

Speech recognition in professional use.

A new way of working: How
AI and cloud applications
are revolutionising the work
environment



Speech recognition in professional use.

A new way of working: How AI and cloud applications are revolutionising the work environment.

With rapid advances in speech recognition and analysis, voice input is being effectively used in more and more areas of application. Not only is it much faster than entering text on a keyboard, it can also provide employees with more flexibility and satisfaction. When choosing a speech recognition solution, companies should consider security and data privacy aspects as well as performance and accuracy.

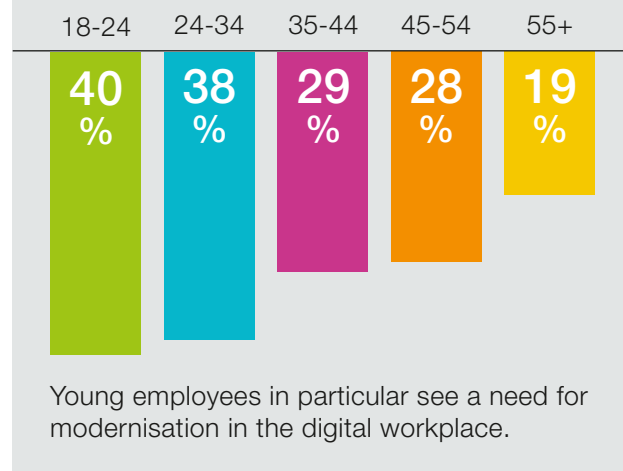
Digitisation is massively changing the way people work and their choice of working equipment. Mobile devices, high-performance networks and cloud-based infrastructures enable employees to access company resources from anywhere, at least in theory, because the reality often looks quite different. Only 16 percent of German employees can work flexibly from home or while on the go, according to a recent study by Initiative D21¹. The refusal or inability to allow their employees to work flexibly can become a real competitive disadvantage for companies.

Companies that meet the demands of today's employees (keyword: "A new way of working") have higher growth. According to a study commissioned by VMware², they are also more likely to be market leaders than competitors with more traditional workplace facilities. A survey³ by the market research company YouGov on behalf of the intranet agency Hirschtec shows that young employees in particular are pushing for changes. While 40 percent of 18 to 24-year-olds see a need to modernise the digital workplace, only 19 percent for over 55-year-olds feel the same.

When it comes to the office environment, bosses need to say goodbye to closely held opinions about work organisation and equipment. Open-space offices, which were once considered a haven for creativity and productivity, are increasingly being criticised. According to empirical studies by Ethan S. Bernstein and Stephen Turban, they inhibit interactions and communication rather than promoting them⁴.

The need to modernise the digital workplace

Percentage of people per age group who rate the need to modernise their digital workplace as high or very high.



