context, etc. Soon they begin writing macros / functions to automate conversation data processing and analysis. With that comes maintaining both the JSON format and the macros / functions, and a realization they do not have the time to build everything themselves. vCon solves this common problem and enables macros / functions to be written to this common format and contributed to the open source project.

Speech Recognition Test Set

ASR (Automatic Speech Recognition) for many languages is still in development. That is why Le Voice Lab exists, a French association that brings together various institutional players (universities, research laboratories, etc.) and private companies whose common interest is to build an independent ecosystem and common standards to enable France and Europe to remain competitive in the global voice market. It's not just Europe, around the world there is a substantial quality gap compared to the English speaking world for ASR.

But which ASR works best for business's customers, and is it good enough for the intended applications? Enterprises are now equipped to easily compare the different ASR engines from global and regional providers. vCon enables a single source of test data to accurately and

repeatedly measure speech recognition performance across 100s or 1000s of samples, to gather statistically meaningful performance data.

As it's a computer file format the vCons can be processed through an Excel sheet or business intelligence application. Businesses can make quantified decisions based on their specific situation. The 'Rolls Royce' ASR may be the best with an accuracy range of 94-96%, but the 'Honda Civic' ASR is good enough at 92-94% for the intended application. The vCons from the different ASRs can be processed through the business application, and the business results compared, not just word error rates.

A business may receive 95% of their voice calls from 3G mobile networks with a range of dialects. They can build their own vCon test set, run them through the ASRs, and with nothing more than Excel compare the results. It could be that ASR in general is not currently up to the task, this will change given the continued performance improvements, but better to make an informed decision and revisit; than assume ASR is inadequate until the gap with competitors becomes clear and leaves your business struggling to catch up.

vCons democratize an opaque industry, which relies on fear, uncertainty, and doubt to stop the buyer making a quantified decision that is best for their situation.

PII Compliance

The problems created by not maintaining PII compliance is more than negative publicity. The fines have been massive, for example the FTC (Federal Trade Commission) fined Facebook $5 billion in 2012, Equifax was fined at least $575 million in 2017 and 2019, and British Airways was fined $230 million in 2018. There's a long list of regulations including: GDPR, HIPAA, CCPA, PCI DSS.

It's not just the big brands that get caught, a PII audit can happen to any business. They are not scheduled, and can be triggered by a complaint that can come from an unhappy ex-customer or even a competitor. There are fines and possible incarceration for not reporting PII breaches as well.

PII can be a person's name, in combination with any of the following information:
• Mother's maiden name
• Driver's license number
• Bank account information
• Credit card information
• Relatives' names
• Home Postal address
• Personal E-mail address
• Home or cellular telephone number
• Personal characteristics / biometric data

- Social Security Number (SSN)
- Date or place of birth
- An individual's title
- Work telephone number
- Official work location/address
- Work email address
- Asset information, such as a car's Vehicle Identification Number (VIN) or title number. Even MAC (Machine Address Code) or IMSI (International Mobile Subscriber Identity)
- Even a vCon record given voice finger-printing
- Other information that would make the individual's personal identity easily traceable

Across all the conversations a business has, both internally and externally, which are often recorded for training purposes or by company policy, there is a significant repository of customer data missed by most PII tools because they tend to focus on text based files.

Existing PII tools search the company's storage for files containing customer data, e.g. usernames and password. vCon makes conversation data more easily available to the existing PII audit tools, so your business is better protected. Note, some PII audits are now warning of the emerging need of including audiovisual data for PII Compliance.

Through the open standard vCon the existing PII Compliance tools can be extended to files containing audio visual conversations that are often overlooked. A business is not trapped into multiple specialized PII audits across its different communication silos.

Customer Experience Improvement

Many of us have seen the data, 58% of customers say that customer service is a very important factor that affects their choice of a brand, source Microsoft's State of Global Customer Service Report.

Call recording has been in place for decades, yet why have contact centers remained the same for decades? "Please listen carefully as our menus have changed." No they have not, the menu has been the same for the past two decades! Why hasn't call recording led to a virtuous circle of improvement?

Here are some of the claims made by call recording vendors:
- Call center managers can review the calls to get a better understanding of how agents handle customer conversations.
- Learn whether support representatives are following the protocols.
- Figure out specific customer support aspects that can be improved.
- With call recordings, call center managers can save time as they do not have to listen to each call in real time.
- Identify the gaps in terms of training and best practices whether they are followed or not.

