

Increasing visibility and discoverability of scholarly publications with academic search engine optimization

With the help of academic search engine optimization (ASEO), publications can more easily be found in academic search engines and databases. Authors can improve the ranking of their publications by adjusting titles, keywords and abstracts. Carefully considered wording makes publications easier to find and, ideally, cited more often. This article is meant to support authors in making their scholarly publications more visible. It provides basic information on ranking mechanisms as well as tips and tricks on how to improve the findability of scholarly publications while also pointing out the limits of optimization. This article, authored by three scholarly communications librarians, draws on their experience of hosting journals, providing workshops for researchers and individual publication support, as well as on their investigations of the ranking algorithms of search engines and databases.

Keywords

academic search engine optimization; scholarly publishing; findability; discoverability; scholarly communication; dissemination



LISA SCHILHAN

Head of Publication
Services
University of Graz
Austria



CHRISTIAN KAIER

Librarian
Publication Services
University of Graz
Austria



KARIN LACKNER

Librarian
Publication Services
& Information
Services
University of Graz
Austria

Introduction: ASEO – what, why and how?

What is ASEO?

Search engine optimization (SEO)¹ is a strategy used in online marketing to improve the findability of websites and documents in search engines. Commercial websites have been applying SEO broadly for years. Academic search engine optimization (ASEO) specifically refers to academic texts and aims both at providing researchers with the best possible support in finding relevant results for their search queries and at helping authors to achieve a better ranking of their own publications in search engines and databases. Elements to be 'optimized' include the wording of title and abstract as well as the choice of keywords and the provision of rich metadata.

2 Why is ASEO important for researchers?

The 'relevance' and 'impact' of research are measured, among other things, by counting the number of views, downloads and citations a publication has received. It is therefore essential for researchers and institutions that their own publications are found easily in the information retrieval systems of libraries, like library catalogues or literature databases, as well as in academic search engines like Google Scholar or BASE – because only what can be found will be read and possibly be cited.

Wide dissemination of research output is also called for in research assessment and is commonly measured by the views, downloads and citations scholarly publications receive. Research funders aiming to ensure the 'impact' of funded projects even require dissemination strategies to be explicitly mentioned in funding agreements. In the Horizon 2020 Grant Model Agreement,² for example, there are no less than seven sections that address the visibility, dissemination and promotion of research results. Specific guidelines have been prepared to provide guidance concerning the use of social media in EU projects as well as instructions on how to use them effectively.

At the same time, the output of scholarly literature is constantly growing.³ This has not only made it difficult for authors to attract attention to their own publications but also contributes to a looming discoverability crisis,⁴ making it harder for readers to identify relevant content. By providing rich and informative metadata, ASEO helps to avoid information overload and improve orientation.

Do it the ethical way

Standards of good scientific practice and research integrity must take precedence over any 'optimization' of publications and their metadata. As opposed to conventional SEO for purely commercial purposes, ASEO is a sensitive domain that must be approached with a sense of proportion and appropriateness.

Authors will have an interest in formulating the titles of their works in such a way that they are easier to find. In this endeavour, a balance must be found between creative freedom, publication culture, research integrity, visibility and findability. In the attempt of optimizing scholarly publications in terms of visibility, research results must not be inflated or distorted, and no false expectations must be raised with regard to content and relevance. It is essential to maintain a balance between increasing visibility and presenting high-quality research. Any 'over-optimization' must be avoided as it would not only complicate the search for relevant research but also harm the reputation of science as a whole and of the individual author(s).

'It is essential to maintain a balance between increasing visibility and presenting high-quality research'

Considering all these aspects, what can authors do to make sure their publications will be found?

ASEO and relevance ranking

Scholarly articles and books are usually searched for in Google Scholar, online catalogues of libraries and literature databases. Most of these search systems are based on artificial intelligence, the aim of which is not only to find items corresponding to the search terms entered but also to rank the results by relevance. This relevance sorting or ranking is based on algorithms that are considered a 'trade secret' of the respective company. With search engines like Google, Primo or EBSCO, for instance, it is crucial for each provider to deliver the 'best' results. The aim is not always to offer the greatest possible number of results but to display the most 'relevant' hits on top of the list of results.

How relevance ranking is done

In order to perform such relevance sorting, a large number of factors are taken into account. The basic mechanisms of the relevance ranking can be identified even though commercial providers usually do not disclose how exactly these factors are considered. One key element

3 is bibliographic metadata: formal data such as titles, subtitles, names of the authors, identifiers and keywords are analyzed and ranked for each search query. The search system assesses the frequency of the search term in the metadata as well as the position of the search term within the publication. When the search query is 'climate change', for example, a document containing the search term in its title will be ranked higher compared to a document where the term only appears in the abstract.

However, ranking is not only determined by the appearance and position of the word in the title, abstract or the full text. The frequency of the term in the metadata or the full text will also be considered. The more often a search term appears in the metadata, abstract and full text, the more relevance points will be assigned to the document by the algorithm. If the full text of the article is made available open access, a wider range of words can be searched, which improves the relevance ranking. The points are aggregated and the documents ranked accordingly.

Depending on the search engine technology, many other factors such as the year of publication – recently published articles and books are considered more relevant and therefore ranked higher – citations and views in relation to the total number of documents found and the journal impact factor may also influence the ranking.

The next sections offer guidance on how to make scholarly articles more visible or, to put it the other way round, on how to avoid a range of mistakes that prevent researchers or search algorithms from identifying literature as relevant.

