



VOICE SEARCH SURVIVAL GUIDE



CONTENTS

| | |
|--|----|
| INTRODUCTION | 3 |
| WHAT IS VOICE SEARCH? | 3 |
| A BRIEF HISTORY OF VOICE SEARCH AND VOICE RECOGNITION | 4 |
| WHY BRANDS SHOULD TAKE VOICE SERIOUSLY | 9 |
| OPTIMISING FOR VOICE SEARCH | 10 |
| STRUCTURED DATA | 10 |
| NATURAL LANGUAGE SEARCH | 15 |
| LONG TAIL KEYWORD TARGETING | 16 |
| INFORMATIONAL KEYWORD TARGETING | 16 |
| CONCLUSION | 17 |

INTRODUCTION

With the content surrounding the voice search phenomenon now reaching ear-splitting levels, we thought the time was right to offer some rational advice on voice search. We have made our feelings known a couple of times in various blogs but, to restate it here – we (along with many of the more level headed of the industry's practitioners), while not wishing to downplay the inevitable importance of voice search, believe that much of the industry has either overreacted or failed to act and that the correct course of action lies somewhere between.

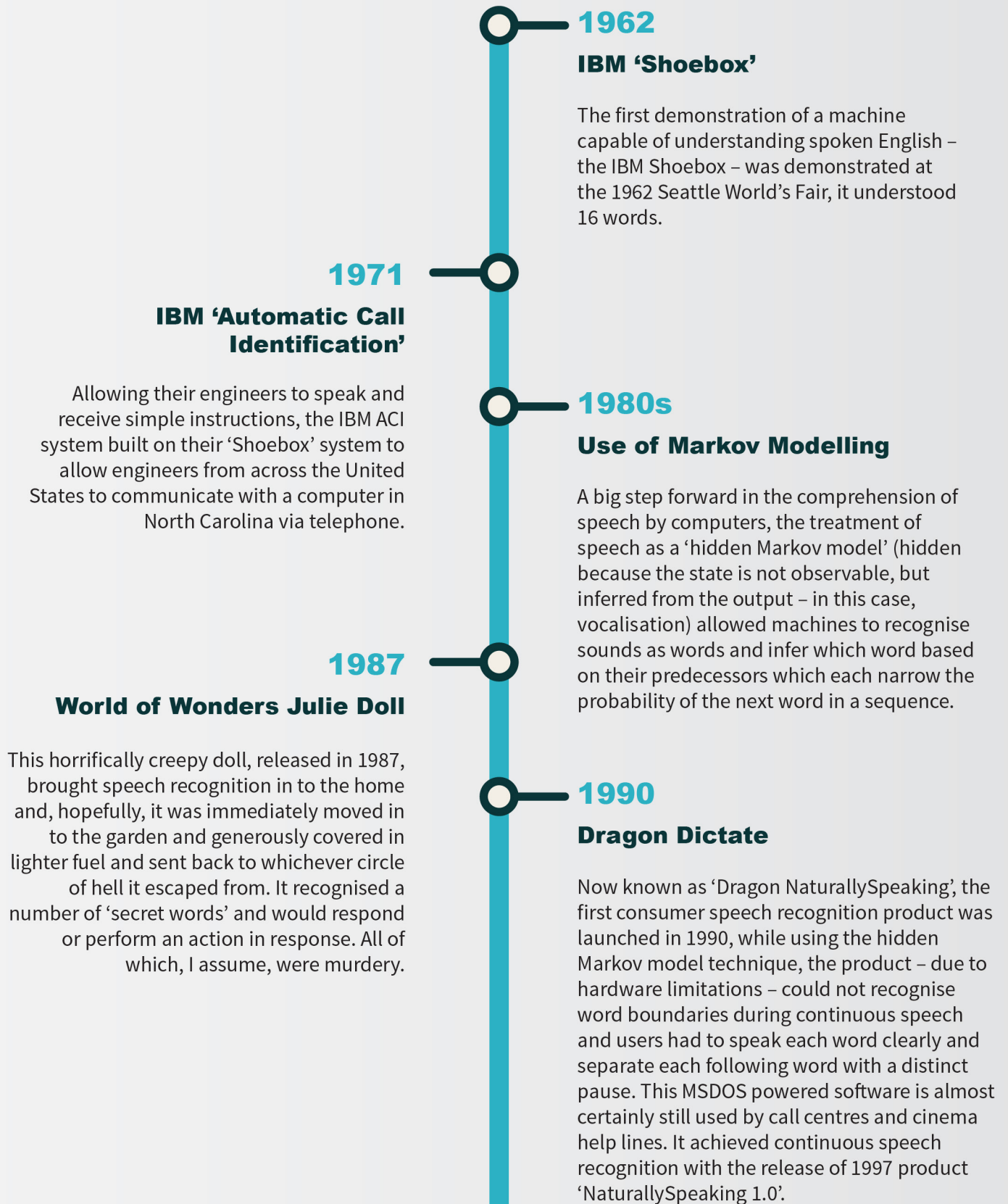
WHAT IS VOICE SEARCH?

