



6 February 2024

Google LLC and Google South Africa
Ballyoaks Office Park 35 Ballyclare
Building E Bryanston
Johannesburg
Gauteng
2191

Attention: Alistair Mokoena and Marianne Erasmus

By Hand and by email

By email: allymokoena@google.com / merasmus@google.com

Dear Mr. Mokoena and Ms. Erasmus

CAXTON: REQUEST FOR INFORMATION

Introduction

1. Caxton and CTP Publishers and Printers Limited and its subsidiary CTP Limited ("**Caxton**") has, over the past two years, unsuccessfully sought to achieve redress of the disputes and economic and commercial disparity between Caxton's news publications and the exploitation thereof by the dominant digital platform operated by Alphabet Incorporated, its subsidiary Google LLC and its Irish and South African affiliates ("**Google Ireland**") and Google South Africa Proprietary Limited (registration number 2006/035611/07) ("**Google SA**") (collectively, "**Google**"). We refer to all these entities as "**Google**" on the basis that Google, in its day-to-day interactions with Caxton, and in the general conduct of its business in South Africa, makes no distinction between the separate identities of these entities.
2. Caxton wishes to achieve an equitable licensing arrangement and accounting for the revenues derived by Google from exploiting Caxton's copyright protected news content, which Caxton disseminates in the exercise of its constitutionally protected right of freedom of expression as enshrined in section 16 of the South African Bill of Rights.
3. Caxton is legally entitled to protect its various rights and is entitled to require, as of right, that its commercial arrangements with Google reflect the value Caxton's content unlawfully expropriated by Google and the content provided to Google by Caxton, under licence. Caxton seeks to protect, enforce and exercise such rights by, *inter alia*, agreeing on a fair remuneration model with Google, for the use of Caxton's proprietary news content ("**the**



Caxton Objective"). The Caxton Objective may be achieved by way of negotiation with and/or by way of legal proceedings against Google.

4. While Caxton remains committed to the Caxton Objective, an intractable obstacle persists. As Caxton had repeatedly emphasised, Google possesses information Caxton reasonably requires in order to formulate its commercial position in the enforcement of its rights and in any negotiations. Yet, despite Caxton's requests for access thereto, adequate information has not been forthcoming, which would render any negotiations to be one-sided, imbalanced and ultimately futile.
5. The abovementioned information asymmetry makes it impossible for the parties to proceed on an equal basis, which is a prerequisite for any commercial negotiation, or for the enforcement of rights and the formulation of claims for damages and or equitable licensing under copyright.

Contextual framework:

6. Google and Caxton are both stakeholders contributing to the South African Competition Commission's Media and Digital Platforms Market Inquiry ("**Inquiry**") initiated by the Competition Commission in terms of 43B(1)(a) of the Competition Act, 89 of 1998. The Inquiry is investigating whether there are any market features inherent to digital platforms markets that impede, distort, or restrict competition, prejudicing the interests of South African publishers.
7. News publishers (such as Caxton) have indicated that digital platforms (such as Google) derive value from its content, but that the information required to measure the value has not been made available by these platforms. Another submission made by news publishers is that the digital advertising market is distorted in that digital platforms have access to data that news publishers require to compete fairly in these markets. News publishers have also explained that digital platforms have become gatekeepers of information who facilitate how



and to whom news and commercial publisher content is distributed using methods and according to algorithms that have not been disclosed.

8. During the Inquiry process, Google has *inter alia* claimed that it does not gain any material value from Caxton's news content and that Google provides significant benefits to Caxton and other news publishers.
9. The information requested in this letter is required by Caxton to interrogate claims made by Google to the Inquiry. Google's claims made to the Inquiry affect not only Caxton, but the entire South African news industry. For Caxton and other media companies to respond to these claims, more information from Google is required, as set out below.
10. Caxton, as a member of the South African news industry, has the right to freedom of the press, and indeed is required to protect this right. Currently, this right is under threat due to, in our view, the distorted features of digital markets and Google plays a dominant role in these markets, which has caused a market failure and has resulted in inherent information asymmetry between market participants. These market distortions form the crux of the Inquiry's investigation, and it is Caxton's right and duty to gather all the information it can to better inform the Inquiry in its investigation.