ASEO for scholarly articles

Title optimization

The most vital element of a document is its title, and authors should consider its wording particularly carefully. The search terms occurring in the title have the highest relevance, not only for search systems but also for users. In a result list containing several hundred hits, readers will make an initial analysis of relevance based on the title when scanning the result list.

Meaningful and succinct

The length of academic titles has been a subject of studies⁵ that recommend aiming for short titles. Longer titles tend to receive fewer citations than short, concise titles. Incorporating the result of a study in the title to make it more declarative⁶ will also increase perception.

The most important keywords should occur at the beginning of the title to allow readers to easily identify the content of a publication. It is not recommended to 'hide' the most important keywords or topic words in the middle or at the end of a title. For instance, the title 'On facing the crucial psychosocial and political-economic dimensions of anthropogenic global warming' starts with some words of little relevance ('On facing the'), so that the most prominent position within the title is 'wasted' on less important words, whereas the most significant words ('anthropogenic global warming') are hidden at the end of the rather long title. Thus, a reader will not get the most important information at first glance. Moreover, search engines like Google will shorten the lengthy title and therefore cut off the final – and in this case most important – title words, as is shown in Figure 1.

'The most vital element of a document is its title'

On facing the crucial psychosocial and political-economic dimensions of ...

24.01.2020 · Anthropogenic global warming is one of the most significant existential threats facing the human species. Nonetheless, most individuals largely ...

Figure 1. Screenshot of a lengthy article title in Google Search displayed on a mobile phone. The most important keywords are at the end of the main title, which are cut off due to the limited space

- 4 A problem that sometimes arises when publishing in a special issue or an edited volume is that titles sometimes lose their most essential key words. Published in the overall context of a special issue a title like this would work: *Defensive strategy and the construction of the wall.*⁷ If disseminated on various platforms and repositories, without the overall context of the book, the title does not indicate which wall the article refers to: it could be the Berlin Wall, Trump's wall at the Mexican border, the Israeli West Bank wall or the Great Wall of China. When scrolling through a list of 300 articles about the Berlin Wall, a result that does not contain the search term at all is not very likely to be clicked on. Furthermore, when searching for, for example, 'Berlin Wall construction', the ranking algorithm would rate the above-mentioned title, which contains only two of the three search terms, quite low, and numerous titles containing all three words would be ranked higher in the result list, even though the article in question perfectly matches the search query in terms of content. However, the ranking algorithm cannot 'know' this by simply analyzing the title.

'when creating a title, make sure that it is not misleading or ambiguous'

Therefore, when creating a title, make sure that it is not misleading or ambiguous when displayed out of context and that both the reader and the ranking algorithm can identify the subject.

Main title vs. subtitle

One strategy for optimization that has less to do with search engine technology and more to do with the search behaviour of users, is to use titles and subtitles. Particularly in the humanities and social sciences there is a deep-rooted propensity for creative main titles designed to arouse interest or to illustrate the topic more vividly, even if the title will not always accurately reflect the content of the article. The subject-specific context is often only made clear in the subtitle. However, search systems classify subtitles as less relevant than the main title.

From an ASEO perspective there should be more emphasis on the title and less information hidden in subtitles. Many databases do not even display subtitles in the list of results, which indicates that these articles will also be poorly ranked. Furthermore, it is difficult for users to identify the subject of an article, especially if the title is very long and therefore cut when using a mobile device. This applies in particular to titles that contain quotations in the main title. An article titled "*There are no cats in America!*": *The Sea Voyage in Don Bluths "An American Tail" and Shaun Tans "The Arrival" as a "Representation of Liminal Migration Experiences"* (see Figure 2) contains sources of error for successful indexing and

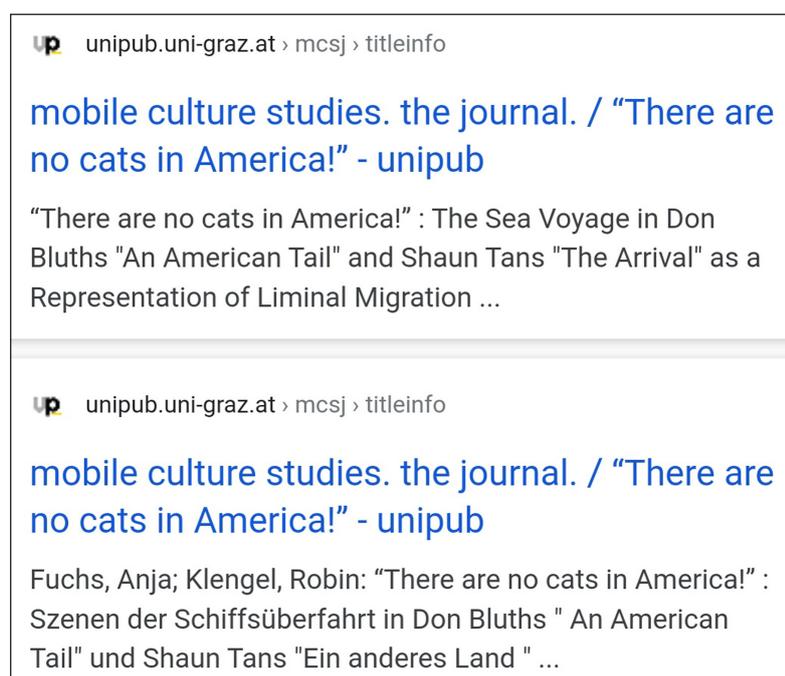


Figure 2. Screenshot of a creative article title in Google Search displayed on a mobile phone. The most important keyword 'migration' is at the end of the subtitle, which is being cut off in the original language (German)

5 ranking by search engines and citation analysis tools and will often be cut off so that the most important word 'migration' will not even be displayed. Therefore, readers who only see the title 'There are no cats in America!' in the list of results will find it difficult to link it to the topic of migration. The user may perceive the article as having limited relevance or may attribute the appearance of the title to a malfunction of the search engine algorithm. Consequently, he or she may not click the title to see the full entry in the database and therefore neither read nor cite it.

