

INDUSTRY WATCH

\$NLP: How to spend a billion dollars

Robert Dale*

Language Technology Group

*Corresponding author. E-mail: rdale@language-technology.com

(Received 16 November 2021)

Abstract

Funding for AI start-ups in general is booming, and natural language processing as a subfield has not missed out. We take a closer look at early-stage funding over the last year—just over US\$1B in total—for companies that offer solutions that are based on or make significant use of NLP, providing a picture of what funders think is innovative and bankable in this space, and we make some observations on notable trends and developments.

1. Introduction

As part of putting together my weekly newsletter on the natural language processing industry,¹ I track funding events in the NLP world. For the period from 1st November 2020 to 31st October 2021,² there were over 300 such events reported across the full range of start-up funding stages.³ For the purposes of the present article, I decided to focus on the 120 or so early-stage companies who received either seed funding or Series A funding in that period;⁴ in broad terms, this restriction in scope correlates with innovative ideas that funders have seen as promising but which have yet to demonstrate an ability to generate long-term profit.

This amounts to a large and diverse space of products and services, so we need a way of structuring the space. It is tempting to attempt to do this either on the basis of some distinction between technology types or on the basis of domains of application. But neither of these works fantastically well here, since many products and services are a combination of multiple technology types – that is often what provides their value-add – and many products and services target multiple domains. So we will start with a slightly uncomfortable high-level structure that is broadly based on technology types:

- Document AI: this category covers technology that works in one way or another with documents. You might think of this as a combination of text analytics and natural language generation, although it also encompasses a few other technology types that fall outside these.

¹Subscribe for free at <https://www.language-technology.com/twin>.

²Ideally, I would have covered a calendar year, but the publication time line for this piece made that difficult.

³I use a combination of human and machine processing to find news of relevant funding events across a wide range of sources; I doubt that my list is 100% complete, but I believe it is fairly comprehensive. If your company lies within the scope outlined here, but I have not mentioned it in what follows, please drop me a note to let me know.

⁴See <https://www.investopedia.com/articles/personal-finance/102015/series-b-c-funding-what-it-all-means-and-how-it-works.asp> for an explanation of the various stages of funding.

- **Conversational AI:** this category covers those applications that engage in an interactive dialog with a user via natural language. I use the category to cover both text-based chatbots and voice-driven virtual assistants.
- **Other Voice Technologies:** there is something of a fuzzy boundary between this and the previous category, but I think it is useful to separate out some voice-related products and services that are not centrally focussed on dialog.

Each of these categories is then decomposed into subcategories, some based on technology and some on domain. It is not perfect, but it is the best I could do; the alternative was a long flat list that you would never read.

With so many companies to cover, the descriptions of each provided here are necessarily extremely brief. I have tried to provide summaries that are at least marginally more insightful than what you might glean from skimming the companies' websites: I reached out to every company mentioned to dig a little into their technologies, and the summaries below are often just the most salient points from the sometimes quite detailed responses received. But there is also a fair number of companies who were, not surprisingly, wary of giving much away, and in some of those cases, I have had to fall back on what can be inferred from the often vague descriptions on websites and in marketing material.

The flip side is that, in the course of doing research for this article, I amassed vastly more information on the companies mentioned here than there is space to convey; so I am planning to use that material to inform deeper dives on some of the categories of products and services discussed here. If you would like to influence which particular areas below I might prioritise, drop me an email at rdale@language-technology.com.

2. Document AI

Of the companies surveyed here, 60 fall within our definition of document AI, made up of text analytics (8 companies), information discovery (11), writing assistance (11), legal tech (18), other domain-specific applications (5) and other document AI (7).

2.1 Text analytics

As we will use the term here, a text analytics application is one that attempts to extract some non-trivial representation of content from the document being processed. This covers document classification, sentiment analysis and extraction tasks such as named entity recognition and relation extraction.

Cortical.io (Series A; US\$6m) emphasise their machine-learned 'semantic fingerprinting' approach to support extraction, classification, comparison and search capabilities over documents. They use this to develop custom solutions, but they also offer two products, Contract Intelligence and Message Intelligence, built on top of this technology; the first of these focusses on legal documents (also see Section 2.4 below) and the second on message streams such as email.

The processing of specific document types is a recurrent theme. **Natif.ai** (seed; unspecified amount) builds information extraction APIs for document types such as invoices, insurance policies and receipts; their stack includes a 'Deep-OCR' layer. **Mindee** (Series A; US\$14m) provide tools for extracting structured data from photos or PDF files; they offer APIs for standard document types such as receipts, invoices, passports, driver licences and identity cards, as well as an API builder for custom requirements.