Source: YouGov/Hirschtec
Sample: 1,000 working people in Germany

1 https://initiated21.de/app/uploads/2019/01/d21_index2018_2019.pdf

2 <https://www.vmware.com/radius/digital-employee-experience-infographic/>

3 <https://hirschtec.eu/studie-digitaler-arbeitsplatz-ganzheitlich-2019/>

4 The impact of the 'open' workspace on human collaboration, <https://doi.org/10.1098/rstb.2017.0239>

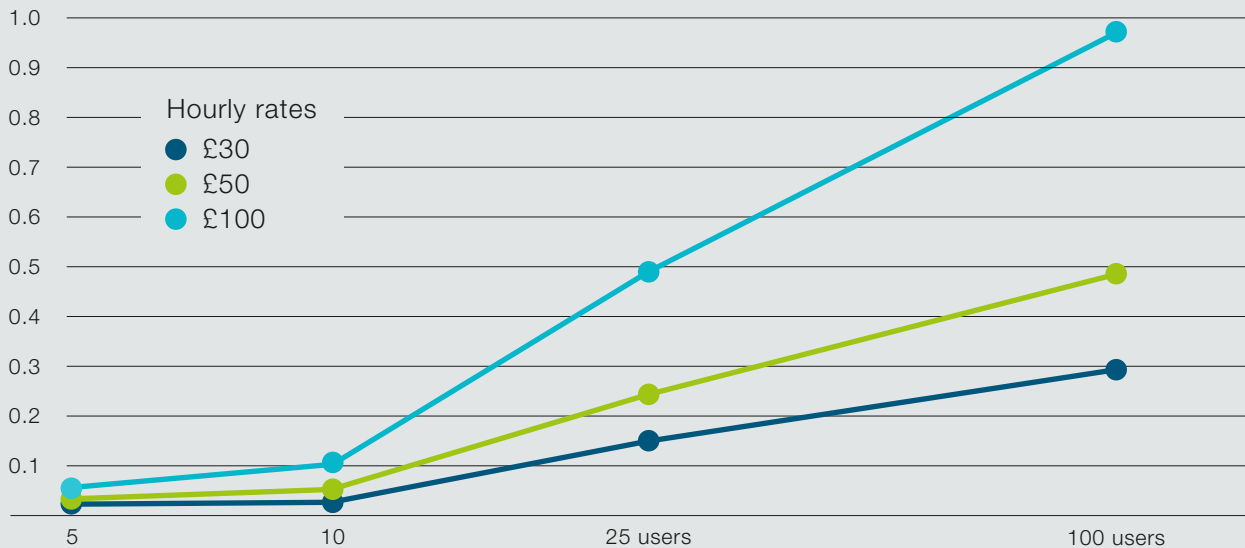
Even traditional tools such as the keyboard could become obsolete. Advances in artificial intelligence (AI) are making them increasingly superfluous. Current speech recognition systems can recognise and implement spoken text with almost no errors. According to a study at Stanford University, text input by voice is three times faster than typing⁵.

With voice input, text-based tasks such as office communication, documentation or drafting briefs can be considerably accelerated. As a result, the investment in a speech recognition solution quickly pays off. A company with 50 employees could potentially save 240,000 pounds per year if each of its employees saves just two hours of work per week through more efficient text entry (assuming hourly rate of 50 pounds).

The significant advances in machine learning (ML) and artificial intelligence (AI) are major reasons for the high performance levels of current speech recognition solutions. Algorithms can now learn independently from experience or can be specifically trained. By interacting with a speaker, they adapt increasingly better to their language characteristics, such as a dialect or accent, and continuously increase the accuracy of the match. Subject-specific vocabularies can be imported in advance, explicitly trained or continuously improved in the dictation process. AI-based components such as Natural Language Processing (NLP) and Natural Language Understanding (NLU) allow modern speech recognition programs to identify the content and context of the speech input. For example, they can respond to instructions such as “next page” or “new section” without the speaker having to use rigid word combinations for control commands.

Potential savings by using speech recognition

Savings in million pounds per year with 2 hours per week



If you gain two hours a week of extra time by using a speech recognition solution, you can save a lot of money. (Source: Nuance)

⁵ https://hci.stanford.edu/research/speech/paper/speech_paper.pdf

Cloud computing also promotes the spread and application of voice-based text creation. A centrally managed and flexibly usable infrastructure considerably simplifies access and can be scaled as desired with little effort. Instead of a complex installation at the customer's location, cloud-based voice recognition can be used quickly and easily with just a few clicks. Nuance is one of the few manufacturers to offer a private cloud solution that enables speech recognition in their own data centres for those organisations that cannot consider cloud computing on external servers. The provider or the internal IT administrator provides the necessary server and storage components. He also installs all necessary updates, including security updates, without users having to worry about it.

It is therefore not surprising that more and more companies are interested in voice-based text entry and want to follow the "A new way of working" concept as a new approach to working in business today. This is the result of a study conducted by market research company Censuswide on behalf of Nuance, which surveyed 400 IT decision makers in the financial and legal sectors. Almost 30 percent of the survey participants consider the keyboard to be unsuitable to facilitate the transformation into the digital business age. Almost everyone involved believed that they would benefit from more freedom at work and that speech recognition would be a valuable tool for their daily work. Nearly 90 percent were also convinced that speech recognition tools would reduce their dependence on transcription offices and other text service providers. More than 80 percent said they could use voice recognition to work more productively.

Beware of using common applications professionally.

Current operating systems for mobile devices and PCs such as Android, iOS, Windows 10 or Mac OS X already have their own voice assistants. So what could be more obvious than using these existing tools professionally?