To avoid this problem, catchy phrases or funny remarks should only be used in the subtitle and the most important keywords should occur in the main title, preferably at the beginning.

'A title that is meaningful to the reader can be difficult to index by automated search systems'

Suspended hyphens

A title that is meaningful to the reader can be difficult to index by automated search systems. Scholarly literature by its nature deals with complex issues and therefore tends to be full of phrases and compound expressions. Suspended hyphens are used to shorten a list of compound words or phrases, such as pre- and postnatal stress. This is sometimes used in the English language but even more often in other languages which have many compound words, like Germanic languages.

For example, as shown in Figure 3, 'pre- and postnatal' can be found in various spelling forms in scientific articles:

The impact of pre- and post-natal smoke exposure on future asthma and bronchial hyper-responsiveness.

Goksör E, Amark M, Alm B , Gustafsson PM, Wennergren G

[Author information](#) ▶

Acta Paediatrica (Oslo, Norway : 1992), 10 May 2007, 96(7):1030-1035
DOI: 10.1111/j.1651-2227.2007.00296.x PMID: 17498194

Effects of pre- and postnatal olfactogustatory experience on early preferences at birth and dietary selection at weaning in kittens.

Becques A, Larose C, Gouat P, Serra J

[Author information](#) ▶

Chemical Senses, 04 Dec 2009, 35(1):41-45
DOI: 10.1093/chemse/bjp080 PMID: 19965899

Effects of pre and postnatal exposure to low levels of polybromodiphenyl ethers on neurodevelopment and thyroid hormone levels at 4 years of age.

Gascon M , Vrijheid M, Martínez D, Forns J, Grimalt JO , Torrent M, Sunyer J

[Author information](#) ▶

Environment International, 14 Jan 2011, 37(3):605-611
DOI: 10.1016/j.envint.2010.12.005 PMID: 21237513

Effects of prenatal and postnatal maternal emotional stress on toddlers' cognitive and temperamental development

Yanfen Lin ¹, Jian Xu ², Jun Huang ³, Yinan Jia ¹, Jinsong Zhang ¹, Chonghuai Yan ¹, Jun Zhang ¹

Affiliations [+](#) expand
PMID: 27665073 DOI: 10.1016/j.jjad.2016.09.010

Figure 3. Screenshots of some titles with different spellings of the term 'pre- and postnatal'

- 6
- pre- and postnatal with a hyphen after 'pre'
 - pre- and post-natal with an additional hyphen between 'post' and 'natal'
 - pre and postnatal without any hyphen at all
 - prenatal and postnatal with both terms in full.

The issue with suspended hyphens is that we as human readers can make the semantic connection but search engines often cannot. Humans will readily combine the parts 'pre-' and 'natal' to pre-natal or prenatal, while search engines cannot always combine them. Figure 3 presents four different spellings of the term 'pre- and postnatal', with the first three titles containing one of the variants of 'pre- and postnatal' with or without hyphens and the fourth title containing the terms 'prenatal' and 'postnatal' written out in full. So, the search for 'prenatal stress' would not consider the first three titles as highly relevant while the fourth title would be ranked highly as it contains the search term in the exact form.

There has been research on how a hyphen in a title can lead to reduced findability. As far back as 1999 John Bowman stated⁸ a lack of findability for titles with hyphens and ampersands (&) in the old online library catalogues (OPAC). Hyphens not only impair the findability and ranking of a title in a search engine or database and therefore may have a negative effect on the number of citations, but they also even affect the citation counting itself. A study published in 2019 by Zhou et al.⁹ shows that articles with hyphens in their titles are cited less frequently. This is due to the fact that articles with special characters are not always cited correctly. The more special characters there are in a title, the more likely authors will be to quote the title incorrectly, which in turn makes it impossible for citation analysis tools to count it as a citation. The citation analysis tools will not assign these misquoted references to the respective article. Furthermore, citation database systems lack robustness in automated citation counts.¹⁰ Search engines and databases which include citation counts in their ranking algorithms will rank these articles lower, which results in reduced visibility and, consequently, fewer citations – a vicious circle. Authors should therefore choose a title that will minimize such errors.

'The more special characters there are in a title, the more likely authors will be to quote the title incorrectly'

Special characters, diacritical signs and formulae

The recommendation to avoid possible errors also concerns the use of diacritical signs in other languages: ä, å, æ, ø as well as mathematical formulae and gender-sensitive language. These should be avoided in the interest of ASEO as far as possible; and special characters such as asterisk, slash and hyphen should be avoided in the attempt to use gender-sensitive language. It is recommended to forgo both the male and the female form and word the title without the use of a gender-specific word. Instead, both variants (female and male) should be used. While this is not a concern for English texts and titles, it is of particular interest for Germanic languages. The German 'Autor*in' (gender-neutral expression for author), for example, indicates that a search engine would have problems finding the variant 'Autorin' (female author) and thus, would either not be found at all or be ranked in a less prominent position.

Keep in mind that the differences between American English and British English may not be regarded in search systems. For example, a test search for 'optimization' generated about 5,000 results whereas 'optimisation' only led to less than 600 results in the library catalogue unikat.