Voice search, or voice enabled search, is a method of searching the internet through an internet connected device using spoken commands rather than the traditional method of inputting text in to a search field. The searches, which are – according to various sources, and implied in advice from Google (including [‘How to prepare for the voice revolution’](#) which recommends tailoring content to cater to long tail, conversational queries) are typically longer than text search (though the often quoted 4.2 to 3.2 voice vs. text is from [research on Yahoo data](#)) meaning that voice search is conversational, long form, spoken interaction with a distributed database (such as the internet) through a search engine or digital assistant.



A Brief History Of Voice Search And Voice Recognition

Unlike certain aspects of search – which have been driven predominately by the search engine giants – the rise of voice search is a mix of user demand, science and technology and search engine impetus driving development together, but where has it all come from?



1993

Speakable Items

Allowing their engineers to speak and receive simple instructions, the IBM ACI system built on their 'Shoebox' system to allow engineers from across the United States to communicate with a computer in North Carolina via telephone.



1996

IBM MedSpeak

The first commercial continuous speech recognition software with a significant vocabulary, MedSpeak (later known as ViaVoice) would be IBM's last entry in to the commercial speech recognition market, and ViaVoice was licensed exclusively to Nuance Communications (makers of NaturallySpeaking). Instead, IBM focused on embedding ViaVoice in to telephony systems.

2006

Why We Can't Have Nice Things

Worldwide spy agency and Edward Snowden's home away from home before the Russian Federation, the NSA begin using speech recognition to identify keywords in recorded conversations.

2007/8

2007/8 – GOOG-411 and a Voice Search for iPhone

2007 saw Google release both a voice search telephone directory and a voice search app for the iPhone, making it the first appearance of voice search on a mobile device.

1993

Sphinx-II

With the largest vocabulary of any contemporary continuous speech recognition system, the Sphinx-II was invented by Xuedong Huang, a Carnegie Mellon University graduate that would go on to lead Microsoft's speech and language initiatives and is presently responsible for the first name that will likely be familiar to all readers – Cortana.

2002

Microsoft Office

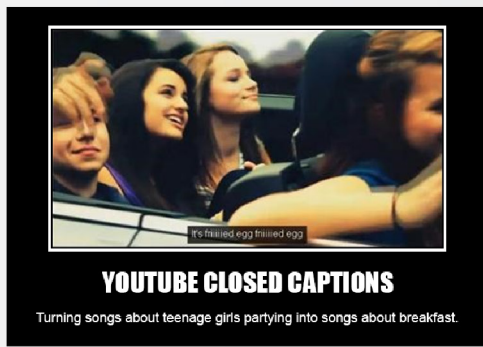
For the first time, 2002 saw the integration of speech recognition in to its suite of Microsoft Office programs.



2007

Windows Vista

While almost universally despised, and for good reason, Vista was the first version of Microsoft's Windows to incorporate speech recognition.



2010

YouTube Close Captioning

In 2010, Google began using voice recognition to add closed captioning (subtitles) to YouTube videos and, while these are frequently ridiculous, they are mostly accurate and have rendered the vast majority of YouTube's content accessible to those with hearing difficulty.

2011

Siri

The first and consistently worst personal digital assistant arrives on the iPhone, causing a plague of comic sketches featuring hilariously misunderstood sentences to spring up globally. While Siri had previously been available as an iOS app; its purchase by Apple and integration with the Apple OS catapulted it into the limelight.