Well, a closer look at the terms of use shows that unforeseeable dangers await. The recorded voice data is usually processed on servers of providers located in the US. If personal data is recorded and transmitted without the data subject's consent and without a contract with the provider for order processing of the data, this may violate the European General Data Protection Regulation (GDPR). In serious cases this can result in fines of up to four percent of a company's global gross annual sales⁶. This could mean that the savings made by using the free tools can be quickly offset.

People subject to professional secrecy, including doctors, lawyers and insurance companies, are particularly exposed to risks. Anyone who discloses third-party information as a member of such a group can be punished with a prison sentence of up to one year in accordance with Section 203, Paragraph 1 of the German Criminal Code (StGB). German lawyers are also obliged to secrecy by Section 43a of the Federal Lawyers' Ordinance (BRAO).

But even if there are no legal concerns against the use of voice recognition software from the end customer sector, companies face numerous disadvantages when using such assistants. There is no guaranteed availability or professional support for business users. Integration into existing software environments or workflows is not always possible. One of the weak points of consumer Apps is their inability to adapt to specialist requirements. Most industries have a specific vocabulary containing terms and abbreviations that common applications can't deal with. Professional speech recognition software, on the other hand, can be expanded with existing subject-specific language packages or can learn them very quickly.

⁶ <https://www.legislation.gov.uk/ukpga/2018/12/contents>

How financial service providers benefit from speech recognition.



The European Union significantly expanded the documentation requirements for banks and insurance companies in 2018 with the financial market directive “MiFID II” (Markets in Financial Instruments Directive)⁷. All communication with customers in connection with orders, such as the purchase of securities, must be documented and kept for five years. Customers may request that the records or copies thereof be made available to them. In the case of complaints or official investigations, these documents also provide important evidence. For the companies concerned by this, it means a considerable increase in personnel. Recorded phone calls must be transcribed and handwritten minutes must be typed up. Here, speech recognition can make a significant contribution to improved efficiency and productivity, but also to more legal certainty.

The example of NATIONAL Bank in Essen, Germany shows how financial service providers benefit from professional speech recognition. Many of the approximately 600 employees of the traditional bank, founded in 1921, spend a large part of their working day writing credit statements, company analyses, minutes from meetings and other extensive text documents. By using Dragon Professional Anywhere (DPA) speech recognition software, the company was able to significantly increase the efficiency of its employees.

The software can transcribe 160 words per minute—a 3-fold increase compared to keyboard input. The error rate is very low. Around 99 percent of the spoken text is recognised correctly. But the employees’ health also benefited from using the speech recognition solution. Instead of typing in texts for eight hours or more while sitting in front of their keyboards, they can now organise their work flexibly and, for example, dictate while standing.

The property and management consultancy Courté J. & Co. KG also benefits from using professional speech recognition. Previously, business letters and briefs had been dictated and then transcribed by third party typing offices or the internal secretariat. This procedure was complex and slow. Days often passed between the drafting of the document and its completion. With the help of Dragon speech recognition, Courté can now create briefs directly during dictation. This has enormously accelerated the process.

⁷ https://www.bafin.de/DE/PublikationenDaten/Jahresbericht/Jahresbericht2017/Kapitel5/Kapitel5_1/Kapitel5_1_1/kapitel5_1_1_node.html

Speech recognition in **law firms** – so that clients can get their cases handled faster.



The increasing digitisation does not stop at the law firms. According to a representative study by the AdvoAssist lawyer network in cooperation with the litigation financier Foris⁸, 69 percent of the lawyers are convinced that the use of digital technologies will “strongly” or “very strongly” influence their law firm in the next five years. Classic lawyer activities such as advising clients, drafting contracts and preparing documentation are also coming under pressure from digital service providers, the so-called LegalTechs. These are mostly Internet portals.

While behind these portals there are often still “human” legal experts who use the Internet essentially as a new sales channel, there is an increasing threat of competition from machines. Systems based on artificial intelligence (AI) are also showing astonishing achievements in the field of law. In some areas, AI is already clearly superior to human lawyers, as an example from America shows. To test the performance of its AI-based analyses, the LawGeex⁹ portal tested its algorithm against 20 lawyers. They had to examine five non-disclosure agreements (NDAs) for legal risks. Lawyers took an average of 92 minutes to complete the task and found 85 percent of the problems. The AI, on the other hand, did the job in just 26 seconds, with an accuracy of 94 percent.