Mathematical formulae can create another set of problems: the representation in basic text code. It is vital to limit characters to basic text code like UTF-8.¹¹ In that case, if special characters like the Φ symbol cannot be displayed properly the problem lies with the end user's device. Since we do not have control over these devices it is advisable to avoid special characters. The search result itself will also be compromised when using special characters. For example, as seen in Figure 4, the book *A Dynamical Perspective of the Φ^4 Model* cannot be found in the Austrian consortium library catalogue if you search for 'a dynamical

7 perspective of the phi 4 model'. In order to find the book, you will need to use the Φ symbol. If the platform that hosts the book cannot display a special character like the Φ symbol, the search functionality is impaired.

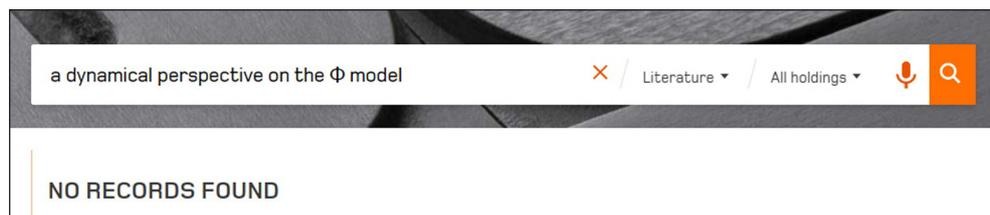


Figure 4. The title contains a phi symbol. Even if searched with a phi symbol, the book cannot be found in the library catalogue although the library holds a copy of it

The display of formulae in titles often also creates problems in lists of results. Mathematical formulae with superscript and subscript numbers and mathematical symbols often cannot be captured by the software of library search engines and databases and must be 'paraphrased'. Formulae are therefore often reproduced in a very cumbersome manner in lists of results and are also a frequent source of errors. Figure 5 shows the correct formulae on an article while Figure 6 shows the same title as an indecipherable string of special characters, found in a literature database.



Figure 5. Mathematical terms in a title displayed correctly on the article

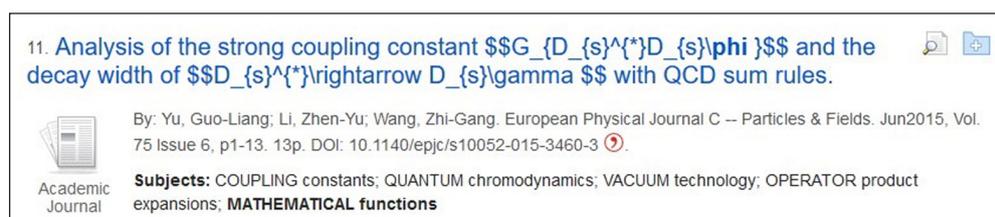


Figure 6. Same title as shown in the database 'Academic Search Premier'. All mathematical terms are indecipherable and not displayed correctly

Summary: Title optimization

Authors of scholarly publications are recommended to use meaningful titles and include the most relevant keywords or phrases in the main title rather than the subtitle. The most important terms should be placed at the beginning of the title so that users and search engines can identify the content of the publication easily. Authors should keep in mind that their potential readers will find their work mostly by searching for specific search terms in academic search engines. Those search terms should be included or at least considered. It is good practice to keep the titles short and easy to identify. Special characters should be avoided since they may cause errors in the display and make an article difficult to search for. Abbreviations impair the search functionality and should only be used if they are well-known. Lastly, it is part of good scientific practice not to overstate your results. The visualization of the most important tips is shown in Figure 7 which can be used in publication training courses.

'use meaningful titles and include the most relevant keywords or phrases in the main title rather than the subtitle'