2011

Introducing Voice Search for Your Computer

In a YouTube video among other places, Google announced the introduction of voice search for your computer – like Google Voice, but for your desktop

2013

Google Hummingbird

In what then Senior Vice-President of Google Amit Singhal called the most dramatic change to the Google algorithm since 2001, the Hummingbird update sought to shift interactions with search, and rankings, to a more human focus – with greater emphasis on natural language, so a page matching the meaning of a search could outperform that which only matched a key term or phrase.

2014

Cortana

Probably the best of its time, and certainly criminally under-deployed, the Cortana digital assistant arrived in 2014. Sadly, as a Microsoft product and tied to Microsoft devices, its abilities are unlikely to be celebrated, but Cortana regularly appeared either first or second in accuracy and ability until it was outpaced by its rivals a few years later.

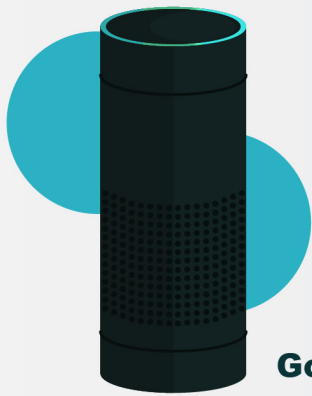
2014

Alexa

While it took a few years to catch on, the Amazon Echo, the first major smart speaker, appeared late in 2014, bringing Alexa – and probably GCHQ and the CIA – into the homes of consumers across the UK and US.

While almost universally despised, and for good reason, Vista was the first version of Microsoft's Windows to incorporate speech recognition.





2016

Google Assistant

While at the forefront of voice search for smart devices – with GOOG-411 and the voice search app for iPhone, it took until 2016 for the Google Assistant smart device to appear, the Google Assistant was integrated in to all android devices, while the smart speaker went on to sell almost 7 million units over the 2017 Christmas period. Google Assistant, despite having a significant gap to close, managed to reach 29.9% in 2018.

2015

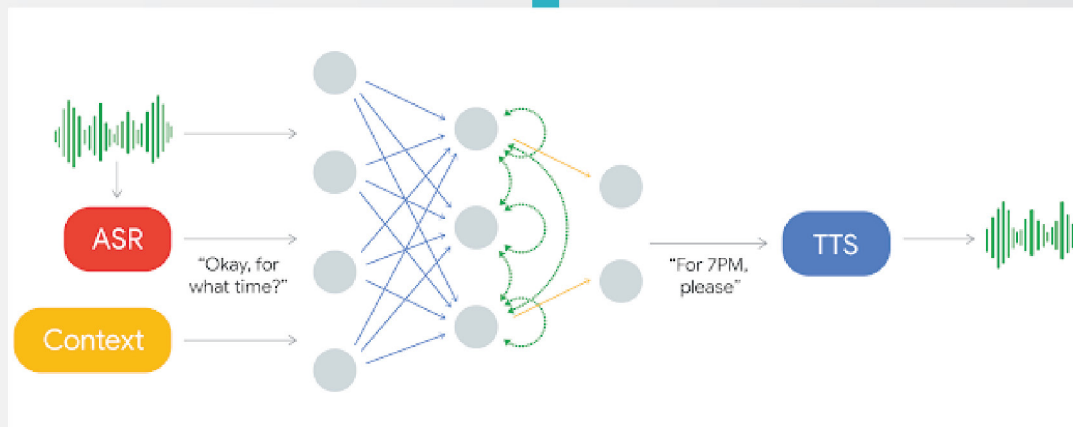
RankBrain

RankBrain emerged in 2015 and has been omnipresent in digital marketing circles ever since. With 15% of all Google searches having never previously been entered, the RankBrain machine learning technology allows Google's search algorithm to return relevant results even on a previously unseen search. As with previous updates that expand the search engine's comprehension, RankBrain, or a subsequent iteration, will be important to the future of voice search.

2017

Voice Search Reaches 25%

Voice search reached 25% of all searches in 2017 and while there is yet to be any real evidence that this can be converted in to a major sales platform for brands, there was significant doubt that mobile eCommerce would take off when the iPhone debuted in 2007.