Some companies focus specifically on text sentiment: **BlueOcean** (Series A; US\$15m) is a toolset for brand management that applies sentiment analysis to text, image and video content with the aim of identifying a brand's strengths and weaknesses relative to its competitors. **Viable**

(seed; US\$3.9m) aggregates, analyses and summarises text-based customer feedback at scale and then provides insights in natural language reports.

Others embed text analytics within broader platforms. **OpenBots** (seed funding; US\$5m) incorporates intelligent document processing within an open-source cloud-based Robotic Process Automation (RPA) platform, supporting OCR, classification and extraction capabilities over a variety of common document types and giving users the ability to create custom document templates. **Rossum** (Series A; US\$100m) provides a platform for its ML-based document processing tools that include aggregation across input channels and a UI for human review and validation.

We are also seeing hybrid human-machine approaches in this space: **Daloopa** (Series A; US\$20m) positions itself as a ‘full-service data extraction company,’ combining human and machine processing to produce customised solutions with accurate extraction results.

2.2 Information discovery

There is a fuzzy boundary between this subcategory and the one above, but I see information delivery as being more focussed on crawling over large document sets, often (but not always) external to the organisation. So you might think of this subcategory as being about enhanced search.

Grata (seed; US\$9.5m) is a B2B search engine for finding and targeting private companies; it uses NLP and ML to extract business information from company websites. **Ferret** (seed; US\$4m) develop text analytics and information extraction tools in service of what they call ‘relationship intelligence’, which aims to help you decide if you want to trust or do business with someone.

Sorcero (Series A; US\$10m) targets the life sciences industry, providing a literature monitoring product that combines deep learning models and industry-specific curated ontologies; these resources are fine-tuneable without coding.

Celential (Series A; US\$9.5m) positions itself as an AI-driven, human-assisted virtual recruiting service: its core resource is its ‘Talent Graph’, built from a wide range of less-used sources, which uses ML and NLP to find matching candidates.

Stravito (Series A; €12.4m) offers a search engine built specifically for market intelligence; it uses ML and ontologies to categorise and add metadata to a company’s internal document and video assets. **MachEye** (seed; US\$4.6m) positions itself as an augmented analytics platform; it emphasises its use of NL query and NLG to understand business questions and generate contextual answers from underlying data stores, including a capability for video generation.

Klevu (Series A; US\$12m) is an e-commerce search engine that uses NLP to enhance queries and augment the content of product catalogues, thus improving the relevance of search results.

Quark.ai (Seed-plus; US\$5m) is a platform that uses NLP to interpret support queries so it can recommend resolutions automatically by returning relevant reference documents to the customer or support engineer. **Allganize** (Series A; US\$10m) offers what it calls an answer bot: a combined search and chatbot solution for automating business workflows for employee and customer support. They provide a typical range of text analytics capabilities operating over a company’s document set, but make these capabilities accessible via a chatbot interface.

There is also a couple of startups that focus on disinformation: **Blackbird.ai** (Series A; US\$10m) aims to surface manipulative disinformation campaigns, misinformation and propaganda, using ‘narrative criteria’ to identify hoaxes and myths on the web; and **Kinzen** (seed; €1.8m) detects and scores information risk in text, audio and video content, using a knowledge graph for disinformation detection and ASR models that are optimised for information risk.

2.3 Writing assistance

The writing assistance space has exploded in the last 18 months with the appearance on the scene of OpenAI’s GPT-3. Earlier large language models were good enough to support simple text

prediction and error correction, but were not confident enough to predict what you might want to say more than a few words out. GPT-3's superior predictive abilities have inspired a slew of writing assistance tools that not only will write several paragraphs of text at the slightest prompting but are also able to offer all sorts of variations on what you have already written.⁵ Not all of the following are based on GPT-3, but any which are based on previous generations of models with lesser capabilities may find it a struggle to survive in such a densely populated space.

AI21 Labs (Series A; US\$25m) is known for its Jurassic-1 family of large language models; but it also has a suite of products built on these models. WordTune is AI21 Labs' co-editor writing assistant that offers rewrite, shorten/lengthen, tone changes and translation functions, as well as grammar and punctuation fixes.

Compose AI (seed; US\$2.1m) has an enticingly simple proposition: it offers a free Chrome extension that claims to cut your writing time by 40% using autocompletion. A paid version also learns your personal writing style.