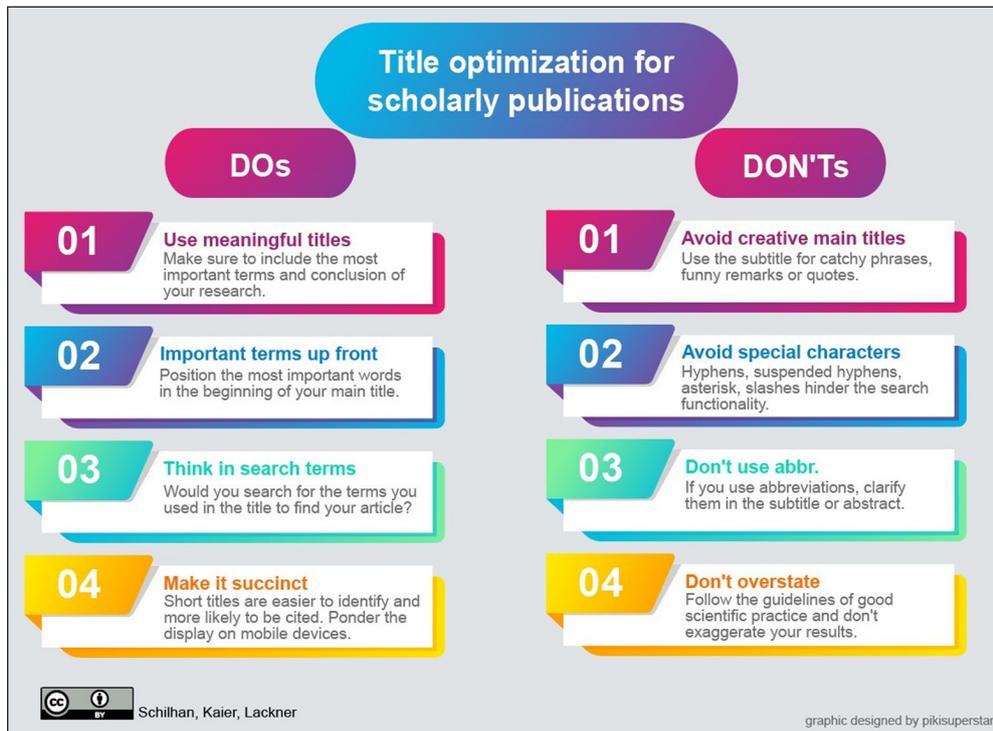


Figure 7. Title optimization for scholarly publications dos and don'ts

Keyword optimization

Keywords that accurately describe the content of a scholarly publication significantly improve its findability. Authors asked to provide keywords when submitting their article should therefore reflect carefully on the selection of descriptive keywords. The use of thesauri may be helpful to identify suitable keywords. Discipline-specific thesauri such as Medical Subject Headings (MeSH) and Embase Subject Headings (Emtree®) can be helpful, and numerous specialist databases offer comprehensive thesauri of their own. Google's keyword tools such as Google Trends can also be used to search for adequate keywords. It is also advisable to consider which terms you would use yourself when searching for a publication on the topic in question.

'Keywords that accurately describe the content of a scholarly publication significantly improve its findability'

Furthermore, it is important to distinguish between 'narrow' and 'broad' terms with regard to the effectiveness of the search. If a 'narrow' (i.e. very specific) term is mentioned in the title, the 'broader' term could be used in the keywords to increase discoverability on search databases. A keyword should always be in the singular. Authors should place themselves in the position of a searcher and consider how they would search in order to find an article. Vague, non-representative words such as 'comparison', 'characterization' and 'criticism' should be avoided. It may be more useful to describe the content of a publication through a reference to its temporal aspect, the place or country, the studied person or methodology.

Ideally, keywords should also provide additional information that is not obvious from the title. For example, authors provided their article 'Maternal mRNA input of growth and stress-response-related genes in cichlids in relation to egg size and trophic specialization' with the keywords 'Haplochromine cichlids/Maternal mRNA/Eggs/Trophic specialization/Adaptive radiation/East African lakes'. The keyword 'East African Lakes' provides important additional information; 'East Africa' might be even better, since a searcher would probably search for 'cichlid East Africa' rather than 'cichlid East African Lakes' for the sake of simplicity.

Keywords should also be meaningful and unmistakable. In the previous example, instead of just using 'eggs', 'fish eggs', 'spawn' or 'roe' would be better. Someone searching for 'eggs' will be more likely to search for literature about birds'/chicken eggs; someone searching for

9 literature about fish eggs will use rather more specific search terms than just 'eggs', because the search term 'eggs' will mainly return hits for birds'/chicken eggs.

Furthermore, the number of keywords used should be kept to a reasonable amount to avoid 'keyword spamming'. Excess keywords will be automatically blocked from the list of results by search engines. Using three to seven keywords per article is recommended in most guidelines.¹²

The tips are summarized in Figure 8 which can be used in publication training.



Figure 8. Keyword optimization for scholarly publications

Abstract optimization

Abstracts are an important element in supporting the findability of an article in scholarly journals and databases. An abstract should be formulated in an informative manner and contain the study design or method and the results, as well as relevant terms.

There are various ways to optimize abstracts in terms of findability, such as the use of synonyms. On the one hand, synonyms increase the probability of being found via different search terms. On the other hand, they reduce the risk of not being found just because a certain search term does not appear in the title, subtitle or abstract and therefore the article is not included in the list of results although the content may match the search query.

- 10 The most important terms referenced in the article should be included at the beginning of an abstract, as not all databases and search engines display the entire abstract. Where there are abbreviations for a technical term used frequently in an article, the term should be written out in full as well as appear in its shortened form in the abstract.

For this article we constructed this abstract:

'With the help of **academic search engine optimization (ASEO)**, publications can more easily be found in **academic search engines** and **databases**. Authors can improve the **ranking** of their publications by adjusting titles, keywords and abstracts. Carefully considered wording makes publications easier to find and, ideally, cited more often. This article is meant to support authors in making their **scholarly publications** more visible. It provides basic information on **ranking mechanisms** as well as tips and tricks on how to improve the **findability** of **scholarly publications** while also pointing out the limits of optimization.'

'The most important terms referenced in the article should be included at the beginning of an abstract'

The most important phrase is right at the beginning, the sentences are short and highly informative. Important keywords like 'ranking', 'academic search engines' and 'findability' are used in order to be found in a search query using these keywords. The abstract gives a clear idea on what the readers can expect of the content.

A summary of the tips for optimizing an abstract is shown in Figure 9.



Figure 9. Abstract optimization for scholarly publications

Tips and tools

Every document carries a large amount of metadata, among them their title, subtitle, date, creator, file format, etc. These metadata are used to evaluate a search query in a search system. The display of an item in a list of results and its ranking depends, among other things, on the degree of overlap between search query and the document metadata. When uploading an article to, for example, an institutional open access repository or ResearchGate via an upload form, it is therefore highly recommended to add as many metadata as possible, especially keywords. Not all metadata fields can be edited by authors, but additional information in the form of metadata should be provided whenever possible, even if the fields are not mandatory.

'The display of an item in a list of results and its ranking depends ... on the degree of overlap between search query and the document metadata'

Metadata optimization of PDFs

With the PDF format for instance, it is possible to add information to the file properties. These metadata should be filled in thoroughly and correctly to improve visibility, especially on search engines. In particular, the file name, title, name of authors and keywords should be included in the PDF file properties.