2018

Google Duplex

In a demonstration that caused audible gasps, at I/O 2018, the CEO of Google, Sundar Pichai, demonstrated the latest in voice recognition and response in their Google Duplex technology which can be heard interacting almost seamlessly but fur a slight oddness of affect with an unsuspecting series of people which whom it books various appointments and reservations.

2018

Cortana and Alexa Begin Working Together

Not, as it may sound (hopefully), the beginning of Skynet, 2018 saw a demo of Alexa enabled devices integrating with Cortana – allowing the Amazon digital assistant access to Cortana's list of skills and accesses (including their PC and MS Office) and vice versa.

2019 and Beyond

The next big step in voice search is overcoming the crowd – it's been fairly well established that, given the opportunity, people will use voice search, but overcoming the problems inherent in speaking to a personal digital assistant in a crowd, on public transport etcetera is a key barrier to the next step.

There are two competing techniques which seem to be the most probable solutions – subvocalisation and ingressive speech recognition. While both remove the embarrassment factor of speaking out loud to reorder toilet paper on the train, they each have their problems. The former, at the moment, requires an odd looking neck brace and the latter requires a device that has to be held almost to the lips.

While both have proven prototypes, success and failure will be determined by the device that goes to market – as we've seen with Google Glass and other wearables, the clunky and faddish nature of the devices has drastically held back their adoption. If, however, the same excitement can be generated for one or other style of device, then that will almost certainly be the input method that wins the race.

From there on, the process will be one of integration – the ability for phones to build voice profiles, allowing you to make phone calls in your own voice without speaking out loud (watch out board meetings, you're about to become even less useful), the ability to link with fitness apps, headphones, glasses and other wearables: 'Okay Google, what's my heartrate? Reroute my run to add an extra mile and play my running playlist'.

While it is difficult to predict outcomes for future technology, it is possible to identify problems and predict solutions and, if a sleek set of wearables hits the market with subvocal recognition capacity that can integrate with popular apps, then there will quickly be a change in the world that we haven't seen since the mobile phone hit mass adoption in the late 1990s.



WHY BRANDS SHOULD TAKE VOICE SERIOUSLY

With the previous section out of the way, we're hoping that much of the convincing is done. It's easy to see the failure of voice to reach maturity as early as people consistently persist in predicting, however, as indicative of voice being a kind of vapourware. However, while brands were investing in Telegraph advertorials in January to tell that this year, like the perennially optimistic Liverpool fan, will be the year, the big leap forward has not arrived and won't this year – and maybe not even next year.

What will and is happening, however, is that exposure to and trust of digital assistants is growing all the time (even at a time when security issues around technology are a point of contention and frequently make international news, sales of smart speakers are still escalating quickly). With this increased coverage and developing trust, the commerce value is growing.

At the moment, predictions are as inconsistent as they are wildly variable as to the present and future value of voice commerce, however it seems reasonable from research to place last year's totals <1% of total retail spend (with one source from the end of 2018 estimating a value in the US of approximately \$2 Bn).

While this is not a tremendous amount as a share of a trillion dollar market, there is sufficient reason to expect this to increase and with the number of smart speakers growing seemingly exponentially, there are fewer reasons to completely dismiss claims that a \$10 spend per device could lead to an \$80 Bn spend in the next few years and for it to grow further from there.

The investment from major tech companies is a sure fire indication that voice should be taken seriously – Amazon, Facebook, Microsoft, Apple, Google and more are all investing significantly in voice and while it may be a new player that eventually captures the voice search market (or an old one if IBM's progress is eventually directed that way), there is little room to doubt that the major tech businesses see this as a wise investment – especially when coupled, as it most often is, with AI and machine learning.

With these factors in mind, it makes sense to prepare – in the ways we can – for an emergent trend that could change the way we live and buy.