- Learn first-hand customer feedback or issues and train your team to handle them better.
- Listening to call recordings one on one with employees will empower managers to identify the skills that need to improve and work upon.
- Based on the recorded audios, call center managers can prepare presentations to teach about the proper way of making and taking calls.

It's rather manual and lacks quantified data and analysis, this reflects the historical limits of ASR. The call center manager will call a meeting based on their analysis, or bring in a consulting firm that analyzed the data, and based on 'industry best practices' make improvement recommendations. The training to implement the improvements is given, and when the call center manager evaluates the results, not much has changed. It's been going on for decades. Often the sample size is too small given the highly manual approach, and sometimes changes to the process have unintended consequences blurring customer feedback.

vCon enables tens of thousands, even hundreds of thousands of calls to be analyzed by a broad ecosystem of innovative companies. Changes in the process can be analyzed by A/B analysis on possible ways to diffuse customer frustration about an overage charge. The vCons for the calls can be analyzed to make sure the A/B script is being followed, sentiment analysis, and customer feedback during / post call can be combined to produce quantified results.

Often great customer service is exemplified by employees going above and beyond what is economic for the business. While some companies like Zappos simply ensure a human approach to the customer. This may not work for all brands and situations. But being able to test and quantify the results over thousands or tens of thousands of conversations gives the business, and most importantly the agents, confidence in the process change. This is an example of a complex human factor problem that will require new approaches to conversation intelligence, enabled through the ecosystem made possible through vCon.

Sharing Conversations with Third Parties

Working with contractors has its ups and downs. During meetings note taking can miss agreements on deliverables and their timing, simply not happen, or not be fully understood by all parties. Everyone leaves with a slightly different perception of the action. Missed deliverables become finger pointing exercises that do little to help the working relationship.

Most conference calls are recorded, yet are only used for those that missed the meeting and listened to at 1.5-2 X speed. A call recording is generally not used by the people who attended, because they already attended. However, the content in that call is invaluable. It's not that we forget what was agreed on the call, it's that we get busy, when the actions / deliverables are in black and white on our to-do list, we are reminded every day.

Imagine after the meeting a vCon is produced. The vCon is processed by a Meeting Minder - Construction Edition app, which sends a summary to all participants on the conference call that

includes the actions and deliverables. The app is trained to the specific vocabulary and processes used in an industry vertical, construction.

Each action/deliverable includes a reference to the section in the conversation. For example, "Action: Plan for Floor 5 will be sent by Joe to Mary at Company X by end of day June 7th, spoken by Anne,<link to specific point in the conversation>". There is no finger pointing, immediately after the meeting the summary was sent to everyone within the meeting minutes, and transferred onto each participant's to-do list. If there was an error, it would have been picked up then.

Practically, ASR on a conference call is far from perfect, especially if someone calls in using a mobile phone. But the actions and summary will be checked by some of the participants to make sure everything is captured adequately.

vCon enables a range of value added services to be created around this open standard, an ecosystem of innovators taking conversation intelligence beyond transcription. No longer does the Meeting Minder app provider need to waste time and expense joining every conference call / collaboration platform's partner ecosystem. With vCon they can implement one format and cover all communication platforms that use vCon. Instead focusing on building best in class industry vertical solutions.

The integration headache for the Meeting Minder app provider results in less competition. They will focus on the top 5 conference / collaboration platforms. The open source conferencing platforms, the regional platforms, and smaller providers are skipped. vCon helps maintain healthy competition and the reach of vCon apps to extend across all programmable communication platforms.

Verbal Agreements - Converted to Writing

As a general rule, the law in the UK considers verbal contracts to be as legally binding as written ones, and therefore they do hold up in court. But this is an exception compared to the rest of the world. Generally, for established business relationships it makes doing business for small projects easy. Often just an email that summarizes the offer, a consideration, and acceptance is enough, the project gets delivered, and everyone is happy.

However, when legal gets wind of such dealings the $5k project that was supposed to be delivered next week doesn't happen as the 'standard' T&Cs require contractor insurance with $X million+ liability coverage that takes more than a week to arrange. Plus all the time and effort in word-smithing far exceeds the project size.