Google's withholding of information prejudices Caxton's rights:

11. Caxton requires that Google must provide the information Caxton needs in order to pursue fair and reasonable compensation, and in order to sustain its constitutionally protected right to freedom of expression and the press, which it cannot exercise in the absence of financial accountability by Google. If Google provides the Inquiry with the information required by South African media at large, such information will not be accessible by market participants due to the confidentiality regime imposed by the Competition Act on which Google has an will indubitably rely, rendering such disclosure meaningless.
12. If any negotiations undertaken as part of the process adopted by Caxton for the achievement of the Caxton Objective fail or do not proceed on an equal basis, Caxton will require the information requested in this letter in order to formulate a legal claim against Google. Google



has publicly acknowledged¹ that it is subject legal claims of a similar nature to the claims Caxton may consider bring against Google. It is therefore reasonable for Caxton to consider formulating a legal claim against Google and requires the information requested in this letter to do so.

13. As this letter sets out above, the failure of digital platforms such as Google to provide the information requested in this letter imposes a limitation upon Caxton's right to freedom of expression – a right that Caxton exercises in order to participate in the markets which the Inquiry is investigating. Without freely exercising its Section 16 constitutional right, Caxton cannot compete freely in the market at hand.
14. Caxton's rights are prejudiced in that it does not have access to the information it needs in order to give effect to its rights under the Constitution or the Competition Act or the Copyright Act.
15. Section 32 of the South African Constitution and the Promotion of Access to Information Act entrenches Caxton's right of access to this information.
16. If Google fails to provide the information requested below, Caxton will proceed to seek the information by means of exercising the rights referred to in this letter.
17. On the above basis we require Google to provide the information requested in "**Annexure A**" of this letter.
18. Please provide the information requested in Annexure A by close of business on 1 March 2024.

Yours sincerely

Paul Jenkins
Non-executive Chairman
Caxton and CTP Publishers and Printers Limited
CTP Limited
Cell - 083 250 1000
Email – paul@jenkins.co.za

¹ At page 13 of Item 1A of the 2022 Annual Report it is stated that: *[Google is] subject to claims, lawsuits, regulatory and government investigations, other proceedings, and consent orders involving competition, intellectual property, data privacy... commercial disputes, content generated by our users, goods and services offered by advertisers or publishers using our platforms, personal injury, and other matters.*



Annexure A

Information requested by Caxton

1. What percentage of Google's gross annual revenue derived from all of Google's product offerings available to the South African market in the preceding ten financial years can be attributable to Caxton's news content?
2. How does Google determine its gross annual revenue derived from all of Google's product offerings available to the South African that could be attributable to Caxton?
3. Does Google have any mechanism for attributing indirect revenues for the use of Caxton's news content?
4. What percentage of Google's total annual global revenue is attributable to Caxton's news content in the preceding ten financial years?
5. How does Google determine the percentage of global revenue that is Caxton's news content?
6. Please provide all sources of Google's revenue that could be directly or indirectly attributable to Caxton either through direct or indirect advertising (or generated through any information collected or any use of any algorithm affected by any search or any individual).
7. What is Google's market share in online search in terms of advertising revenue and audience in South Africa?
8. How does Google determine its' market share in online search in terms of advertising revenue and audience in South Africa?
9. What is Google's revenue split between:
 - 9.1 advertising;
 - 9.2 app store commissions;
 - 9.3 subscriptions; and
 - 9.4 profit before and after tax?



