How the Autocomplete Function Influences Our Opinion Making and whether this is Unlawful from A German Law Perspective

Paulina Schindler1* and Lennard Lehmann2
1Friedrich Schiller University Jena, Chair of Business Informatics, Germany
2Friedrich Schiller University Jena, Chair of Public Law, esp. public commercial law, Germany
*Corresponding author

Abstract.

The autocomplete function in search engines can make the search process easier for users. However, this function can also influence the opinion formation of an individual and thus also of society. The technical implementation possibilities and possible limitations of the reliability of these search suggestions are presented in this paper. Subsequently, effects are presented which can have an impact on the user in connection with the autocomplete function of search engines. The possible individual and societal effects of autocomplete show that this function is suitable for influencing users. Due to its manipulability, this can also be done by third parties. A subsequent assessment of German law shows that this is not inadmissible. In order to nevertheless reduce influence over users and ensure diversity of opinion, increased transparency and supplementary regulatory provisions could be considered.

Keywords: autocomplete, media influence, diversity of opinion, regulation, search engine

1. Introduction

The Internet is a mass medium (Burkart, 2019). Since search engines play a central role in making the Internet accessible (Schweiger, 2004), they are of great social importance as intermediaries (Paal, 2015). Search engine operators can be classified as so-called media intermediaries (Kühling, 2021). In contrast to other intermediaries, they usually do not present their own editorially designed content, but only the content of third parties after an interaction of the user, in this case the input of a search query. Search engines can thus also be described as information intermediaries (Dörr & Natt, 2014), acting as intermediaries between users and third-party content.

Search engines are used by a large number of people. More than half of Germans stated in a survey that they use search engines at least several times a week (VuMA, 2021). Thus, numerous people are also regularly confronted with search suggestions that can be offered as part of a search process. So-called "search suggestions" are also referred to as "autocomplete" (Lewandowski & Quirmbach, 2013). Autocomplete is a variant of a search function in programs, whereby possible extensions are automatically suggested during the entry of a search term. In principle, they occur not only in classic text-based search engines that index web pages on the Internet, but are also used for other purposes, such as for online stores or online
directories. However, since these are reduced to a specific offering within a single web offering, they have less relevance to society as a whole and are therefore not considered in greater detail below.

The market-leading search engine Google (StatCounter, 2021) itself states that autocomplete reduces the necessary entries by about 25% and thus saves a total of more than 200 years of writing time per day (Sullivan, 2018). Therefore, the goal is to optimize the search process. Since autocompletion specifies possible search areas that can be used directly and arouse interest, the user's search can be guided by it (Lewandowski & Quirmbach, 2013). A not only positive influence on the opinion formation of a user and in the long term of society is thus conceivable. An example of this is the offering of search suggestions that can have a damaging effect on a person’s reputation - this has already led to legal disputes worldwide (Karapapa & Borghi, 2015).

Current examples of the autocomplete function in search engines in Germany can be seen in Figures 1 and 2. While the search for "China" directly refers to the corona pandemic and the location of the searcher, the search for "Asylanten" ("asylum seekers") leads to extensions with partly negative connotations such as "politikversagen" ("political failure") or "irrsinn" ("insanity").

It is questionable whether the autocomplete function is an impermissible influence on the opinion-forming process (which makes regulation by the state appear to be necessary), which norms already exert a regulating influence on search engines, and whether these are sufficient to protect the opinion-forming process. These questions are addressed in this article.