OthersideAI (seed; US\$2.6m) focusses on turning summaries and shorthand notes into written emails via its HyperWrite product. **Copysmith** (seed; US\$10m) targets marketers, content creators and e-commerce platforms, with its tech being tuned to write copy for ads, product descriptions, social media posts, landing pages and blog posts.

Copy.ai (Series A; US\$11m) positions its LM-based text generator within a broader eco-system for helping people start and run businesses; the first use cases are related to marketing and copywriting.

In another hybrid solution, **Contents** (Series A; US\$6m) offers LM-based content generation, but with optional proofing and checking by humans.

More broadly, **Cohere.ai** (Series A; US\$40m) offers an alternative to the GPT-3 language model, with an emphasis on responsibility.

Text Blaze (seed; US\$3.3m) probably does not belong here since it's basically a mechanism for creating and using text snippets to eliminate repetitive typing – no real NLP there. But its use of customisable templates echoes the way that many current data-to-text NLG products work, and I have a soft spot for simple technologies that do something really useful, so it gets an honourable mention.

There are also a number of apps that attempt to determine whether your existing text is fit for purpose. **LitLingo** (Series A; US\$7.5m) deploys models that help businesses by detecting language use that falls into specific categories, such as risk of litigation or regulatory non-compliance. **Instoried** (venture round; US\$8m) is an augmented writing platform that analyses the sentiment expressed in your marketing content and makes recommendations for adding empathy, 'supercharging conversions' in the process. **Pluralytics** (seed; US\$1m) reads your marketing content, tells you who it appeals to and why and then suggests words and phrases to improve engagement with your target audience 'while remaining authentic to your brand voice'.

2.4 Legal tech

As well as those companies offering what we might think of as domain-independent products and solutions, there are also many who target specific domains. The most strongly represented in terms of document AI is legal tech.

A key focus here is contract lifecycle management (CLM), which covers contract initiation, authoring, process and workflow, negotiation and approval, execution, ongoing management and compliance and subsequent contract renewal. NLP can play a role in a number of these stages; most common are the analysis of existing contracts and support for the authoring of new contracts.⁶

⁵See Dale (2021) and Dale and Viethen (2021) for some discussion of these kinds of applications.

⁶See Dale (2019) for an earlier review of NLP technology in the legal world.

Lexion (Series A; US\$11m) is an end-to-end contract management system built around the idea of a ‘smart repository’ that is populated using classification and information extraction over existing documents. Amongst other utilities, the company offers a Slack chatbot that can retrieve documents for you. **Malbek** (Series A; US\$15.3m) is another CLM platform that includes both a contract repository and contract authoring tools that are tightly integrated with Microsoft Word. **SimpliContract** (seed; US\$1.8m) similarly emphasises smart contract search and storage in combination with easy contract authoring. **Legislate** (seed; £1M) is an end-to-end contract creation and management platform for non-lawyers; it uses knowledge graph technology to generate, negotiate and manage documents at scale. **Arteria** (Series A; US\$11m) applies AI to the drafting, negotiation and analysis of contracts, with an emphasis on the leveraging of its underlying structured data.

Focussing on the contract analysis side, **Zuva** (Series A; US\$15.75m) makes DocAI, an information extraction application that uses 1200+ pre-built machine learning models to identify specific clause types in legal documents. **Semeris** (pre-seed; US\$600k) focuses on contract analysis in the financial sector; a key feature is the ability to compare how the language used in transaction documents varies from deal to deal and evolves over time, abstracting over variations in surface form.

On the authoring side, **Clearbrief** (seed; US\$3.5m) provides support for legal drafting by identifying discrepancies between the document you are working on and the sources you cite, along with other forms of citation support. **Henchman** (pre-seed; €1m) is an add-in for MS Word that identifies and categorises previously written clauses in your contract repository, making them easily to retrieve for re-use. **Contract Mill** (seed; €1m) is a no-code document automation platform that supports sharing and re-use via a clause library. **Definely** (seed-plus, US\$3m) focusses in on defined terms and references, providing the relevant definitions and information in a side panel so you do not have to keep jumping backwards and forwards in the document. **BlackBoiler** (venture round; US\$3.2m) exploits the Track Changes markups in an organisation’s MS Word document repository to create bespoke editing models, enabling it to make company-specific revisions to previously unseen documents. **10BE5** (pre-seed; < US\$1m) automates capital market-related drafting and diligence workstreams; its first product, N2N, is essentially an NLG application for financial disclosure documents.