⁸ <https://www.advo-assist.de/rechtsmarktforschung/>

⁹ <https://www.lawgeex.com/resources/aivlawyer/>



In addition to the increasing competition from the LegalTechs, the lawyers are also worried about junior staff. Christoph Möllers, lawyer and member of the vocational training committee of the Federal Chamber of Lawyers (BRAK), warned during the event “Zukunft der Anwaltschaft” (Future of the Attorney’s Office), that the profession of legal assistant is threatened with extinction¹⁰. Career prospects are too unattractive; competition from the judiciary and the legal departments of companies is too great. As a result, lawyers increasingly rely written documents created by third party typing offices. In addition to the legal security and confidentiality problems, this cumbersome process also leads to long waits and high costs. These are considerable competitive disadvantages in light of the fast and cost-effective digital competition.

Speech recognition software is therefore an ideal tool for lawyers to become faster and more efficient and thus to be able to compete with their digital rivals. For example, Newstead and walker relies on Dragon voice recognition by Nuance. In the past, their lawyers used analogue or digital dictation machines to record their briefs that were then transcribed by a typist. This process took several hours. In the meantime, lawyers had to keep the facts in mind in order to be able to check the finished version for correctness. This is an enormous cognitive strain at times with three or more parallel cases. Today, their lawyers can dictate their text directly and check it immediately. Then they can complete the process and immediately and fully concentrate on the next task. With the enormous saving of time and costs, the lawyers can also devote more time to legal cases and react more quickly than before.

¹⁰ <https://www.brak.de/thesen-zur-zukunft-der-anwaltschaft/>

Digitisation in public authorities – an obstacle course.



In public administration, digital transformation is currently still particularly difficult. As the Tagesspiegel reported, referring to a study by the consulting firm Ernst & Young¹¹, making appointments online is only possible with one in five German public authorities. At the same time, fees can be paid online in only six percent of cases and only three percent use the functions of e-ID cards. A key obstacle to more digitisation is the personnel structure. There is no consistent alignment of the personnel strategy with the requirements of digitisation, according to the administrators interviewed.

Not only the HR strategy, but also the overall staffing is preventing a faster transformation. There is a lack of administrative staff, especially in large cities. The existing specialists are so involved in operational tasks such as the recording and documenting processes throughout the work day, that they neither have the time nor the energy to pursue strategic development.

The use of intelligent speech recognition solutions can make a significant contribution to relieving staff and creating the necessary freedom for further development. It starts with a call being answered in the contact centre. With the help of speech recognition and automated communication control¹², Nuance research shows that 54 percent of calls can be handled directly without human intervention.

The average processing time is reduced by 42 percent. In addition, routine tasks such as the creation of briefs and notices or communication by email can be significantly accelerated.

For example, several UK government bodies use Dragon speech recognition software. Field staff can dictate their reports on site or on the way to the office in a mobile dictation machine. When they arrive at their workplace, the finished information is already available. The employees can then get on with new tasks instead of having to write a log first. In addition, a promptly dictated text usually represents the facts much more accurately and correctly than a document written from memory an hour later.

Authorities also have a key role to play in integrating people with disabilities into the labour market. If physical restrictions make it difficult or even impossible to work using a keyboard and screen, speech recognition software can be of great help to make such employees' everyday work easier or to enable them to engage in professional life in the first place.

¹¹ <https://www.tagesspiegel.de/wirtschaft/digitalisierung-verwaltung-weiterhin-wenig-digitalisiert/23770040.html>

¹² <https://www.nuance.com/de-de/industries/government.html>

Conclusion.

Digital transformation poses great challenges for companies of all sizes and in all industries, but also for public administration. One of the biggest challenges lies in creating enough time and leeway for strategic direction and further development in daily business. Speech recognition can make a significant contribution here, because it speeds up the process of capturing and processing documents and messages. It is more ergonomic, so therefore it reduces work fatigue and it considerably simplifies text-based workflows. This gives employees more time and energy to train themselves, to develop new business and to continue with the digital transformation of their activities.