OPTIMISING FOR VOICE SEARCH

As previously stated, we can't prepare fully for a future that is – if not unclear, then certainly a little murky. However, by extending current trends, we can infer that there are a few areas that brands can invest in that will be important and achievable when it comes to preparing for voice search in whatever form it finally arrives while also making sense immediately. These are:

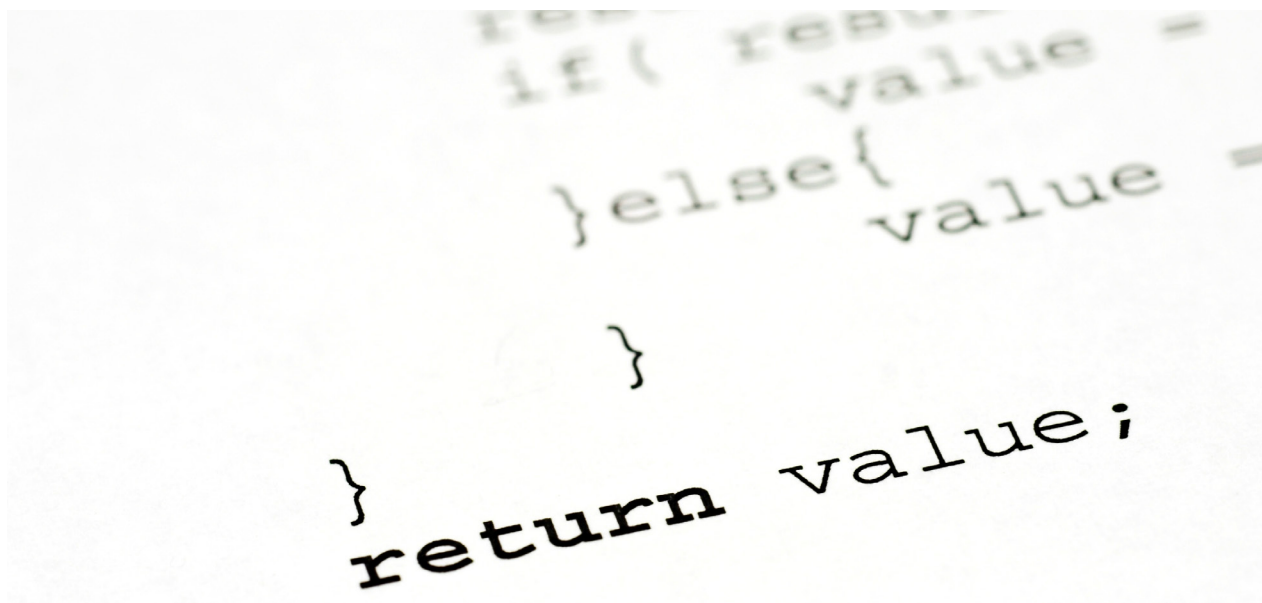
- Structured data
- Conversational or natural language
- Long tail keyword targeting
- Informational keyword targeting

STRUCTURED DATA

What is structured data?

Rather than being a specific subset or variety of data, when we refer to 'structured data' we are talking about an organisational construct of data. While prose may convey information, it tends to do so in an organic way - conveying information in a looser, more conversational manner. However, if you were to study the prose, and distil its meaning in to, for example, a table – this would represent 'structured' data, essentially the same information in an easier to digest format.

Where search engines are concerned, it is easier for an algorithm to parse information if it is offered within a scaffold or framework of structural information. This tends to be done using HTML, microdata and JSON-LD cues that provide the search engine with additional pointers that it can use to determine the nature of the data it is processing.



Why it matters

Along with schema.org, Google has been guiding the way we structure data on the web for years – with structured data recommendations for everything from addresses to eCommerce products; and while ‘write for people, not robots’ has been the most frequent advice for copywriters (since the Panda update), but that is no longer the best advice – or not entirely.

When the Panda update was rolled out in 2011, Google was looking to tackle keyword stuffing, scraped and thin content etc. and while Panda has been incorporated in to the core algorithm, there has been a blossoming of technology that has somewhat muddled the waters of its original task.

As voice search becomes conversational, and seeps into everyday life (as it will do), and Google answers more queries directly in search engine results pages (SERPs), page one – and more specifically, the top of page one – is going to be vital for certain query types.

The more Google’s machine learning technology and, consequently, the Google algorithm are able to understand, the more they are going to render visits to pages irrelevant (for informational queries, at least, though there are plenty of people arguing that it may be the same for transactions).

Schema markup is going to be the foundation of voice first search, but it is – as can be seen from recent developments and various Google patents – also directly involved in another of the trends that have presented themselves over the last few years, that of Google’s efforts to become an ‘answers engine’.