Beyond contract repositories and document authoring, **Pactum** (Series A; US\$11m) is a contract negotiation application that aims to automate the entire negotiation process from the initial email that begins the negotiation to contract generation and signing, and **Josef** (Unspecified; US\$2.5m) is a legal platform that enables lawyers to automate, build and launch their own legal chatbots or services, ranging from bots to handle client interviews to document automation.

Finally, on the information discovery side of things, **Jus Mundi** (Series A; €8.5m) is multi-lingual search engine for global legal information, built on a legal KB constructed from a large document corpus; **Regology** (Series A; US\$8m) offers a platform that actively tracks regulatory updates to support a comprehensive ‘law library’ knowledge base, allowing companies to dynamically monitor changes to business regulations and helping to ensure compliance; and **Trellis Research** (Series A; US\$14.1m) is a research tool for litigators, providing a ‘smart search’ capability across a database of state trial court records.

2.5 Other domains

Of course, there are other domain-specific applications out there.

In fintech, **Informed.IQ** (Series A; US\$20m) provides an automation capability for lenders, collecting and analysing loan documents using ML models trained on millions of consumer, auto and mortgage credit applications; and **Zelros** (Series A; US\$11m) offers a range of functionalities that build on insurance-specific ML models for extracting information from voice, emails and contract documents.

In healthcare, **HealthTensor** (seed; US\$5m) diagnoses and documents conditions based on the information in Electronic Health Records; and **Mendel** (Series A; US\$18m) transforms unstructured EMR data and clinical literature into compliant analytics-ready data; features include an OCR capability and the ability to redact personally identifiable information.

And in HR, **retrain.ai** (Series A; US\$9m) uses NLP and ML to read job boards at scale, with the aim of gaining insight into where the job market is going.

2.6 Other document AI

Finally, there are a few companies that we would consider to be document AI, but which don't fit neatly into the categories above.

It is still possible to have a startup focussed on machine translation, although you need a new twist to make yourself visible above the many existing MT providers.

Bering Lab (seed; unspecified amount) trains its NMT engine for specific domains, offering models for a wide range of industries; the company also offers post-editing by domain experts, echoing the hybrid human-machine approach we have seen elsewhere. **Language I/O** (Series A; US\$5m) offers an NMT aggregation layer that provides access to a number of third-party NMT vendors; the best performing engine for your language pair on the day of use is automatically selected. A post-processing phase then detects terms and phrases that may need to be corrected, based on the customer's domain data. **Toucan** (seed; US\$4.5m) aims to help you learn a new language while browsing the web: it injects foreign-language terms into the text you are reading.

In the privacy space, **Private AI** (seed; US\$3.15m) uses transformer models to detect over 50 different direct identifiers (like names, SSNs, credit card numbers) and quasi-identifiers (like age and location); it also provides models for pseudonymisation. **Xayn** (Series A; US\$12m) uses a combination of compressed models to provide a self-learning personalised search experience via on-device processing.

On the document automation side of things, **Narrativa** (unspecified round; US\$1.3m) is an NLG company that integrates traditional template-based document generation along with capabilities based on various deep learning models for specific use cases.

And **Reclaim.ai** (seed; US\$4.8m) is a smart calendar application that uses NLP to analyse calendar content in order to provide context-sensitive and event-specific scheduling.

3. Conversational AI

Of the companies surveyed here, 36 fall into our category of Conversational AI, made up of CAI toolkits (6), CAI solution providers (4), e-commerce applications (9), healthcare tech (7), other domains (5) and a category we call 'assistance and analytics' (5).

3.1 CAI toolkits

Perhaps surprisingly, there still appears to be space in the market for yet more chatbot and conversational agent development tools. A lot of these are positioned as no-code solutions; dialog design lends itself well to graphical editing. But we have known that for a very long time, and it is a fact that has been leveraged by most of the incumbents; so I think it will be interesting to see how many of these startups survive in the longer term.

Botpress (Series A; C\$15m) is an open-source text-based chatbot-building platform aimed at those who are not NLP experts; it provides visual development tools and a managed NLP engine for intent and entity recognition. **Humley** (seed; £700k) is a no-code chatbot development tool that supports 13 languages and comes with a number of pre-built conversational assistants for specific use cases and verticals. **Landbot** (Series A; US\$8m) pitches itself as a simple no-code tool

for building conversational websites, with a specific focus on lead data capture, lead qualification and simple personalisation.