Speech recognition is also a step towards the future of work (“A new way of working”) with more freedom and autonomy for employees. However, companies and authorities should not neglect data privacy and security when choosing a solution. This is the only way to prevent an unpleasant surprise and to ensure long-term and sustainable use.

What to look for when choosing a speech recognition solution.

The following criteria are particularly important when deciding what to purchase:

Speed:

Your spoken text should be transcribed almost in real time. Speech recognition, which “ponders” after each sentence means you lose your train of thought and will delay the creation of the text rather than accelerate it.

Accuracy:

Speech recognition is of little help if the transcribed document is full of errors and has to be post-processed. Therefore, make sure it has a proven high recognition rate. Current market leading programs can correctly recognise up to 99 percent of dictated text.

Customisable:

The software should be quick and easy to adapt to the conditions and requirements of your environment. This ranges from training on individual speaker profiles and the technical vocabulary commonly used in your industry to integration into existing workflows, for example in document or approval management.

Flexible deployment:

Depending on requirements and budgets, deployment via a public or private cloud is possible. In some cases, direct installation of a local desktop solution can still make sense today. The provider should therefore offer you all the options and not restrict your freedom of choice with a small selection of deployment and licence models.

Easy to install and maintain:

Speech recognition software is of little use if the time and productivity advantages are negated by complex, time-consuming and labour-intensive administration. Therefore, make sure that the solution of your choice is easy to configure and install. In addition, automatic updates reduce the administrative workload.

Scalability:

You will only be prepared for the future if the speech recognition solution can be quickly and flexibly adapted to increasing needs without sacrificing performance.

Security:

During dictating, sensitive information is often passed on and processed. Make sure that the communication between the components of the solution is protected by a strong 256-bit encryption. Strong encryption should also be used when storing data.

Data privacy:

When processing personal data, European General Data Protection Regulation (GDPR) requirements must be observed. When choosing a cloud solution, make sure that the provider's data centres are located in the European judicial area and are certified according to common standards such as ISO 27001. The German Federal Office for Information Security (BSI) Cloud Computing (C5)¹³ requirements catalogue also provides a valid basis for assessment and decision-making.

Integration capacity:

The solution you choose should be installed locally as well as seamlessly integrated into existing thin client infrastructures, virtualised environments and workflows.

13 https://www.bsi.bund.de/DE/Themen/DigitaleGesellschaft/CloudComputing/Anforderungskatalog/Anforderungskatalog_node.html

Checklist for choosing a speech recognition solution.

- What deployment models are there?
- Do I have to install and manage the software myself or can I get it as a service?
- How many users can work with the software at the same time? What scaling options are there?
- How is the software structured? Does it have to be installed on every client or is there a central server component?
- Can the solution be installed on a local server without the cloud?
- How long does the installation take? How costly is maintenance?
- How accurate is speech recognition? Can the software adapt to individual speakers?
- Can the speech recognition dictate directly at my cursor in my case management systems?
- How quickly can the software learn subject-specific vocabularies?
- Does the software need to be trained? If so, how much is it?
- Can the software be integrated into existing workflows? How complex is the integration?
- Can the speech recognition software automate repetitive tasks by the creation of voice commands?
- How secure is the software? Are transmissions and storage encrypted?
- Are all data privacy issues observed? What certifications and evidence are there for this?
- If the software is provided as a service (SaaS) from a public cloud: What certifications can the cloud provider provide? Is the data encrypted during transmission and storage? Are the data centres in the European judicial area?

Glossary of Terms.

A new way of working

This expression describes the new way of connecting life and work. Due to the widespread availability of the Internet and cloud applications, physical presence is no longer required for many jobs. Employees no longer want to work in fixed structures. They prefer “agile” ones where they benefit from more freedom and self-determination. Hierarchies in companies are reduced and strategic decisions are made jointly.