By using schema markup to make an easily digestible scattered database of the web, constructing ever increasing and comprehensive ontologies, Google is not only ensuring that it will be able to answer most queries, but also that many website visits will be avoidable (furthering the work already done by in SERPs answers and rich results).

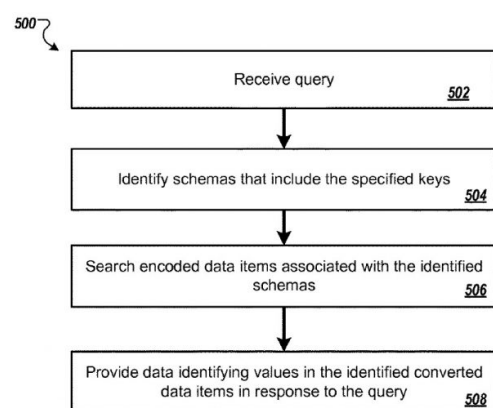


FIG. 5

It's no surprise, therefore, to see Google and Bing double down on the use of existing schema and in the creation of extensions. Whether or not spoken queries lead to the necessity to use a screen as part of the follow up, or a visit to a site, Google as an answer engine is going to fundamentally alter the main function of many a brand website.

With all this in mind, there will be a need to upskill staff or hire more JavaScript knowledge for in-house and agency teams to ensure that the correct and latest markup is being used – as well as to participate in the construction of industry specific extensions in the hope of being at the forefront of the coming changes, rather than having to react to every new extension.

The search and digital marketing industry therefore needs to meet machine learning in the middle, and invest more time and money in learning to speak the language of the algorithm as we wait for it to master ours.

Why and how you should implement speakable (and connected) markup

It should come as no surprise that Google and schema.org, as voice search becomes a more pressing issue, have announced that they will be supporting 'Speakable' markup (beginning in the US before roll out to the rest of the world).

What is speakable markup?

Speakable markup is a form of structured data that highlights sections of an article as useful for text-to-speech (TTS), using an id value reference and either CSS selectors (like a class attribute – commonly denoted by the hash symbol, like those used when [giving same page links](#)) or XPath's.

Why should you use speakable markup?

The speakable beta is only available to US users and only to those sites registered as news publishers at the moment (and works, as the below markup on Search Engine Land shows), so if it's not available to the masses, why implement it now?

```
in Google Maps that the results can fluctuate a lot when new listings show up. In any event, have you seen any local search changes around the same time of this Wednesday algorithm discussion at Twitter and Local Search Forums.", "author": {"@type": "Person", "name": "Barry Schwartz"}, "publisher": {"@type": "Organization", "name": "Search Engine Roundtable", "logo": {"@type": "ImageObject", "url": "https://www.seroundtable.com/images/publisher_logo.png", "width": 511, "height": 60}}, "speakable": [{"@type": "SpeakableSpecification", "xpath": "V/V/div[@id='maincontent']/V/V/div[@class='title left']/V/V/h1", "V/html/head/meta[@name='description']/V/@content"}]]</script>
<link href="https://plus.google.com/u/0/112116680648779331229/posts" rel="author"/>
```


It should come as no surprise that much of your future audience will be controlled by machine learning algorithms that will parse your site for communication via either rich results or by your choice of digital assistant.

While that doesn't exactly explain why you should implement the markup, the truth is that voice search and the age of the digital assistant is on its way, and while (unless Amazon manages to sort it out) it is unlikely to impact on your sales figures for a while yet, it will begin to impact your traffic in the next few years – especially if you offer any kind of informational content – which virtually all brands do to build authority and audience.

It's for this reason that you need to implement Speakable markup (and other applicable information markup schemas – like [HowTo](#), or even types of location markup) – a lot of content on a brand's site is there to build authority or address specific user queries and to, therefore, build an association between the brand and the product or service they supply. The available real-estate for this kind of content rapidly and vastly decreases in the age of the personal assistant.

With only one spoken result at present and with it extremely unlikely to exceed three – simply because of the time and demand on attention it would require from the user. If you want your brand to continue to build that recognition as an expert in the field, you're going to need to be ready as soon as the markup is rolled out to the general population.

How to implement structured data

Schema is available for free on schema.org (and speakable specific schema has been in various states of construction for well over a year), so can be implemented by any party wanting to get a handle on structured data.