Agara Labs (seed; US\$4.3m)⁷ offers a voicebot that is pre-trained to resolve the most common e-commerce queries and a no-code environment for app development; a key feature is the use of a direct speech-to-intent ASR model trained on several hundred hours of customer support phone call recordings, claiming a 20% improvement over other approaches. **Symbi** (seed; US\$4.7m)⁸ advertises a wide-ranging set of APIs for integrating ‘conversation intelligence’ into applications, although many of the components are still in beta. Some of the features provided are targeted at conversational analytics, and others at what we refer to in Section 4.2 as ‘meeting analytics’.

The visual development platform offered by **Voiceflow** (Series A; US\$20m) emphasises its support for team working and the conversational application lifecycle, encompassing design, prototyping and testing.

3.2 CAI solution providers

The companies in this category will build you a conversational AI if you would rather not do it yourself.

Hyro (Series A; US\$10.5m) is a conversational AI platform that focuses on what its makers call ‘adaptive conversation’: eschewing the common intent-based approach, the company will build you a chatbot that uses knowledge graphs and what appears to be good old-fashioned computational linguistics to obtain scalability and reusability. **Senseforth** (unspecified round; US\$14m) offers solutions built using a conversational AI bot store with pre-built models and domain knowledge for a range of verticals including banking, insurance, retail, healthcare, telecom and hospitality.

GUURU (Series A; US\$5m) emphasises its SmartRouting query routing capability: it analyses incoming questions and routes them to minimise cost, with recurring questions being answered by a chatbot, and others being redirected to a user community or sent to your agents, as determined by the specifics of your use case.

Founded in 2006 as a full-service digital and creative agency, **RAIN** (Series A; US\$3m) does not really count as a startup, but its recent funding has allowed it to diversify from being a developer of branded Alexa Skills to building its own SaaS voice-first product (currently in stealth) for a segment of the deskless workforce.

3.3 CAI in eCommerce

Little chatbots that pop-up on shopping websites to remind you that your cart still has stuff in it, or to suggest that you might like to add these orange polka dot socks to complement your lime green track pants, are ubiquitous. But conversational AI in service of eCommerce has many other facets.

At the simpler end of the spectrum, **Charles** (seed; €6.4m) integrates chatbot functionality on a wide range of messaging apps with a range of eCommerce backends. **Goodcall** (seed; US\$4m) targets small businesses with a simple and easy to set up cloud-based conversational AI; its narrow focus around common ‘knowledge skills’ like providing opening hours means it can be set up quickly using a no-code interface.

Heyday (seed; C\$6.5m; subsequently acquired by Hootsuite for C\$60m) provides a conversational AI platform targeted at retailers: its chatbot recommends products in response to user searches. **Atom** (seed; US\$3.4m) is a conversational sales automation tool that integrates with

⁷Just as this article was being written, Agara was acquired by Coinbase for an undisclosed amount.

⁸In November 2021, the company raised another US\$17m in a Series A round.

popular messaging apps; amongst other things, it attempts to predict most likely sales so it can prioritise handing off to human agents.

Webio (seed; €500k) provides chatbot technology focused on credit, collections and payments messaging, supporting blended chatbot/live agent conversations. **Satisfi Labs** (Series A; US\$3m) pitches itself as a knowledge management platform for conversational search and commerce, with ‘expert assistants’ targeting entertainment, hospitality, sports and tourism.

Some products target specific pain points in the eCommerce lifecycle. **Orum’s** goal (Series A; US\$25m) is to automate the hardest parts of outbound calling: it dials multiple numbers in parallel, detecting voicemails, filtering out bad numbers and navigating phone directories before handing off to live salespeople.

And then there is conversational advertising. **Cavai** (venture round; £6.5m) offers a platform that provides support for building chatbot-like interactive ads. **Instreamatic** (Series A; US\$6.1m) goes a step further, delivering voice ads that you can talk to; the app maintains historical context to take account of previous conversations.

3.4 CAI in healthcare

Another major area for the deployment of conversational AI technology is in healthcare.

Andor Health (Series A; amount undisclosed) provides healthcare workers with a virtual assistant which integrates with a patient’s electronic medical record; it also identifies keywords and other critical health content in physicians’ communications, automatically recommending notes and triggers that can be added to patient charts. **Corti** (Series A; US\$27m) listens-in to patient consultations to suggest additional questions that can be asked; or it can listen to call center traffic to watch for signs of critical illness that require escalation.