When providing a PDF, the archivable PDF/A standard should be used to ensure long-term and cross-system readability of the publication. A PDF file can be converted to PDF/A using freeware, Microsoft Word, Open Office, Adobe Acrobat and other programs. These metadata are used to support discoverability on search engines.

Metadata optimization of images

Images and graphs in publications should also be optimized for findability. It is important to note that text in images can only be recognized by search engines if the image is saved in a vector graphics format such as .svg and .eps. These file types can be generated in vector graphic programs like Windows PowerPoint, Adobe Illustrator, Inkscape or Corel Draw. Text in files with the .bmp, .jpg and .png formats are not machine-readable. As an alternative, it is recommended that additional information be entered in the file metadata. Some search engines can recognize and evaluate captions. A detailed caption is, therefore, recommended.

With other research output such as software, codes, data and reports in proprietary formats, all metadata should be completed in the best possible way.

Text formatting

To help search engines 'understand' the content of a document, it is advisable to use the available features of text formatting. A heading defined as a heading will generally be recognized as such by search engines. Similarly, a bibliography will be recognized as such when it is referred to as either 'bibliography' or 'references'. If an article includes the usual components, such as a title written in large font at the beginning, names of authors and a numbered bibliography section below, it is categorized as a scholarly article by search engines and recognized, for example, in Google Scholar.

'The DOI of a publication should be part of every form of dissemination'

Persistent identifiers

One form of metadata that is widely adopted is the persistent identifier 'digital object identifier' (DOI). It makes sure that a publication can be permanently accessed and thus found via a stable link, even if its location (Uniform Resource Locator [URL]) has changed. The DOI of a publication should be part of every form of dissemination. Besides research information systems and publication databases, bibliometric tools such as Altmetrics or PlumX also use DOIs as a data source. Especially when disseminating an article via social media, it is essential to specify the DOI to maximize its visibility.

As an identifier for researchers, the ORCID iD¹³ has become a standard. Being able to unambiguously identify researchers is essential since the correct assignment of scholarly

12 publications to their authors is key for the academic careers of individual researchers, but also to identify the research output of institutions. However, the attribution of publications to their authors is prone to errors for various reasons, such as:

- name similarities and ambiguity (especially with frequently occurring names)
- name changes (for example due to marriage)
- name variations
- different spellings (like Meyer, Myer, Meyr, Meier, Meijer, Mayer, Maier, Mayr, Mair)
- foreign language characters (like ã, ā, ç, è, ł, ø, ô, š)

Creating and maintaining an ORCID profile, including one's own ORCID iD in each publication, as well as adding it when uploading publications to platforms such as ResearchGate or to open access repositories, contributes significantly to the correct assignment of publications and to the visibility of their authors.

ASEO for books

Journal articles are the dominant form of publication in many disciplines today, but books continue to play an important role in others. Still, the circulation and sales figures of (printed) scholarly books have been declining for years for various reasons.¹⁴ It is therefore particularly important to ensure that books are easy to find, display and classify to foster a broad perception.

Many of the strategies and suggestions already presented can, and should, be applied to articles and books alike. However, books (and especially monographs) are distributed and perceived differently from articles, with book series sharing characteristics of both. Monographs and edited volumes are often viewed and purchased individually, rather than as part of a series, and therefore need to be 'advertised' in a somewhat different way.

Article metadata, library catalogues, databases, repositories and academic search engines have been the focus of this article so far. With regard to books, directories of commercial traders and search engines such as Amazon are highly relevant as well. Amazon is the most important product search engine for books from a publisher's point of view, as most of the people searching there are prepared to spend money.¹⁵ Essentially, the Amazon SEO works in a similar way to the Google SEO: the search term should occur in the title or blurb of the book and the more relevant a title appears to the search algorithm, the more prominently it will figure in the result list.

While publishers usually keep the right to decide on the title of the books, they do ask authors – and editors – to fill out marketing questionnaires and to provide a blurb for their academic book. By providing blurbs of an appropriate length, several paragraphs and possibly sub-headings, authors can support the visibility of their books. This 'product text' should ideally be able to replace a 'sales talk'.¹⁶

'Amazon is the most important product search engine for books from a publisher's point of view'

Books are commonly included in databases and advertised even before they are actually published, and titles and texts in publisher previews may read differently from those of the final published work. It is therefore important not only to devise adequate and meaningful titles and descriptions but also to keep them updated – as well as optimized for discoverability – as the manuscript of the book evolves. Not all publishers seem to be fully aware of the problem, however,¹⁷ and authors should remind their publishers to pay constant attention to how their books are displayed in the relevant databases.

When publishing an edited volume, not only the title, blurb, keywords and metadata of the book itself should be optimized, but also those of the individual chapters, since they are often disseminated on various databases, platforms and repositories without the

13 overall context of the book. Therefore, the advice given in the section 'ASEO for scholarly publications' of this article can also be applied to chapters of edited volumes.

The publisher's role in ASEO

Publishers are typically authorized and commissioned by their authors to take over the distribution of the works they publish. This entails a responsibility on the publisher to ensure that publications will be discovered. In addition, publishers have a strong motivation to ensure that 'their products' will be found in times of information overload and increasing competition for the time and attention of their 'customers'. As academic publishers position their 'brand' and increase their reputation to a large extent by the 'products' they publish – books, journals and journal articles – and the sales figures, downloads and citations these achieve, branding is important and 'their' output must be as visible as possible. Authors and publishers obviously share an interest here, and close co-operation between publishers and authors absolutely makes sense.