2. Autocompletion as a function

2.1 Technical implementation of autocompletion

When a user enters text in the search field of a search engine, some search engines, automated by means of algorithms, dynamically suggest additions for the currently entered search content while the entry is still in progress (Ghatnekar, 2013). These can be adopted for a further search, but it can also be performed without using these suggestions. Google, by far the world leader
in search engines, usually with over 90% market share (StatCounter, 2021), offers such
autocompletions automatically (Google, 2021, A). This cannot be disabled by the user (Google,
2021, B). From a legal perspective, the search suggestions are not third-party content, but the
respective provider's own (cf. BGH, NJW 2013, 2348). Although the suggestions for additions
are based on the search entries of other users, the decision as to which of these are displayed
and, above all, in what order and weighting, lies solely with the provider. As a result, these are
attributable to the provider. Therefore, the provider cannot invoke liability relief with regard to
the various supplementary suggestions pursuant to §§ 8-10 of the German Telemedia Act
(TMG), since it is not a case of third-party information. In principle, there are various ways to
complete search queries. For example, the entered terms can be completed by appending terms
to them (pre- and post-supplementation (Lewandowski & Quirmbach, 2013)), or by keeping
the entered words in order or mixing them (Google, 2021, A). The pure suffixing, in which a
given word is supplemented, is the simplest due to the possibility of using an alphabetical list
(Lewandowski & Quirmbach, 2013). Additional spelling corrections can be made (Sullivan,
2011) to increase the search success of users and to avoid empty hit lists. If search queries are
to be completed, it is usually done automatically by means of algorithms in the case of dynamic
generation, simply because of the quantity and variability of search queries. This requires a
preselection of search suggestions on the part of the search engine. Since search engines are
usually designed as a black box, finding detailed information on the technical structure and
algorithms is a challenge (Hartl, 2017). Due to its success, the approach of the market leader
Google is likely to set the trend for the operation of other search engines as well. Therefore,
Google’s autocomplete function will be considered in more detail in the following.

Google generates possible completions of entered search terms through a combination of
factors (Google, 2021, B). These include searches by other users including content that is
currently popular in the user's region. If the user is logged into Google, the suggestions may
also be based on past searches by that user (Google, 2021, B). Personalized search suggestions
then appear above suggestions otherwise deemed relevant by Google (Sullivan, 2011). Unpopular,
ew terms or terms that violate Google's guidelines (Google, 2021, C) can also be
left without completion by the algorithm. If a term is new but has recently experienced a strong
increase in searches by users due to its popularity, it will also be classified as relevant and
displayed (Sullivan, 2011). In principle, existing documents or online reference works can also
be used to generate search suggestions. The order of the displayed search suggestions can be
determined according to criteria such as correspondence with the user input, popularity,
topicity, spatial proximity or even economic interests (Lewandowski & Quirmbach, 2013).

An important aspect of autocompletion is the speed of providing search suggestions
(Chakraborty, 2020), since their goal is to shorten a search process in time. For this reason,
optimization of the autocompletion process is necessary. To efficiently store and quickly
retrieve characters and words in the background on the server side, tree structures such as tries
(Chakraborty, 2020), also known as prefix trees, are suitable because their structure makes
them faster than other types of trees, such as binary search trees (Germann et al., 2009). In this
search tree, several strings can be stored at the same time and thus texts can also be indexed
(Qin et al., 2013). By storing prefixes only once in the tree, the tree becomes more compact
and thus performance is improved (Germann et al., 2009). Starting from an empty string at the
root, existing prefixes can be supplemented by further characters (Qin et al., 2013). From the
root, a trie is then searched from top to bottom. To reduce response time, such a trie can be
created periodically and stored in a database (Chakraborty, 2020).
By using caches, the provision of search suggestions can be fast and up-to-date, even in the presence of high request rates. By storing which results are already displayed to a user, they are not displayed again (Gomes, 2010).

2.2 Limitations of the reliability of the search suggestions

Due to the fact that the technical procedure of a search engine is usually not verifiable for users, its use is based on trust. Users depend on a reasonable selection of search suggestions and search results to conduct a successful search that informs them comprehensively. They are directed by the search engines’ control of the flow of information and their viewpoint or their autonomy is limited. Since users are neither specifically informed about such decisions on the part of the search engine (transparency) nor can they completely bypass them in their search, this effect is particularly relevant for opinion formation (Bracha & Pasquale, 2008). Thus, search engines have a special responsibility in providing information. In this context, some users expect the neutral rendering of information (Biermann, 2012; Dörr & Natt, 2014) whereas other users expect personally tailored search suggestions (Lao, 2013). While neutrality is often a goal for search engines, its implementation in reality is difficult (Bracha & Pasquale, 2008). Therefore, the reliability of search suggestions in terms of their representation of an objective overall picture for society is subject to different expectations and limitations. Based on the social ideal of objective reporting that encompasses all available information, limitations arise in this regard simply due to the technical circumstances of offering search suggestions by search engines.