Artificial Intelligence (AI)

AI refers to information processing systems that capture the meaning of language, texts or images and can draw conclusions from them. AI systems know the principle of cause and effect, recognise concepts and context, and can make hypotheses for the future based on their observation. They are also able to test these hypotheses on the basis of new data and to learn from the results. AI systems communicate with people in a natural way via speech or gestures and can react appropriately to the actions and instructions of their counterpart.

Cloud Computing

According to the definition¹⁴ of the National Institute of Standards and Technology (NIST), cloud computing has five characteristics:

On Demand Self Service: Users can access resources directly at any time.

Broad Network Access: Access is via a network infrastructure. Standardised protocols are used, which enable access via all common client platforms (PC, smartphone, thin clients).

Resource Pooling: Resources are shared and are available to users dynamically depending on requirements.

Rapid Elasticity: The allocation of resources is quick and flexible. Almost any amount of performance is available to users at any time. Scaling is possible at any time. Scaling effects can be exploited efficiently.

Measured Service: The cloud operator monitors the systems automatically and continuously, thus ensuring service quality.

Cloud Deployment Models

Private Cloud/On Premise: With this deployment model, the cloud infrastructure is available to a company or department for its exclusive use. It can be located in the company’s own data centre, but can also be operated by a provider.

Public Cloud: Deployment model in which the cloud infrastructure is accessible via the public Internet and can be used by everyone. User accounts are logically separated and additionally protected against unauthorised access through encryption during transmission and during storage. Users have no influence on which data centre, in the often worldwide infrastructure, will process their data. However, data processing can often be restricted to certain zones, such as the EU.

Cloud service models

Software-as-a-Service (SaaS): Service model in which the provider provides applications from their cloud environment. The user usually pays a monthly or annual fee and can then use the solution directly without having to worry about updates, protection or scaling.

Platform-as-a-Service (PaaS): The cloud operator provides a platform with the operating system, middleware and development tools on which a customer can develop and/or operate their own applications.

Infrastructure-as-a-Service (IaaS): In this case, the customer accesses (mostly virtualised) hardware components such as servers, storage and RAM and uses them to put together their own drive environment. Installation and management of the operating system and applications are the responsibility of the user.

¹⁴ <https://csrc.nist.gov/publications/detail/sp/800-145/final>

Machine Learning (ML)

Machine learning algorithms can analyse large amounts of data, recognise patterns and derive rules for classification from them. There are three main types of machine learning: In “unsupervised learning”, the algorithm evaluates data without human intervention, while in “supervised learning”, a human “trainer” specifies the training data and evaluates the results. “Reinforcement learning” is a middle ground. The algorithm learns independently, but its results are measured against target values set by humans. By repeatedly trying out and changing the parameters, the system tries to get closer to these target values and thus optimise its learning function.

Natural Language Processing (NLP)

A method in which spoken language or text is examined for key words in order to recognise the meaning and to be able to formulate a suitable answer. For example, a Chatbot could use the word “weather” to display the weather forecast for the next few days or the keyword “costs” to indicate the price of a product.

Natural Language Understanding (NLU)

A semantic analysis of a spoken or written text, which makes it possible to recognise the subject of a document and to react appropriately to the context. This allows moods and evaluations to be recorded and intentions to be distinguished. For example, would the interlocutor like to ask a question, order something or complain.

Speech-to-Text/Text-to-Speech

AI services that translate spoken language into text and vice versa.

Thin Client Infrastructure

Distributed infrastructure in which the main functions are located in the central server component. The client only requires a browser through which the user can access the server components. The thin client approach enables mostly hardware-independent work, since all essential steps are carried out on the server. It also makes it easier to deploy applications and reduces maintenance.

Nuance Communications (NASDAQ: NUAN) is the pioneer and leader in conversational AI innovations that bring intelligence to everyday work and life. The company delivers solutions that understand, analyse, and respond to people – amplifying human intelligence to increase productivity and security. With decades of domain and AI expertise, Nuance works with thousands of organisations globally across healthcare, financial services, telecommunications, government and retail – to empower a smarter, more connected world.



For more information, please visit
www.nuance.co.uk/dragon,
LinkedIn or Facebook

For specific information, you can also contact our Nuance Premier Partners to ensure you will get the right advice and professional services for a successful deployment.

Nuance Communications UK Limited

33 Soho Square
London W1D 3QU England