Bot MD (Series A; US\$5m) integrates a variety of localised hospital information sources to enable answering of clinicians’ queries via a chat interface. **Botco.ai** (seed; US\$3.6m) is a HIPAA-compliant chatbot solution that connects with existing CRM and electronic health record systems, supporting conversations between patients and providers for tasks like booking appointments.

Wysa (Series A; US\$5.5) provides a mental health wellness platform; its chatbot leverages cognitive-behavioural techniques to help users self-manage stressors. The service also makes use of professional therapists in another hybrid solution.

Authenticx (Series A; US\$7.5m) aggregates healthcare organisations customer conversations across multiple channels to provide data for conversational analytics (see Section 3.6). **Virti** (Series A; US\$10m) also provides conversational analytics, but this time in the context of a virtual human training app, with an initial focus on healthcare.

3.5 CAI in other domains

In the human resources space, **BrightHire** (Series A; US\$12.5m) provides an ‘interview intelligence platform’ that integrates with Zoom, supporting recruiters and interviewers with an automated assistant to guide the conversation and capture key highlights in real-time. **Humanly** (seed; US\$4.2m) combines automation of the repetitive elements of conversations with job candidates along with conversational analytics applied to human interviews, with the aim of measuring phenomena like interviewer patience and unconscious bias and their correlation with offer acceptance rate and candidate sentiment.

In education, **EdSights** (Series A; US\$5m) makes a text-messaging chatbot focussed on educational engagement: it aims to detect when students are struggling so it can connect them to helpful on-campus resources. **Merlyn Mind** (Seed and Series A; US\$29m) provides an intelligent voice assistant for teachers via a hardware device with on-board speech processing that enables easy access to a wide range of classroom technology.

ConverseNow (Series A; US\$15m) provides conversational AI that focuses specifically on food-ordering, supported by deep domain knowledge to handle the complexities of food recognition and food dialogs.

3.6 Agent assistance and analytics

There is an increasing trend towards products that are less directly concerned with having a machine interact directly with a human and more about analysing the content of human–human conversations.

A major theme here is what we might call ‘agent assistance’. **Level AI** (Series A; US\$13m) monitors calls with live agents, providing real-time assistance by offering suggested answers to caller queries; it also gathers call analytics data for monitoring agent performance. **Thankful** (Series A; US\$12m) also offers deep learning-based chatbot technology along with an agent assist capability that watches chats with live agents and offers recommended actions and potential replies. **Ultimate.ai** (Series A; US\$20m) is a no-code chatbot development platform with domain-specific components; again, this offers response suggestions for human agents. The app integrates with popular CRM systems.

Another trend in the space is the integration of conversational analytics and supporting tools: these generally provide dashboards that let you aggregate call data and content to get a big-picture view of what is going on in calls, or real-time tools that assess how a given call is progressing along dimensions that you care about. **Staircase AI** (seed; US\$4m), which calls itself a ‘relationship intelligence’ company, analyses digital engagements across a wide range of channels to identify customer issues, red flags and missed opportunities. **Aveni Detect** (venture round; £1.1m) aims to extract insights from conversations between customers and service providers to support quality assurance, using a combination of deep learning models and rule-based approaches. With a focus on the financial sector, the application aims to detect client vulnerability, complaints and concerns around adviser conduct and flags these live during the call.

4. Other voice technologies

This category might be considered to overlap a little with the conversational AI space, but I think it is useful to consider the applications described here separately. We break this area down into three subcategories: transcription (6 companies), meeting analytics (5) and voice synthesis (8).

4.1 Transcription

As speech recognition has improved considerably in recent years, new opportunities for using ASR have been opened up. **Ava** (seed; US\$4.5m) combines ASR with human scribes – yet another hybrid solution – to provide captioning for the deaf and hard of hearing. **Fireflies.ai** (Series A; US\$14m) offers an AI meeting assistant that records, transcribes and makes searchable meeting notes. **Subly** (seed; US\$1m) focusses on automatically transcribing, translating and adding subtitles to videos.

Transcription has always been a key use case in healthcare. There are some interesting new applications here. **Talkatoo** (venture round; unspecified amount) provides transcription targeted specifically at vets, employing a bespoke veterinary-specific vocabulary crafted by practising veterinarians. **InsiteFlow** (seed; US\$2.3m) is a provider of an EHR-integrated platform that helps manage clinical decision solutions; the platform transcribes decisions of clinicians and staff into EHRs. **DeepScribe** (seed; US\$5.2m) targets medical record-taking; it translates informal conversation into doctoresse and integrates this into the EHR.