Limitation by the search engine itself

In principle, offering search suggestions can not only accelerate a user’s search process, but also hinder it in the case of unsuitable suggestions (White & Marchionini, 2006). A selection of the search suggestions to be displayed is made by the search engine (either by manual definition of a provider or automated according to criteria). Because no human plausibility check can take place during suggestion generation by algorithms before display among other things due to time constraints, the search suggestions can also be useless or undesirable for the user. Users are not informed about what information has gone into forming the search suggestions. Thus, the source and the process of creation remain unknown. At the same time, the user has no control or monitoring options.

Search suggestions can be deliberately withheld by the search engine if they contradict internal guidelines (e.g. in the case of hate speech) or if other interests are to be enforced, e.g. for reasons of cooperation with a company or to suppress possible competition (Duhigg, 2018; Tschang, 2008). This can happen unnoticed by the user (Dörr & Natt, 2014) and is usually not subject to external quality control, since the exact selection criteria and algorithms are not publicly available. In this context, the selection of the remaining search suggestions is not necessarily neutral and independent (Dörr & Natt, 2014). If the search suggestions are additionally personalized, the user is presented with an even more narrowly tailored selection during the search, thus limiting the scope of his or her search spectrum in advance. Because search suggestions are limited in number, the ones finally displayed have particular weight for a user’s search. Web page views can also be influenced by the autocomplete function, regardless of how useful the website’s offer is to the user. If numerous users are given certain
word combinations for a search by autocomplete suggestions, it is conceivable that web pages that (coincidentally) match these will be called up more frequently than others, as these are listed further to the front for this search word combination. Web pages that are not listed on the first search results page (Jansen & Spink, 2003) or even on the first page further down (Dean, 2019) are paid much less attention to. Therefore, the user could be indirectly deprived of websites that would be better tailored to his original needs.

Limitation by third parties

Influence can also be exerted by third parties, whereby this can take place either directly or indirectly. In the indirect approach, the aim of third parties is to influence the search suggestions displayed by a search engine by exploiting the internally defined selection criteria. The creation of artificially introduced search suggestions, also called “fake queries” (Sullivan, 2011), can be done for various reasons, such as to displace search suggestions with negative connotations (Munter, 2018) or for advertising purposes (Kunz, 2015). This can also lead to damage to the reputation of third parties (Kayser-Bril, 2020; Lapowski, 2018; Patalong, 2007). There are providers for the implementation of such projects, e.g. by having different people search for given terms over a longer period of time against payment. This is intended to suggest to the search engine algorithm (which selects search suggestions) that the topic is one that is frequently searched for by independent persons and is therefore relevant to other users. This way, these terms are included in the list of search suggestions.

It is also possible to influence the display of search suggestions for the user in a more direct way, for example by looking for ways to change the search suggestions displayed to a user. A concrete example of this is to append words to existing search suggestions by manipulating the messages sent from the server to the client (Białczak et al., 2011). While indirect influence is effective for a larger group of users, direct influence may be more targeted to a single user. Due to the fact that indirect influence has a larger area effect and does not rely on individual vulnerabilities, it seems to be more relevant from an overall social perspective.

It turns out that search suggestions can only reflect a part of what actually exists on the Internet in connection with the search term. Instead, search suggestions are subject to the internal selection procedures of the search engine itself and, in part, to external manipulation by third parties. The fact that users are mostly unaware of this and that it is practically impossible to check and thus also impossible to avert it can reduce the reliability of the search suggestions.

3. Influencability of human opinion formation

When content is displayed as a suggested complement to a user's input during a search process, it can influence the user's attitude and opinion formation, since human memory is like a reconstruction process that is subject to both internal and external influences (Van Damme & Smets, 2014). Not only is presenting untrue statements by means of search suggestions to users a risk to opinion formation, but also steering their opinion formation per se in a direction that they would not otherwise have taken themselves, for example, by directing the user's attention to certain topics. Existing judgments of a user can thus also be unconsciously
reinforced and untrue information can be suggested by the search suggestions unintentionally. Therefore, there is a risk of unconscious influence, which can also have effects on society as a whole. Various effects can occur that reinforce this impact.