Below, you can see the implementation you would need to replicate (minus the “TYPES...JSON:” section) if you wanted to use ‘speakable’:

```
{
  "@context": "http://schema.org/",
  "@type": "WebPage",
  "name": "Jane Doe's homepage",
  "speakable": {
    {
      "@type": "SpeakableSpecification",
      "cssSelector": ["headline", "summary"]
    },
    "url": "http://www.janedoe.com"
  }
}
TYPES: speakable
PRE-MARKUP:
An example of 'speakable' markup (JSON-LD only, initially).
MICRODATA:
TODO
RDFA:
TODO
JSON:
{
  "@context": "http://schema.org/",
  "@type": "WebPage",
  "name": "Jane Doe's homepage",
  "speakable": [ "#myhead1", "#thesummary" ],
  "url": "http://www.janedoe.com"
}
```

This is placed in the header section of your page's html (or in the site html if you can work out a way to standardise the naming of schema sections site-wide).

“@context”: “http://schema.org/”,

Provides Google with the information it needs to contextualise the code as relating to schema markup.

“@type”: “SpeakableSpecification”,

Gives Google a point of reference for your schema (allowing it to know what to expect and what it’s required to do with it for each of the schema types).

“cssSelector”

Tells Google where to look on the page (you can also use the xpath property for this purpose), this selector must then be added to the appropriate text inline – so, were we to use the above example, the code would read:

```
1 ▼ <style type="text/css">
2 ▼ /*markup begins
3
4   "@context": "http://schema.org/", ---- refers Google to schema as the source of the markup
5   "@type": "WebPage", ---- classifies the page type
6   "name": "Jane Doe's homepage", ---- names the page
7   "speakable": ---- classifies the information as speakable
8 ▼ {
9   "@type": "SpeakableSpecification", ---- refers Google to a specific type of schema
10  "cssSelector": ["headline", "summary"] ---- tells Google which sections have been nominated as speakable
11 }
12  "url": "http://www.janedoe.com"
13
14  markup ends*/
15 </style>
16
17 <h2 id="headline">This would render your title speakable</h2>
18 <p id="summary">And this would do the same for the summary paragraph following.</p>
19
```

There is nothing stopping even the smallest brands implementing structured data through markup in an effort to optimise for modern search and in anticipation of the next step.

And another thing...

As machine learning and various digital assistants grow in importance in search, the industry is going to need to begin or increase the frequency at which it cross skills its content creators.

Increasingly – and for much longer than brands can afford to wait – the age of assistance (as Google employees have begun to consistently refer to it) will require us to communicate our needs, wants and intentions as content creators to the various bots and crawlers that will deliver our content to the end user.

For that reason, brands need to ensure they’re riding the markup train rather than waiting on the tracks to be hit by developments.

For more information on structured data, you can download our beginner’s guide [here](#).

NATURAL LANGUAGE SEARCH

Speak to your consumer online as you would speak to them in store or over a desk, in a meeting room or over a coffee. While I wouldn't state that the death of the keyword has arrived – or is even imminent – it is no longer necessary for brands to pitch for a keyword in the heavy handed manner that the search and digital marketing industry has done until relatively recently.

The ability of search engines – all search engines, not just Google – to parse natural language and infer meaning from text has advanced some distance in the years since keyword densities were regularly measured and monitored.

What is natural language search?

Natural language search bridges human and computer approach to search – using machine learning and algorithm updates such as the previously mentioned Hummingbird, to address more complex searches and search intent. This type of search can be in text or speech, but with the intention, by search engines, being to return relevant, quality results for longer searches by analysing the whole query, rather than selecting specific keywords, and matching it to a similarly detailed analysis of appropriate webpages.

How to cater for natural language search

This is predominately done using the following two methods, but the key thing to remember on a page by page basis is that context is incredibly important. While keywords are still important, it is also important to consider context clues – so that the rest of the information on a page reinforces the page's relevance to the keyword target.

How natural language applies to voice

While the way people input searches has changed a lot in the last couple of decades, natural language is one of the biggest shifts – and it owes a lot to the way we speak. We expect search engines to understand us to an extent unthinkable in the late 1990s or early 2000s and, as the search engines have almost managed to do this, we have begun using longer queries – expecting even smarter responses – and voice search has consistently (as mentioned earlier) averaged longer queries than ever.