Using a vCon on the call, chat, or email thread to discuss and agree the project, and using an app such as Meeting Minder - Contract Edition, a verbal or text agreement can easily be converted into a written one, with lightweight T&Cs begrudgingly accepted by legal for small

projects. Now legal are not unhappy and the business can focus on operations, not document editing, with just a call creating a lightweight contract for review and signing.

Note ASR is not necessarily required in this use case. The conversation could be transcribed by a person, in some medical use cases such manual transcription continues to be required. A contract could have a value to make manual transcription economic or potentially preferred.

Insights from Customer Conversations

The customer is constantly providing information to your business: on which competitors they compare you to, what features they value in your product, what problems they have with your product, how they use your product, what features they do not use, how they like to be billed, how they view the pricing, why they moved from / to a competitor, how they use your product, how they mash-up your product and others to meet their needs, how they work around gaps in your product, etc.

Every conversation (voice, email, video, SMS, web chat, IP messaging, social, etc.) across sales, support, customer care, etc. is capturing this data. Extracting this data is not that easy, it's still a work in progress. vCon enables an ecosystem of analytics and data mining companies that can extract such information across all interactions. Determining that a customer is talking about a complementary product not a competitive product is not easy. But insights continue to improve.

vCon enables all customer data to be used to generate insights from the conversations stored but not fully mined. Surveys are often performed to capture minable data, often the surveys have leading questions or the survey is filled as fast as humanly possible. This does not provide accurate information, and why most surveys seem to enforce a marketing view not necessarily a customer view. The conversations taking place with sales, support, customer care are the most accurate customer data a business has access to, and remain relatively poorly mined.

Authenticating and Certifying Conversations

A carrier / telco knows both participants of a conversation were in Germany as the call was made over their network by devices physically connected to their network. Certifying the conversation was made in Germany places additional restrictions on how the data can be used to protect both parties. Personal data is constitutionally protected in Germany. This means individuals have the power to decide when and to what extent personal information is published.

To comply with the Dodd-Frank act's call recording regulations, companies must keep all communication records made through the telephone, voicemail, email, and others, and these records must be uniformly time stamped. A carrier could provide this as a service, recording all conversations made by a business's mobile phones, and sharing with the business as a vCon. The conversation could additionally be processed by the carrier or CPaaS/UCaas/CCaaS to

confirm no PII was shared in this conversation, or certify no keywords specified by the business were found. A business would only want to pay once for that service and attach the certification to the vCon.

Phone numbers can be spoofed. But within a carrier's network they have knowledge of the device's identity, its location, its phone number, its SIM (Subscriber Identity Module) card identity, the owner's account, how long they have had that number or device, etc. There is no other agency that can confirm the identities of the parties on a call with such confidence. A carrier could certify the identities on a vCon.

Example Implementations

Example Twilio Implementation

To implement a vCon engine for a platform, the process is fairly simple: register for completed conversations, fetch the data when they complete, then create the vCon and distribute it. For instance, the figure below shows how we would integrate with Twilio.