Metadata are not only the basis for discovering relevant articles and books for libraries, bookshops and academics – the primary 'customers' of scholarly literature. Metadata provided by authors in the initial stages of proposal or submission also help publishers to determine how, and in which community and subject, a publication can be promoted. If metadata are consistent, book series and serial publications, as well as journals and articles, are easier to find for readers and to advertise by publishers.

'Many publishers ... have started informing and educating their authors on ASEO'

Many publishers not only ask their authors to provide further information for marketing purposes, for example by filling out marketing questionnaires, but have started informing and educating their authors on ASEO.¹⁸ Authors are increasingly asked to share and promote their works via their own networks, social media, blogs and services like Kudos in order to attract a broader audience. They should also consider offering summaries of their works in different languages.

In spite of the eminent role of ASEO, smaller publishers in particular still do not maintain their metadata accurately and consistently and keep missing opportunities for attention.¹⁹ As an author striving to maximize visibility it is therefore important not to passively rely on what a publisher does but to get involved and actively check and suggest improvements for the marketing and promotion of one's work.

Publishing cultures and limits of ASEO

When trying to enhance the visibility of scholarly publications it is essential to be aware of, and to comply with, the 'secret rules' of academia and specific disciplines. What is considered appropriate and reputable with regard to dissemination and advertising of scholarly publications varies widely between (research) cultures – there is no 'one size fits all'.²⁰

Strategies for drawing attention to one's publications vary between disciplines: in the humanities it is common practice to use creative wording and quotations in titles, subtitles or subheadings. While this is less than ideal from an ASEO perspective (as discussed in the section 'Title optimization'), many authors in the humanities consider it to be an appropriate and promising strategy for attracting attention and illustrating their topic and readers are accustomed to these practices. In general, there is a different, more playful use of language in the humanities than in the sciences, where, generally speaking, simpler, standardized language is used. There is much more variation and diversity of languages in the humanities, and not one dominating language of scholarly communication.

While journal articles are often structured in a uniform way (see for example IMRAD²¹ – Introduction, Methods, Results, and Discussion) in the sciences, with even subheadings being more or less standardized, these elements can be designed and arranged more freely in other disciplines, unlocking additional potential for ASEO.

14 The use of abstracts and keywords is still less common in the humanities than in the sciences but is increasingly becoming mainstream. Books and contributions in edited volumes are traditionally relevant forms of publication in the humanities, which also affects ASEO practices and opportunities (see above 'ASEO for books'): individual books and book series need to be 'marketed' rather than just periodicals and their articles. As a consequence, 'impact' needs to be defined differently in book affine publication cultures: there is no such thing as a journal impact factor for books or publishers.

While it is important to consider all these cultural and subject-specific differences, it is obvious that databases and search engines – unless they are highly subject-specific tools themselves – will ignore the specific background of a publication. The challenge for authors, publishers and ASEO experts alike is to be aware of, and find a balance between, these often-conflicting requirements while improving the visibility of publications.

Summary

Nowadays, mastering and applying academic search engine optimization is essential for researchers across all stages of their academic careers: good visibility and findability of publications are prerequisites for considering them in performance assessment and for career development.

ASEO complements more 'traditional' forms of marketing. Various elements of scholarly publications can and should be optimized for search engines and databases, especially their titles, abstracts and keywords. Comprehensive and correct metadata improve the visibility of publications and help other researchers find relevant literature.

As an author it is important to support, monitor and complement any measures for ASEO possibly taken by publishers, while observing the principles of research integrity as well as the conventions of the respective discipline.

Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other *Insights* articles can be accessed here – click on the URL below and then select the 'full list of industry A&As' link: <http://www.uksg.org/publications#aa>.

Competing Interests

The authors have declared no competing interests.

References

1. "Search engine optimization," *Wikipedia*, last modified January 6, 2021, https://en.wikipedia.org/wiki/Search_engine_optimization (accessed 28 January 2021).
2. Horizon 2020 Online Manual, "Communicating your project," https://ec.europa.eu/research/participants/docs/h2020-funding-guide/grants/grant-management/communication_en.htm, (accessed 28 January 2021).
3. Susanne Gaertner, "Fear of Missing Out Vs. Information Overload – Researcher Challenges in a Digital and Open Research Era," October 30, 2019, <https://www.wiley.com/network/latest-content/fear-of-missing-out-vs-information-overload-researcher-challenges-in-a-digital-and-open-research-era> (accessed 21 January 2021); Esther Landhuis, "Scientific Literature: Information Overload," *Nature* 535, no. 7612 (2016), DOI: <https://doi.org/10.1038/nj7612-457a> (accessed 21 January 2021); Andrea Polonioli, "In Search of Better Science: On the Epistemic Costs of Systematic Reviews and the Need for a Pluralistic Stance to Literature Search," *Scientometrics* 122, no. 2 (2020), DOI: <https://doi.org/10.1007/s11192-019-03333-3> (accessed 21 January 2021).
4. Peter Krater et al., "Overcoming the discoverability crisis," 2020, DOI: <https://doi.org/10.5281/zenodo.3935964> (accessed 21 January 2021).
5. Adrian Letchford, Adrian, Helen Susannah Moat, and Tobias Preis, "The Advantage of Short Paper Titles," *Royal Society Open Science* 2, no. 8 (2015): 150–266, DOI: <https://doi.org/10.1098/rsos.150266> (accessed 21 January 2021); Zhi Quan Zhou, Zhi Quan, T. H. Tse, and Matt Witheridge, "Metamorphic Robustness Testing: Exposing Hidden Defects in Citation Statistics and Journal Impact Factors," *IEEE Transactions on Software Engineering*, 2019, 1, DOI: <https://doi.org/10.1109/TSE.2019.2915065> (accessed 21 January 2021).
6. Michael J. Rossi, and Jefferson C. Brand, "Journal Article Titles Impact Their Citation Rates," *Arthroscopy the journal of arthroscopic & related surgery official publication of the Arthroscopy Association of North America and the International Arthroscopy Association* 36, no. 7 (2020): 2025–29, DOI: <https://doi.org/10.1016/j.arthro.2020.02.018>, (accessed 28 January 2021).
7. Nick Hodgson, "Defensive Strategy and the Construction of the Wall," *Archeology* 70, no. 3 (2017): 30–31.
8. John H. Bowman, "The catalog as barrier to retrieval – Part 1: Hyphens and ampersands in titles," *Cataloging & Classification Quarterly* 29, no. 4 (2000): 39–60, DOI: https://doi.org/10.1300/J104v29n04_04 (accessed 21 January 2021).