10. What is the nature of inter-group charges and amounts involved, particularly between Alphabet Inc, Google Ireland Limited and Google LLC (or any of its worldwide subsidiaries) and Google South Africa Proprietary Limited?
11. How much has Google donated to South African political parties in the last ten years?
12. How much has Google's provided to South African government entities in the form of sponsorships in the last ten years?
13. How does Google determine the methodology for revenue sharing in respect of Google News Showcase?
14. It has been alleged by news publishers globally that Google's Bard AI was trained on a dataset that included news, magazine and digital publications.
 - 14.1 Was any of Caxton's journalistic output used in the training of Bard AI?
 - 14.2 Has Google ever used any data from Caxton or its users for the purposes of constructing any algorithm or including a search algorithm, Bard AI, rubric, global rate setting algorithm or any other algorithm employed in any business activity of Google, Alphabet or Jigsaw?
 - 14.3 How was the value of the inclusion of this data in any way assessed or included in the profits or revenues attributed to Caxton?
 - 14.4 How does Google determine what data is used to train Bard AI?
15. In Alphabet's 2022 annual report ("**2022 Annual Report**") it is stated at page 12 that: *"In the U.S., we launched a new feature to help readers discover local and regional news from different states so they could better understand election coverage of the races they most care about."*
 - 15.1 Why would Google launch new features that are reliant on journalist output if news was not important to Google's business model?
 - 15.2 What steps does Google take to ensure the content related to elections is unbiased and accurate?



16. At page 15 of the 2022 Annual Report it is stated that: *"People in affected areas turned to Search to find the latest news and information on where to find shelter and aid"*.
- 16.1 How would Google be in a position to provide users with these search results and services without the latest news and information being produced by news publishers and served on Google surfaces?
17. At page 31 of the 2022 Annual Report it is stated that: *"Communities around the world rely on local news to thrive and stay connected. More than 90% of the publications that are part of the Google News Showcase, our curated online news experience and licensing program, represent local, community, and regional news"*.
- 17.1 Does Google acknowledge the profound importance of local news to consumers?
- 17.2 How does Google determine the which news titles can be included in Google News Showcase?
- 17.3 Is Google inclusive of indigenous languages in its product partnerships?
18. We understand that Google collects users' data. It is unclear to Caxton what data Google collects on users that use Caxton's domains, as such we ask that Google provides the following information:
- 18.1 Does Google collect any user data from users while they are reading Caxton's journalistic output?
- 18.2 If so, what data does Google collect from users while they are reading Caxton's journalistic output?
- 18.3 Does Google use this user data to re-target audiences with advertising solutions?
- 18.4 Does Google collect any user data through any tech solution that Caxton might use e.g. GAM?
19. In Google's submission to the Inquiry dated 15 November 2023 ("**Google's Submission**") a number of claims are made by Google which Caxton requires clarity on:



- 19.1 At paragraph one of Google's Submission it is stated that "[w]hen the user is searching for news, our algorithms try to connect them with publishers of relevant, original, trustworthy and high quality journalism. In this way we drive traffic to publisher websites, helping to grow readership, and connect our users with relevant and useful information available on the web." (emphasis added). Please provide the following information:
- 19.1.1 How is Caxton's journalistic content affected by Google's algorithms, in terms of search results on the Google Search Engine Results Page?
- 19.1.2 Are Google's algorithms capable of identifying what content is news content?
- 19.1.3 How does Google's algorithms determine what content is news content?
- 19.1.4 How many referral clicks has Google directed toward Caxton's news sites in the past 12 months?
- 19.1.5 What are "free referral clicks" and how does Google determine this?
- 19.1.6 Has Google ever intentionally deprioritised Caxton news content on the Google Search Engine Results Page, if so why?
- 19.1.7 Does Google have any protocols for deprioritizing news in search results? What are the circumstances under which this is done?
- 19.1.8 How does Google determine whether news content should be deprioritized in search results?
- 19.1.9 What percentage of Google users conduct a search for news and merely review the results on the Google Search Engine Results Page and do not click through to a publisher's website? And how does Google determine this?
- 19.1.10 Does Google crawl Caxton's domains over time to scan Caxton's news sites in order to display recent content on Google Search Engine Results Page and Google News within seconds of publication?
- 19.1.11 What mechanisms does Google use to ensure the news it displays is "relevant, original, trustworthy and high-quality journalism"?