1. On startup, the vCon library registers for events from the Twilio platform.
2. When a conversation ends, Twilio delivers a onConversationRemove event using a webhook.
3. Based on the session identifier, the vCon library traverses the Twilio API to fetch transcripts, wav files and identities.
4. The vCon library creates and signs the JSON and stores it in a local file system.
5. Optionally, the vCon can be delivered via webhook to any third party platform.
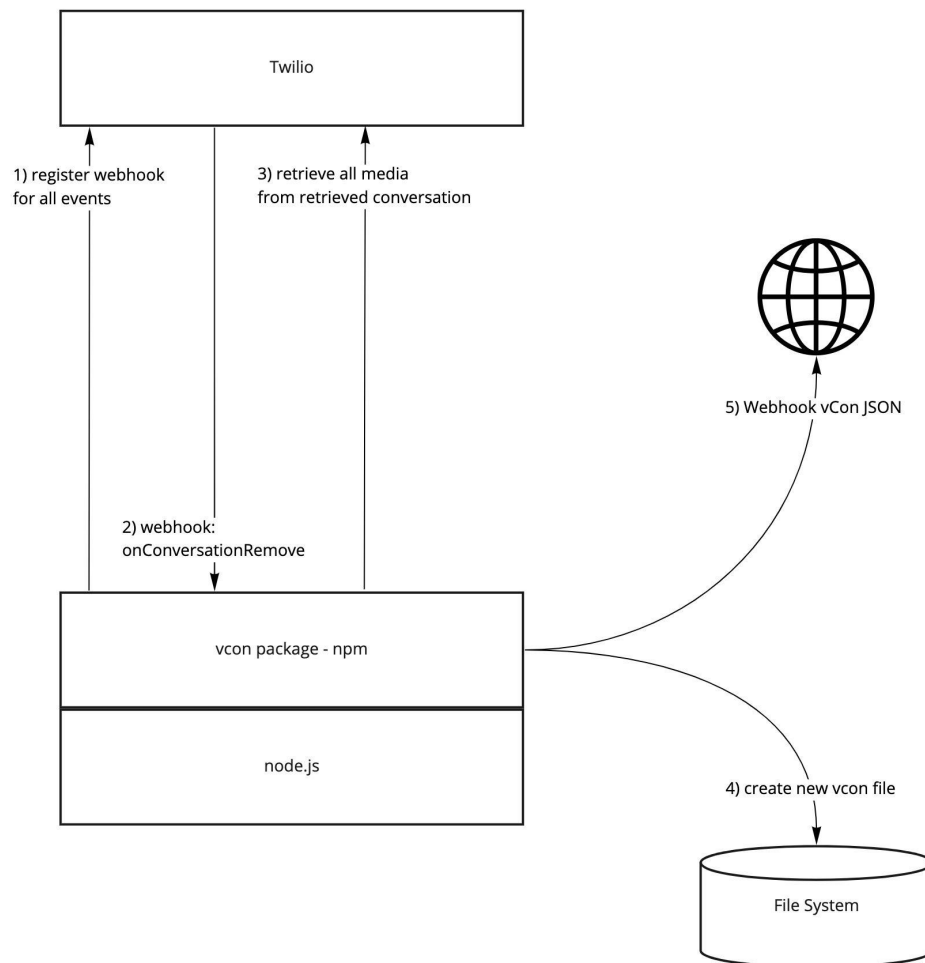
Vision and Industry Impact

This whitepaper provides an introduction to vCon, and an invitation to join the creation of an open standard whose aim is to greatly improve the programmable communications industry. Check out the vCon github repository and get involved, it's open to everyone.

Sharing conversation data is a significant problem faced by programmable communications developers today. vCon makes working with conversations easier, hence the addressable market of developers expands ten thousand fold across web and enterprise developers. This happened one decade ago when telecoms became easy to use with simple web-centric APIs. This will happen again for conversations thanks to vCon.

An innovation ecosystem of specialists in conversation intelligence will flourish solving specific business pain points across operations, compliance, privacy, security, ethics, etc. Being able to access customer data often trapped within communication platforms. Think of vCon as 'robot food', enabling conversation data to be presented in a common format and more easily cleaned for training of machine learning. ASR and conversation AI solutions do not meet the needs of some businesses with respect to accuracy, vCon will help our industry close the gap with respect to the hype.

Analogy to Shipping Containers - A Common Unit Transforms an Industry

For thousands of years, methods of shipping goods across the seas and oceans remained the same. Products came to port in wooden crates, sacks, and kegs by wagons or, later, by trucks and trains. Ships were then loaded and unloaded crate by crate, sack by sack, and keg by keg. It was a time-consuming and labor-intensive process. Theft was a perpetual problem. Often a ship spent more time in ports, loading and unloading, than it would spend at sea.

The advent of World War II brought new logistical challenges in supplying millions of U.S and allied troops overseas and innovative approaches were needed to efficiently supply the war effort. During this period, the introduction of small, standardized boxes full of war material increased the American convoys' capacity to deliver wartime necessities.

After the war, a trucking entrepreneur named Malcom McLean bought a shipping company and, in 1956, started the practice of transporting product-filled truck trailers that were lifted directly from truck to ship. Whole containers, not just small parcels, now moved efficiently onto ships. This transportation process, called intermodalism, allowed products to be shipped around the world quickly, cheaply, and efficiently by using cargo containers that more easily fit on trucks, trains, and ships.

The arrival of containers and intermodalism[1] revolutionized the shipping industry. Containers could be efficiently stacked, allowing more and more goods transported across the seas. Labor costs dropped dramatically and, since containers were sealed, theft declined. Over time, the marine transportation industry and the size of ships, trucks, trains, docks, and ports increased and expanded to handle the growing use of containers. The impact on global commerce was enormous, leading to a boom in international trade due to lower transportation and handling costs.