9. Zhou et al., "Metamorphic Robustness Testing." Klicken oder tippen Sie hier, um Text einzugeben.
10. Zhou et al., "Metamorphic Robustness Testing."
11. "UTF-8," *Wikipedia*, last modified on January 25, 2021, <https://en.wikipedia.org/wiki/UTF-8> (accessed 28 January 2021); "Character encodings for beginners," *W3C Questions & Answers*, last modified April 16, 2015, <https://www.w3.org/International/questions/qa-what-is-encoding.en> (accessed 28 January 2021).
12. "Top tips: making your article visible with SEO," Elsevier, <https://www.elsevier.com/connect/authors-update/top-tips-making-your-article-visible-with-seo> (accessed 22 January 2021); "Search Engine Optimization: For Authors," Wiley, <https://www.wiley.com/legacy/wileyblackwell/pdf/SEOforAuthorsLINKSrev.pdf> (accessed 22 January 2021); "SEO tips for book authors", Springer Nature, <https://www.springernature.com/de/authors/campaigns/seo-tips-for-book-authors> (accessed 22 January 2021).
13. ORCID – Open Researcher and Contributor ID, <https://orcid.org/> (accessed 26 January 2021).
14. Eelco Ferwerda, Frances Pinter, and Niels Stern, "A Landscape Study on Open Access and Monographs: Policies, Funding and Publishing in Eight European Countries," 2017, DOI: <https://doi.org/10.5281/zenodo.815932> (accessed 28 January 2020); Michael Jubb, *Academic Books and Their Futures: A Report to the AHRC & the British Library*, London, 2017, https://academicbookfuture.files.wordpress.com/2017/06/academic-books-and-their-futures_jubb1.pdf (accessed 22 January 2021); Universities UK Open Access Coordination Group, *Open Access Monographs*, 2018, <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/documents/2018/open-access-monographs-report.pdf> (accessed 22 January 2021).
15. Mareike Schubert, "Produkt-Suchmaschine Amazon: Wie Kunden ihre Bücher finden", *Buchreport*, March 21, 2019, <https://www.buchreport.de/news/pp-amazon-seo-produkt-suchmaschine-wie-kunden-ihre-buecher-finden/> (accessed 22 January 2021).
16. Mareike Schubert, "Die optimale Produktdarstellung auf Amazon", *Buchreport*, September 5, 2019, <https://www.buchreport.de/news/pp-die-optimale-produktdarstellung-auf-amazon/> (accessed 22 January 2021).
17. Mareike Schubert, "Produkt-Suchmaschine Amazon".
18. De Gruyter, "Search engine Optimization"; Elsevier, "Top tips"; Wiley, "Search Engine Optimization"; Springer Nature, "SEO tips".
19. Michael Tamblin, "Money on the Table: Opportunities Missed in the eBook Supply Chain", May 14, 2019, <https://www.michaeltamblin.com/news-1/2019/5/14/money-on-the-table-opportunities-missed-in-the-ebook-supply-chain> (accessed 22 January 2021).
20. Sara Kjellberg, and Jutta Haider, "Researchers' Online Visibility: Tensions of Visibility, Trust and Reputation," *Online Information Review* 43, no. 3 (2019): 426–39, DOI: <https://doi.org/10.1108/OIR-07-2017-0211> (accessed 22 January 2021).
21. "IMRAD," *Wikipedia*, last modified on January 4, 2021, <https://en.wikipedia.org/wiki/IMRAD> (accessed 26 January 2021).

Article copyright: © 2021 Lisa Schilhan, Christian Kaier and Karin Lackner. This is an open access article distributed under the terms of the **Creative Commons Attribution Licence**, which permits unrestricted use and distribution provided the original author and source are credited.



Corresponding author:

Lisa Schilhan

Head of Publication Services

University of Graz, AT

E-mail: lisa.schilhan@uni-graz.at

ORCID ID: <https://orcid.org/0000-0002-1425-850X>

Co-authors:

Christian Kaier

ORCID ID: <https://orcid.org/0000-0002-8750-6666>

Karin Lackner

ORCID ID: <https://orcid.org/0000-0001-6096-1717>

To cite this article:

Schilhan L, Kaier C and Lackner K, "Increasing visibility and discoverability of scholarly publications with academic search engine optimization", *Insights*, 2021, 34: 6, 1–16; DOI: <https://doi.org/10.1629/uksg.534>

Submitted on 29 September 2020

Accepted on 04 November 2020

Published on 17 March 2021

Published by UKSG in association with Ubiquity Press.