The strength of the effect of the influence also depends on the characteristics of the user confronted with it. Thus, less analytically and critically thinking user groups are likely to be more susceptible to being influenced (Pennycook & Rand, 2019). However, other users are not immune to the impact of the effects of such (usually unanticipated) influence, as they operate at a lower cognitive level (Pennycook & Rand, 2020).

Table 1: Phenomena of the autocomplete function in their effect on individuals and society

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter bubble</td>
<td>It is suspected that filter bubbles can develop on the Internet when a user is automatically presented with content by algorithms that corresponds to his or her existing opinion and thus reinforces it. This misrepresentation of the reality surrounding the user in the form of similar content can lead to the view that one's own opinion is the dominant one. Isolation of users from each other can be the result. Users are often unaware of the filter bubble (Pariser, 2012). Personalized search suggestions could trigger this. Results of other studies indicate that filter bubbles occur less frequently or at least are less problematic in their effect than initially assumed (Bruns, 2019; Flaxmann et al., 2016; Garrett, 2017; Zuiderveen Borgesius et al., 2016). Because of the Internet, users have been not only with content corresponding to their existing opinions, but also with content to the contrary.</td>
</tr>
<tr>
<td>Framing effect</td>
<td>Even small changes in the way information is transmitted or in the context in which the information is presented can put the transmitted content in a different light and lead to a change in decision behaviour (Stocké, 2002). A user's originally planned search query can be placed in a new context by suggestions from autocompletion, so that the information presented is perceived differently than it would have been originally.</td>
</tr>
<tr>
<td>Google effect</td>
<td>It is less remembered what the information was than where it can be found (Krosnick &amp; Alwin, 1987; Sparrow et al., 2011). This favours frequent use of the search function of search engines, so that users are frequently confronted with automated search suggestions. This can act as a kind of catalyst for the effects of the autocomplete function.</td>
</tr>
<tr>
<td>Illusory truth effect</td>
<td>Information that is picked up repeatedly through repetition is more likely to be believed to be true, even if it is not (Fazio et al., 2019). The Google effect, through which repeated confrontation with suggestions occurs during autocompletion, can favour this. Repeated confrontation with such information by the autocomplete function of a search engine can trigger the illusory truth effect. It also works when the information presented is not very plausible (Fazio et al., 2019) or even when such misinformation has been warned against (Pennycook &amp; Rand, 2020). Thus, with repeated presentation, the autocomplete function has the possibility of having a strong influence on a user's opinion formation.</td>
</tr>
<tr>
<td>Long-term effect of rumors</td>
<td>Once rumours have been spread, they take a long time to be resolved (in social media, among other places), especially if they are incorrect (Zubiaga et al., 2016). Even after a possible subsequent negation or correction, the effect of recorded information is usually not completely reversed (Johnson &amp; Seifert, 1994).</td>
</tr>
<tr>
<td>Phenomenon</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Misinformation effect</td>
<td>If untruthful reporting follows an event, the user's memory of this event is distorted (Blank &amp; Launay, 2014). Search suggestions for an event can influence a user's memory by presenting further (possibly incorrect) information. Negative information is given a higher weight by people than positive information (Wojciszke et al., 1993).</td>
</tr>
<tr>
<td>Primacy effect</td>
<td>When researching a subject with a search engine, one of the first pieces of information a user picks up is likely to be offered by the autocomplete function. Since the first information a user takes in about a subject has a stronger impact than the subsequent information (Krosnick &amp; Alwin, 1987), this can have a great influence on the formation of a user's opinion, based on the course of a search process alone.</td>
</tr>
<tr>
<td>Reputation heuristic</td>
<td>Instead of reviewing the published content of a source, the credibility of a piece of content is based on the reputation of the source behind it, as this is easier to verify. If the search engine is perceived as a trustworthy, credible source, suggestions by it may be perceived as particularly credible (Metzger et al., 2010), even if the suggestions generated are not necessarily true.</td>
</tr>
</tbody>
</table>

The effects from Table 1 show that opinion formation can be influenced. Due to the mass use of search engines, these effects can in principle take effect among a large number of users. Since search suggestions can be generated from the searches of other users, a self-reinforcing effect is conceivable. For this reason, the ability to influence human opinion formation through search engine autocompletion can be regarded as a mass psychological phenomenon that must be examined under a law perspective.