---

[1] Moving large-sized goods in the same steel-based containers through two or more modes of transport. It's a typical way of moving goods in modern times. Intermodal transfer may involve truck, rail, ship, and then truck again.

A vCon is equivalent to a shipping container for conversations. No longer do people have to listen to conversations and handcraft insights. A global industry can use vCons to solve business problems at scale.

An ecosystem of vCon app providers can automate simple tasks such as actions from a meeting. Imagine after the conference call a vCon is produced. The vCon is processed by a Meeting Minder app, which sends a summary to all participants on the conference call that includes the actions and deliverables. Each action/deliverable includes a reference to the section in the conversation. For example, "Action: Plan for Floor 5 will be sent by Joe to Mary at Company X by end of day June 7th, spoken by Anne,<link to specific point in the conversation>". Making the container standard, vCon for conversations, enables global scale and innovation. Insight from call recording remains a highly manual industry today.

Containerization resulted in global trade accounting for about 60% of Global GDP (Gross Domestic Product) today, from 22% in 1967. See the picture below from the McKinsey Report, Container shipping: The next 50 years. vCon impact will be much faster, see the Open Banking example on the timeline we expect.

Exhibit 2 **Global trade took off in the 1800s and picked up again—this time with the help of container boxes—after a hiatus from 1914 to 1945.**

World exports and imports of goods and services

Share of world GDP, %

Estevadeordal et al., "The rise and fall."[1]   Klasing et al., "Quantifying."   Penn World Tables (v. 8.1)

First wave of globalization

Second wave of globalization

59

29

18

22

10

6

3

3

1

1500   1600   1700   1800   1900   1945   2015

1870   1913   1967

[1]Simple average of lower- and upper-bound estimates.

Source: Antoni Estevadeordal, Brian Frantz, and Alan Taylor, "The rise and fall of world trade, 1870–1939," *Quarterly Journal of Economics*, May 2003; Esteban Ortiz-Ospina and Max Roser, "International trade," OurWorldInData.org; Mariko Klasing and Petros Milionis, "Quantifying the evolution of world trade, 1870–1949," *Journal of International Economics*, January 2014; Penn World Tables Version 8.1; McKinsey analysis

When SMS was only in Silos - Removing Silos Created Multi-Billion Dollar Industries

SMS is ubiquitous, you can reach almost everyone on the planet with a text. Application to Person SMS is used by most businesses for alerts, notifications, security with 2 factor authentication, and chat. But in the beginning SMS only worked within a carrier, you could not send SMS between carriers in the same country. SMS remained trapped in silos.

SMS as a commercial service in the UK was launched in 1995, however, texting didn't take off. By 1998 it was finally possible to exchange an SMS between the four major British mobile phone networks. At that point SMS exploded from 1 billion in 1999 to 30 billion in 2005. With lots of stories in the press about children racking up $1000 mobile phone bills through SMS. In the US carrier interoperability did not happen until 2002, and followed a similar explosion.

vCon enables a range of value added services to be created around this open standard. No longer does the Meeting Minder app provider need to waste time and expense joining every conference call platform's partner ecosystem. They can not, in practice they focus on the top 5, and that limits competition and the smaller providers have a competitive disadvantage. With vCon they can implement one format and cover all communication platforms that use vCon. The app can focus on building best in class industry vertical solutions. Removing silos enables industries to explode like SMS!

Open Banking, a Template for what vCon can Achieve

More than five million people in the UK are actively using open banking services, four years after being mandated by the competition regulator. It took 10 months to grow the number of users from one million to two million in 2020, and it has taken just four months to grow from four million to more than five million by the end of 2021. It is estimated that by September 2023, 60% of the UK population will be using Open Banking. In just 4 years open banking reached ubiquity.

On January 2018, the Competition and Markets Authority (CMA) forced the nine largest UK current account providers (Allied Irish Bank, Bank of Ireland, Barclays, Danske, HSBC, Lloyds Banking Group, Nationwide, RBS Group, Santander) to open up their data. This group of banks is referred to as the CMA9.

Open Banking enables consumers and small and medium-sized enterprises (SMEs) to share their bank and credit card transaction data securely with trusted third parties who are then able to provide them with applications and services which save time and money.