4.2 Meeting analytics

In Section 3.6, we mentioned some applications that provide conversational analytics. We separate out meeting analytics tools here as a separate category, although they are essentially the same technology applied to conversations that are generally multiparty. The growth in applications of this type has no doubt been fuelled by the COVID-driven uptake in the use of tools like Zoom.

Vowel (Series A; US\$13.5m) pitches itself as a collaboration tool for meetings: it creates a running searchable transcript, provides related productivity tools and integrations and embeds this within an associated meeting ecosystem that supports agendas, integration with calendars and other organisational features. **Sonero** (pre-seed; US\$300k) transcribes virtual meetings, then extracts action items, important topics, key points and questions and answers.

Read.AI (seed; US\$10m) analyses audio to deliver real-time in meeting metrics, covering speaker talk-time, sentiment and engagement. **Aircover** (seed; US\$3m) performs real-time transcription of a sales call or video conference and then analyses this to provide in-meeting sales support.

Poised (seed; US\$4.5m) positions itself as an AI-powered communication coach: it provides personalised feedback and lessons by observing your online meetings, measuring speaker share, how much you use filler words or hedges and speaking pace. Feedback is provided in real time so you can make immediate corrections.

4.3 Voice synthesis

Voice synthesis appears to be a popular area for startups, with, once more, the improvements provided by deep learning models leading to a number of interesting extensions of basic TTS technology.

LOVO (seed; US\$4.5m) offers 180+ voice skins in 33 languages for synthesising speech in a number of genres, such as audiobooks, games and documentaries; it also lets you build a customised voice skin using 15 minutes of data. **WellSaid Labs** (Series A; US\$10m) creates life-like synthetic voices from human samples and provides a user-friendly interface for adding voices and avatars to produce voiceovers from scripts. **Synthesia** (Series A; US\$12.5m) produces videos with audio commentary generated from your textual scripts or PowerPoint files, complete with the avatar of your choice.

Papercup (unspecified round; £8m) translates videos by generating voices that sound like the original speaker, but speaking in the target language. **Sanas** (seed; US\$5.5m) provides real-time accent translation, allowing you to speak in any accent you like without any noticeable lag.

Amai (seed; US\$600K) produces custom voices that can be deployed on edge devices; a key feature of the offering is an editing tool that makes it easy to apply specific emotions to elements of a script and to modify timing. **Humelo** (seed; US\$2.6m) similarly offers a voice synthesis and editing program that allows you to create and edit voices with control of emotion, duration and pitch.

Supertone (corporate round; US\$3.6m) provides a range of real-time voice enhancement technologies, including singing voice synthesis, voice cloning and voice design.

A number of these companies have demos on their websites, which are well worth a look and a listen.

4.4 Other voice solutions

Finally, we have a few other companies working on voice tech applications that do not fit elsewhere above.

There is some interesting work in voice biomarkers: **Kintsugi** (seed; US\$8m) is talk therapy software for mental health; it uses voice biomarkers to measure and predict well-being, using

a neural network model that aims to discover speech patterns indicative of depression. **Ellipsis Health** (Series A; US\$26m) uses machine-learned models for both linguistic content and acoustic and prosodic patterns to detect, measure and monitor the severity of depression and anxiety.

Phonic.ai (seed; US\$2.2m) provides a voice and video survey platform, offering ASR and human transcription in 32 languages, plus automated, manual or hybrid response coding and sentiment and emotional analysis.

Picovoice (seed; US\$500k) offers voice technology embeddable into edge devices, including a voice control solution for smart products with voice activation for audio applications, voice command and keyword spotting. It also offers an NLU engine that can perform on-device processing.

5. Some concluding notes

So there you go: a whistlestop tour of 119 NLP startups that got funding in the last 12 months.

To wrap things up, here are some random observations based on the preceding review, including some cross-category themes.