4. Legal evaluation of the phenomena of the autocomplete function

Due to the influence on opinion formation, it can be asked whether the autocomplete function in its current form is opposed by legal norms in order to counteract the phenomena named in Table 1.

4.1 Relevant standards

A special law for regulating search engines or for safeguarding the formation of opinion does not exist de lege lata. In particular, the German Network Enforcement Act (NetzDG) applies exclusively to social networks pursuant to § 1 (1) sentence 1 NetzDG. Only selected provisions in the German Interstate Treaty on the Media (MStV) serve to safeguard diversity of opinion. This is a treaty between the federal states (Martini, 2020). As a reaction to the digital progress, media platforms and media intermediaries are also covered in addition to traditional linear media (Martini, 2020). This means that search engines are also covered by the term "media intermediary," which is kept open in principle (Paal & Heidtke, 2020).

4.2 Legal evaluation of the phenomena from Table 1

With regard to the danger posed by filter bubbles and framing, the regulations in the German Interstate Treaty on the Media (MStV) are not sufficient. The personalization of supplementary suggestions in combination with similar queries from other users does not constitute direct
discrimination within the meaning of § 94 (1), (2) MStV, since the search suggestions displayed are not (third-party) journalistic-editorial content (cf. BGH, NJW 2013, 2348). At most, there is an indirect obstruction in two respects: First, the search suggestion cannot be classified as third-party content and therefore, no third-party content is discriminated. Second, it is the user’s selection that causes the search results to be displayed and could cause an obstruction in the sense that certain journalistic-editorial content does not even appear in the list of results because a term suggested by the search engine is added to the search entry. Finally, such discrimination would also have to be systematic (Frey et al., 2021; Zimmer & Liebermann, 2020) and unfair, and must therefore not pursue legitimate interests such as combating hate speech on the Internet. These requirements will hardly be verifiable in practice.

It is true that the misinformation effect can be counteracted by the fact that false assertions regularly do not enjoy any protection and can be attacked if they constitute a violation of personal rights (cf. BGH, NJW 2013, 2348). However, insofar as the search supplement suggests a false fact or at least a connection that cannot be proven but does not infringe any third-party rights in the process, no action can be taken against this. § 93 and § 94 MStV are not applicable in this case due to the lack of discrimination against journalistic-editorial content.

The further phenomena of reputation heuristic, the long-term effect of rumors, the Google effect and the primacy effect, which occur during the search input, are merely phenomena that occur generally in connection with searches and are not legally tangible.

As a result, the regulations of the MStV to date are not suitable for containing the described social phenomena and their effects on opinion formation (Schwartmann et al., 2019).

5 Autocomplete as an influence on opinion formation in need of regulation

5.1 Necessity of regulation according to constitutional law

The ability to influence opinion formation thus raises the question of whether state regulation is necessary. In order to answer the question, the constitutional framework has to be analysed. The freedom of opinion and information under Article 5 (1) sentence 1 of the Basic Constitutional Law (GG 5.2) and the freedom of the press and broadcasting under Article 5 (1) sentence 2 GG are the main constitutional provisions. According to Article 5 (1) s. 1 GG, everyone has the right to express and disseminate his opinion freely in speech, writing and pictures and to inform himself without hindrance from generally accessible sources. From this follows not only freedom of expression, but also that the process of opinion formation has to be protected by the state (Kreile, 2017). According to Article 5 (1) s. 1 GG there is also a right to access information, which is referred to as freedom of information (Grabenwarter, 2020). The various freedoms of communication represent an "objective principle" that serves the principles of democracy, the rule of law and the welfare state (Grabenwarter, 2020). Fundamental rights are addressed to the state. It is thus the responsibility of the state to ensure through appropriate measures, that the fundamental right of freedom of expression (Kellner, 2019) and, concomitantly, diversity of opinion are guaranteed (Dörr & Natt, 2014). It can therefore be concluded that the state can, if necessary, also intervene in a regulatory manner to protect diversity of opinion. The same can be inferred from the freedom of broadcasting. The
freedom of broadcasting under Article 5 (1) sentence 2 GG protects broadcasting itself from state control and influence, but at the same time serves to shape broadcasting as part of the constitutionally protected opinion-forming process by means of a positive order within the framework of technical possibilities in such a way that a comprehensive range of information is available (see BVerfG 1981, 1774 (1774 et sqq.)). But digital services such as search engines cannot be assigned to the concept of broadcasting, and the Federal Constitutional Court has not yet been able to continue its jurisprudence guaranteeing diversity of opinion with regard to media intermediaries. Nevertheless, it can also be applied to intermediaries that they play an important role in the process of opinion formation in society, which should then be regulated by the state if diversity of opinion is restricted or cannot be guaranteed. Other authors even speak of a first-right obligation for the legislator to shape the regulatory model for media intermediaries (Kreile, 2017; Schwartmann et al., 2019).