- 19.1.12 How does Google mitigate against the unlawful use of Caxton's news content (plagiarism) by unscrupulous news outlets?
20. At paragraph three of Google's Submission, it is stated that "*[w]e do not earn material revenue from news.*" On this point, we ask that Google informs us:
- 20.1 of how much direct revenue Google has made from "*news*" in the preceding decade?
- 20.2 How does Google determine direct revenue?
- 20.3 How does Google determine what is "*material revenue*"?
- 20.4 How does Google define "*news*"?
- 20.5 How does Google define "*news queries*"?
21. At paragraph four of Google's Submission, it is stated that "...*Google Search only makes advertising revenue when users click on ads.*", further to this at paragraph six Google claims that "*[s]earch queries with a news intent account for a very small portion of overall searches on Google Search, and are typically not commercialised.*" Lastly, at paragraph seven, Google states that "*Google Search does not earn or seek to earn meaningful advertising revenue from displaying news results on the Google Search Engine Results Page.*" This seems to be in stark contrast with the statement made in Google Inc's 2010 Annual Report: "*[Google] believe[s] the factors that influence the success of [it's] advertising programs include the following: The relevance, objectivity, and quality of our search results and the relevance and quality of ads displayed with each search results page number of searches initiated at our websites and our Google Network members' websites and the underlying purpose of these searches (for instance, whether they are for academic research, to find a news article, or to find a product or service).* (emphasis added). Based on these comments, we ask that Google provides the following information:
- 21.1 Does Google derive any benefit from Caxton's journalistic output being available within Google's ecosystem / on Google surfaces?
- 21.2 How does Google define "*news searches*"?
- 21.3 What percent of Google search results can be classified as news searches and how does Google determine this?



- 21.4 Does Google consider Caxton's journalistic output to be content that is trustworthy?
- 21.5 How does Google determine whether Caxton's content is trustworthy and does Google use Caxton data to test if other sources are trustworthy in any way?
- 21.6 To the extent that Google labels content as trustworthy is this done using any of Caxton's data?
- 21.7 Would it be reasonable to conclude that the generally high level of trust that people have in news has a positive effect on Google's brand equity as a platform and thus creates an advantage for Google?
- 21.8 Does Google consider Caxton's journalistic output to be quality content that is relevant to South African's searching for South African information and news?
- 21.9 Does Google consider Caxton's journalistic content to be up-to-date and published on a regular basis?
- 21.10 What percentage of searches made on Google in South Africa could be classified as "information searches"?²
- 21.11 What percentage of South African Google search results have included Caxton's journalistic content as part of the results on a Google Search Engine Results Page?
- 21.12 How does Google determine what news item or what item people are looking for at a search?
- 21.13 It could be argued that Google is able to achieve power and scale through network effects by connecting producers and consumers. Further, that Google benefits from network effects by attracting more users and advertisers to its platform, which in turn increases the supply and demand and sets a positive cycle in motion. To this end, the more users that use Google, the more data it can collect, and analyse, to improve its search results and advertising campaigns. This in turn leads to a better experience for users and a higher ROI for advertisers, which sequentially attracts more users and

² An "information search" describes a search which is used to find information about a certain topic and would exclude commercial searches and navigation searches.



advertisers and strengthens the cycle. Does Google disagree with any of these statements?

- 21.14 How many keywords have been sold in the proceeding financial year in South Africa?
- 21.15 How much revenue did keywords contribute to the gross revenue generated in South Africa in the proceeding financial year?