1. My impression is that text analytics technologies have progressed beyond the broad-purpose toolsets of previous years to focus more narrowly on high-quality analysis of specific document types (Cortical.io, Natif.ai, Mindee, OpenBots). We are also seeing OCR increasingly present as an element of the tech stack these companies offer, likely as a result of deep-learning models providing more acceptable accuracy levels.
2. Similarly, in the conversational AI space, we are seeing a lot of focus on domain-specific elements, particularly in the chatbot-building corner (Humley, Senseforth, Goodcall, Satisfi Labs). Here, one way for a new entrant to mark themselves out is to offer deep but narrow coverage of a domain right out-of-the-box, making app construction increasingly a process of dragging and dropping pre-built elements onto a design canvas with minimal configuration required.
3. Large language models are everywhere. They are most evident in the writing assistance space (Section 2.3), where they allow solutions that were not conceivable before, whereas in other areas, LLMs allow us to do better those things that we were already doing. But it is interesting to see that in many of these other areas, vendors will commonly draw attention to how they combine machine learning approaches with manually curated rules and ontologies.
4. Quite a few startups are emphasising the no-code nature of the solutions they offer (Contract Mill, Humley, Landbot, Agara Labs, Goodcall, Ultimate.ai). This feels like a natural maturation of the technology base: as we better understand the similarities and differences amongst the most common use cases, development interfaces can adopt higher level abstractions that are easier to use.
5. It is now not uncommon in many areas to see hybrid solutions that combine machine and human processing to drive better quality results than can be achieved via machine processing alone (Daloopa, Contents, Bering Lab, Wysa, Ava, Phonic.ai). Human validation is typically pitched as an optional extra, so that you can rely on just the machine tech if your use case either requires it, perhaps for speed of processing reasons, or if your users are willing to tolerate lower levels of accuracy.
6. Legal tech is a busy area, with 18 of the companies surveyed here making offerings in this domain (Section 2.4). That is the largest of the subcategories we have used to partition the space. These subcategories are, as we have already acknowledged, a little ad hoc, but the prevalence of legal tech applications in the document AI space is inescapable.

7. Whereas legal tech dominates the document AI category, e-commerce and healthcare are the key application areas for conversational AI. But what is interesting in conversational AI is the blurring of boundaries already hinted at in the top-level categorisation used above. Originally focussed on conversations with one human and one machine participant, the technology now encompasses multiparty conversation in meetings; it supports humans talking to humans, by offering assistance and analysing performance; and it supports increasingly seamless handover from machine agent to human agent (Sections 3.6 and 4.2).
8. Voice synthesis might be the area that provides the most visible – well, audible – improvements compared with what went before (Section 4.3). The naturalness of some of these synthesised voices is excellent; there are still occasional glitches, but compared to what was on offer even five years ago, we are rapidly approaching a point where synthesised voices will be unnoticeable in everyday use.
9. We have not emphasised this much in the review presented here, but ecosystems are important and a major source of value. Some of the applications discussed here are marked out by their provision of a key technology or technologies embedded within an end-to-end platform (OpenBots, Rossum, Lexion, Legislate, Pactum, Voiceflow) or accessed via an intuitive dashboard interface (Authenticx, Level AI, Staircase AI); in other cases, the ecosystem might just be supporting documentation, ancillary resources or user communities (as in a number of the writing assistance tools in Section 2.3). But it is increasingly the case that just offering a new version of a core technology is not enough: many technologies are becoming commoditised, and just claiming that yours is 10% better will not be enough to close the deal.

Finally, to sum up the numbers: The pre-seed and seed funding average across the companies mentioned here is about US\$3.5m, with a minimum of US\$20k and a maximum of US\$14m. The Series A funding average is around US\$14m, with a minimum of US\$1.3m and a maximum of US\$100m. Those are fairly typical averages across many sectors, with Rossum's US\$100m Series A round being a significant outlier.

This did not figure into my determining the scope of the present article, but the 119 companies surveyed here happen to represent a total investment of somewhere just north of US\$1 billion. That is a decent sum of money for new ideas in NLP. But – sobering thought – it is also only around what Elon made in a day from selling less than 1% of his Tesla shares.⁹

References

- Dale, R. (2019). Law and word order: NLP in legal tech. *Natural Language Engineering* 25, 211–217. doi: [10.1017/S1351324918000475](https://doi.org/10.1017/S1351324918000475).
- Dale, R. (2021). GPT-3: What's it good for? *Natural Language Engineering* 27, 113–118. doi: [10.1017/S1351324920000601](https://doi.org/10.1017/S1351324920000601).
- Dale, R. and Viethen, J. (2021). The automated writing assistance landscape in 2021. *Natural Language Engineering* 27, 511–518. doi: [10.1017/S1351324921000164](https://doi.org/10.1017/S1351324921000164).

⁹See <https://www.cnbc.com/2021/11/13/elon-musk-sells-another-1point2-billion-of-tesla-stock.html>.

Cite this article: Dale R (2022). \$NLP: How to spend a billion dollars. *Natural Language Engineering* 28, 125–136. <https://doi.org/10.1017/S1351324921000450>