The Open Banking ecosystem in the UK now extends far beyond the CMA9 - currently comprising more than 330 regulated firms made up of over 230 third party providers of services and more than 90 payment account service providers. Examples of services provided include: lowering overdraft costs through bundling, help with budgeting and affordability checks, debt advice and management, income smoothing and budgeting for the self employed and gig workers, optimizing charitable donations, etc..

The stodgy old industry of banking has been transformed through the creation of an open API that enables consumers and small and medium-sized enterprises (SMEs) to share their bank and credit card transaction data securely with trusted third parties. A vibrant and regulated ecosystem has been created. vCon enables conversation recordings to be used by an ecosystem of service providers. Regulation is not required, the vCon open standard will enable the many smaller providers to differentiate, and through enterprise adoption encourage the larger platforms to follow suit. Innovation comes from the smaller companies!

At [TADSummit](#) over several years we've tracked Open Banking. It's a template for other industries including open telecoms. Miles Cheetham was Head of Propositions at Open Banking and gave several presentations. We recommend this presentation on [Open Banking…Energy…Telecom](#) from TADSummit 2020.

PDF File Format

PDF (Portable Document Format) is a reliable format to export documents that you want to share with others or release publicly, so that the information that you want to convey is conveyed in exactly the same way as it was intended to. Its intent is similar to vCon.

PDF, ISO 32000, is a file format developed by Adobe in 1992 to present documents, including text formatting and images, in a manner independent of application software, hardware, and operating systems] Based on the PostScript language, each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, fonts, vector graphics, raster images and other information needed to display it. In 2008 it became an open standard.

The basic types of content in a PDF are:

- Typeset text stored as content streams (i.e., not encoded in plain text);
- Vector graphics for illustrations and designs that consist of shapes and lines;
- Raster graphics for photographs and other types of images
- Multimedia objects in the document.
- In later PDF revisions, a PDF document can also support links (inside document or web page), forms, JavaScript (initially available as a plugin for Acrobat 3.0), or any other types of embedded contents that can be handled using plug-ins.

PDF combines three technologies:

- A subset of the PostScript page description programming language but in declarative form, for generating the layout and graphics.
- A font-embedding/replacement system to allow fonts to travel with the documents.
- A structured storage system to bundle these elements and any associated content into a single file, with data compression where appropriate.

The benefits of PDF include:
- Portability - since it became an open standard in 2008, many apps both read and allow save-as or export as PDF.
- Compatibility - it just works, no missing images or weird fonts.
- Reliability - its the lingua franca of documents.
- Ease of Creation - it's supported across many applications.
- Security - password protection
- Version Independence
- Annocation for collaboration
- Versatility - across documents, presentations, graphics, even data files.

PDF is part of many established applications, just like there are many established programmable communication applications, e.g. CPaaS, UCaaS, CCaaS, that could use vCon. Also there are many specialized PDF applications, and with vCon there could be many specialized vCon applications, e.g. meeting minder use case, PII tools, contract creator, communication graph creators, etc.

What Does this Mean for Programmable Communications?

Through the creation of a common unit of exchange and openness, industries have been transformed. Global trade in the 50 years from 1967 to 2017 went from 22% to roughly 60% of global GDP. SMS in 5 years grew 3000%, and created multi-tens of billion dollar industries across CPaaS and A2P SMS. Open banking in the UK became ubiquitous in 4 years.

vCon will do the same. Conversations currently trapped in silos of CRM (Customer Relationship Management) or simply stored in the company's data lake will become common units, where innovators will create new value, solve business problems, and deliver valuable insights because their business lives/dies on delivering that. Rather than some 'AI marketing spin' for the sale of a contact center or cloud PBX.

Verified Market Research estimates the Global Customer Analytics Market size at USD 5.24 Billion in 2020 and is projected to be $21B by 2028, a CAGR (Compound Annual Growth Rate) of 20%. Driven principally by siloed solutions, e.g. Tableau in SFDC and Google BigQuery.

Through vCon, an ecosystem of analytics companies can thrive by focusing on solving specific business problems, not generic capabilities.