LONG TAIL KEYWORD TARGETING

While historically, SEO has been a campaign to pitch for your trophy keywords – the 1 or word search queries that represent the peak search volume and intent within your industry. However, as the search terms lengthen, so do the variations in searches. This means that you can look to rank for a series of long tail keywords – many of which will be less competitive – that will earn you similar levels of traffic to the trophy term, but will better cater to searcher intent and, as mentioned, when voice begins to increase in prevalence and with it the longer queries, will see improving volume over time.

INFORMATIONAL KEYWORD TARGETING

This ties in with a lot of the information mentioned previously – with Google seeming to head toward its goal of ‘answers engine’ quicker each year, the 10 blue links model of search is likely to take a fairly big hit (and this can be seen in widely reported dropping CTRs for organic search results). As such, it will be increasingly important to use high quality informational content to build your brand through association. This is especially true in voice – where there is only one result at present, and where there are likely to be no more than three in future just due to time and attention constraints.

For this reason, attempting to hit the rich result for industry relevant search queries will be an important part of voice search optimisation. This will include the production of guides and how tos – content which will help users and help to build a strong connection between your brand and the products and services that are important to your bottom line.

These queries should be viewed as exercises in brand building – while you may want to get as much traffic as you can to your site, these pieces of content should serve mainly to build that association; so that Joe’s Pushbikes – thanks to its helpful ‘best ride’ guides and cycle maintenance blogs – is able to build a strong connection between the searcher and the product, improving the likelihood of the user specifically seeking out Joe for future bike and accessory needs.

CONCLUSION

While the voice search revolution may be some way off, the need for brands to prepare for voice – as many spectacularly failed to do for mobile – is imperative. By being the first or best in your industry early on in competing for voice search queries and technologies, your brand can steal a yard in what is an increasingly competitive online space.

By understanding the increasingly rapid advance of voice recognition and the growing popularity of voice search and by preparing for the future by implementing these techniques, it is possible for brands to avoid a similar level of panic as the one they expressed prior to Google's Mobilegeddon algorithm update.

Just as Google optimised their SERPs for mobile, they will do so for voice, and preparing now is an investment in the future.



GET IN TOUCH WITH US TODAY

ABOUT US

Click Consult is a multi award-winning digital and search marketing agency with a focus on organic (SEO) and paid search (PPC). Part of global consumer brand business Ceuta Group, we have a team of 70-plus specialists and a portfolio of more than 60 clients worldwide.

Our complementary services include content marketing, outreach, social media, conversion rate optimisation (CRO) and international/multilingual search marketing. We can also offer training and consultation to support your teams or existing strategy.

Click was named Search Agency of the Year and Digital Agency of the Year 2018, adding to our long list of other awards and accolades, and also ranks within Econsultancy's 'Top 100 Digital Agencies', and Prolific North's 'Top 50 Digital Agencies'. We're also a Google Premier Partner, a Bing Select Partner and feature in The Drum Recommends.

YOUR BESPOKE ORGANIC SEARCH (SEO) ANALYSIS AND REVIEW

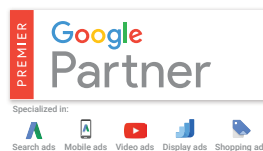
REQUEST NOW



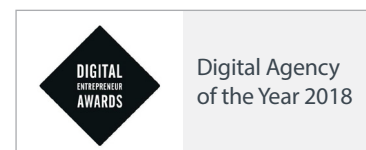
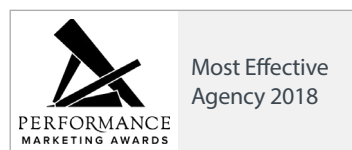
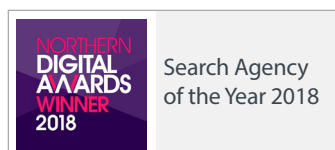
Follow us on [Google+](#), [Facebook](#), [Twitter](#) or [LinkedIn](#), or to find out what Click Consult can do for you, call:

0845 205 0292

OUR ACCREDITATIONS



OUR AWARDS



OUR TECHNOLOGIES

monitor **TRAX** rank **TRAX** feed **TRAX** link **TRAX** pro **TRAX** page **TRAX**