The aforementioned phenomena, especially the misinformation or framing effect, lead to the promotion of a certain opinion and the suppression of others. A free formation of opinion is also not possible because, due to a lack of transparency, it is not possible to see what the supplement proposal is based on (Kühling, 2021). This is a danger to the diversity of opinion, which obliges the state to take measures to protect the diversity of opinion.

5.2 Need for regulation and adaption of existing algorithms

Since the MStV is not suitable for preventing the described negative effects on the formation and diversity of opinion, it must be asked what kind of regulation is required. Provided that search suggestions are used on the basis e.g. of their popularity, comparable search queries, or correct spelling errors, it is an improvement of the user experience. The frequent appearance of search suggestions based on an algorithm that filters by popularity among other users represents a social process that, in the absence of unilateral preference or disadvantage, does not require any regulation. Nevertheless, the influence of search suggestions should not be underestimated. Since search suggestions are bound to be located at the beginning of a search, the primacy effect is to be expected. This underscores the extent of the potential impact of the information provided by a search engine. Due to the Google effect, frequent use and thus frequent confrontation with this content can be expected. Both phenomena thus underline the social relevance of the autocomplete function and thus the need for regulation.

Transparency

A better insight on the user side and thus more transparency could be achieved by marking search suggestions according to which parameters result in the specific suggestion or which are the decisive factors for this. This could be implemented, for example, with symbols next to a search suggestion. An arrow rising upwards could indicate the supplementary suggestion(s) that are currently being entered more frequently by other users (trend). In contrast, connected dots could symbolize that a suggestion is based on similar search inputs. Finally, a stylized person could stand for the personal suggestions.

Regulatory options

A two-pillar model could be considered as a further regulatory option (Schwartmann et al., 2019). Based on such a model, the search suggestions could contain both terms that the search engine operator selects on the basis of personalization and other internal criteria (pillar 1) and
search suggestions that are intended to represent the given topic as diversely as possible, regardless of personalization (pillar 2). Combined with the previous remarks on transparency and labelling of the corresponding suggestions, this could counteract the phenomena described, especially filter bubbles.

6 Conclusion

It can be seen that search suggestions, as soon as they are displayed, can have an effect on the opinion-forming process during a search process. Only the option of no longer displaying search suggestions could avert this. As soon as search suggestions are generated, the responsible algorithm necessarily influences opinion formation, since it is not possible to display all possible suggestion combinations simultaneously and equally. Weighting or pre-filtering is necessary for this function.

Because the opinion of users can be influenced, this topic is of social relevance. However, this influence is not per se inadmissible, as it is not targeted discrimination, as long as there is no systematic deviation from the defined criteria.

In the future, the influence on the opinion-forming process could be weakened by additional transparency. It is conceivable here, for example, to make the criteria used for selecting search suggestions accessible to users (and to enable verification by experts). In addition, consideration should be given to a regulation on the security of diversity of opinion, which does not only link to the presentation of journalistic-editorial offerings, but already takes into account the upstream influencing of users by search supplement suggestions. These involve the risk of circumventing existing regulations.

Securing search suggestions against external influence can provide additional security, both in the case of indirect influence through manipulated search queries and in the case of the direct connection from the search engine to the user.

References


Dean, B. (2019). We analyzed 5 million Google Search Results. Here’s What We Learned About Organic Click Through Rate. [Online]. Available: https://backlinko.com/google-ctr-stats