Employee Analytics is a smaller and less clearly defined market, generally < $1B, mostly HR tools. However, we often experience a gap in meeting follow-up after a conference call. We've built the technology to make collaboration ubiquitous. But the workflows around meetings remain trapped in the 1990s. vCon enables innovators to focus and solve specific industry vertical problems as the meeting data is not trapped within a system or across enterprise platforms, rather available in an easily shareable unit.

Sharing conversation data is a significant problem faced by programmable communications developers today. vCon makes working with conversations easier, hence the addressable market of developers expands ten thousand fold across web and enterprise developers. This happened one decade ago when telecoms became easy to use with simple web-centric APIs from CPaaS. This will happen again for conversations thanks to vCon.

An innovation ecosystem of specialists in conversation intelligence will flourish solving specific business pain points across operations, compliance, privacy, security, ethics, etc. Think of vCon as 'robot food', enabling conversation data to be presented in a common format and more easily cleaned for training of machine learning. ASR and conversation AI solutions do not meet the needs of some businesses with respect to accuracy, vCon will help our industry close the gap with respect to the hype. vCon is the next leap forward in the programmable communications industry.

How to Get in Contact

https://github.com/vcon-dev/vcon

Appendix 1. Acronyms

A2P SMS   Application to Person Short Messaging Service, e.g. Amazon delivery notification.

AI   Artificial Intelligence, were machines that mimic and display "human" cognitive skills, e.g. speech recognition.

API   Application Program Interface, set of definitions and protocols for building and integrating application software.

ASR   Automatic Speech Recognition, speech to text service.

CAGR   Compound Annual Growth Rate

CCaaS   Customer Communications as a Service, contact center provided on demand.

CDR   Call Detail Record. data record produced by a telephone exchange or other telecommunications equipment that documents the details of a telephone call or other telecommunications transactions (e.g., text message) that passes through that facility or device. The record contains various attributes of the call, such as time, duration, completion status, source number, and destination number.

CMA   Competition and Markets Authority (UK)

CMA9   Nine largest UK current account providers (Allied Irish Bank, Bank of Ireland, Barclays, Danske, HSBC, Lloyds Banking Group, Nationwide, RBS Group, Santander)

CPaaS   Communication Platform as a Service. A Company that connects to multiple public and private communication networks, and exposes their services, like voice, video, messaging, phone numbers, through an API.

CX   Customer eXperience, anything related to how a customer experiences a business's service.

DSL   Digital Subscriber Loop, a family of technologies that are used to transmit digital data over telephone lines for broadband internet access.

FTC   Federal Trade Commission

GDP   Gross Domestic Product

iCal   Internet Calendaring and Scheduling Core Object Specification (iCalendar) is a media type which allows users to store and exchange calendaring and scheduling information such as events, to-dos, journal entries, and free/busy information

IMSI   International Mobile Subscriber Identity

JOSE   JSON Object Signing and Encryption

JSON   JavaScript Object Notation, an open standard file format and data interchange format that uses human-readable text to store and transmit data objects.

JWS   JSON Web Signature

MIME   Media Type, e.g. MP3 (audio).

MP3	MPEG-1 Audio Layer III, an audio compression standard of the Moving Picture Experts Group (MPEG). Default for audio over the internet.
PBX	Private Branch Exchange, a telephone system within an enterprise that switches calls between users on local lines, while enabling all users to share a certain number of external phone lines.
PII	Personal Identifiable Information
POTS	Plain Old Telephone System
PSTN	Public Switched Telephone Network, means same thing as POTS
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol, a signaling protocol used for initiating, maintaining, and terminating real-time sessions that include voice, video and messaging applications.
SME	Small and Medium-sized Enterprise
SMS	Short Messaging Service.
SQL	Structured Query Language, programming language for managing data in databases
SSN	Social Security Number
T&Cs	Terms and Conditions
UCaaS	Unified Communications as a Service, employee communications provided on demand..
URL	Universal Resource Locator, e.g https://www.bbc.com/news
vCard	vCard, also known as VCF (Virtual Contact File), is a file format standard for electronic business cards.
vCon	Virtual Conversation file, is a file format standard for conversations. Enabling the sharing of conversation data: transcript, video, audio, participants, metadata like timestamps and location, tamper protections, certifications, etc.
VIN	Vehicle Identification Number
WebRTC	Web Real-Time Communication is a free and open-source project providing web browsers and mobile applications with real-time communication (RTC).
XLS	Excel Spreadsheet, also the filename extension *.xls