

Scores for Google on the 2023 Foundation Model Transparency Index

Background

1. Please see the paper describing the Foundation Model Transparency Index in order to understand what this document includes. The paper provides necessary background on (i) what these indicators are and why they were chosen, (ii) our standardized process for scoring the transparency of foundation model developers, and (iii) what these scores mean in context.
2. This document contains only information that was publicly available before September 15, 2023. It has not been updated and should be interpreted as a snapshot of transparency as of September 15, 2023.
3. In order to assess the transparency of foundation model developers, we used a rigorous, standardized [search protocol](#) to find publicly available information related to these indicators. You can find more information about this search protocol in the paper describing the Foundation Model Transparency Index.
4. We evaluate every company in this same way—you can find scoring documents for the other companies [here](#).
5. We evaluate each company on 100 indicators of transparency. You can find the definition of each indicator and additional information about how each indicator was scored [here](#).
6. Scores for each indicator are either 0 or 1. If the score is a 0, we do not provide a source for the score because our standardized search protocol (which includes many relevant sources) did not yield enough information to award a point. If the score is a 1, we provide a source that includes the information we cite in the justification for the score.
7. We evaluate each company on the basis of its flagship foundation model; in the case of Google, we evaluate PaLM 2.
8. In advance of releasing the Foundation Model Transparency Index, we reached out to Google for comment (along with the 9 other companies we evaluated) and offered an opportunity to provide feedback on the index and the organization's scores.

Scores for Each Indicator

1. Upstream → Data → Data Size
 - Score: 0
 - Justification: No information found related to the size of the data.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
2. Upstream → Data → Data Sources
 - Score: 0
 - Justification: While Google provides some information on PaLM 2 data sources in the PaLM 2 Technical Report, it is too generic and high-level to warrant the point. The sources must be named/identifiable at some lower-level.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
3. Upstream → Data → Data Creators
 - Score: 0
 - Justification: While there is an analysis of how people are represented within the data in the PaLM 2 Technical Report, there is no discussion of who produced the data.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
4. Upstream → Data → Data Source Selection
 - Score: 0
 - Justification: While there is some discussion of why multilingual data was included in the PaLM 2 Technical Report, there is no discussion of why the majority of sources (e.g. books, code) were included, nor why others were excluded.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

5. Upstream → Data → Data Curation
 - Score: 1
 - Justification: Section 3 of the PaLM 2 Technical Report states “We employed several data cleaning and quality filtering methods, including de-duplication, removal of sensitive-PII and filtering. Even though PaLM 2 has a smaller proportion of English data than PaLM, we still observe significant improvements on English evaluation datasets, as described in Section 4. We attribute this partially to the higher data quality in the PaLM 2 mixture.” While this is not specific to particular data sources, as written it appears this was done across all pre-training data independent of underlying source.
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
6. Upstream → Data → Data Augmentation
 - Score: 1
 - Justification: Data is augmented with control tokens that indicate toxicity of text using PerspectiveAPI.
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
7. Upstream → Data → Harmful Data Filtration
 - Score: 0
 - Justification: While filters are described in the PaLM 2 Technical Report, how they relate to harmful data and details on the filter are not provided. Note: While not relevant for this indicator, to avoid potential confusion, we highlight the related use of Perspective API for tagging (as opposed to filtering) text with toxicity-related metadata, including the fixed version of the API.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
8. Upstream → Data → Copyrighted data
 - Score: 0
 - Justification: No information found regarding the copyright status of the data used to build PaLM 2.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
9. Upstream → Data → Data License
 - Score: 0
 - Justification: No information found regarding the license status of the data used to build PaLM 2.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

10. Upstream → Data → Personal Information in Data

- Score: 0
- Justification: No information found related to the inclusion or exclusion of personal information in specific parts of the data.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

11. Upstream → Data Labor → Use of Human Labor

- Score: 0
- Justification: Google discloses that it uses human labor for some stages of the pipeline in the PaLM 2 Technical Report, but it is unclear if it is involved in building pretraining data, data filtration, and instruction-tuning.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

12. Upstream → Data Labor → Employment of Data Laborers

- Score: 0
- Justification: Employer for data laborers is disclosed only for a few stages in the pipeline, e.g. pp 84 in the PaLM 2 Technical Report.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

13. Upstream → Data Labor → Geographic Distribution of Data Laborers

- Score: 0
- Justification: No information found about geographic distribution of data laborers.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

14. Upstream → Data Labor → Wages

- Score: 0
- Justification: No information found about the wages for people who perform data labor.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

15. Upstream → Data Labor → Instructions For Creating Data
- Score: 0
 - Justification: Instructions for two evaluation datasets are disclosed, but this does not apply to all data phases. The PaLM 2 Technical Report reads "We include two CrowdWorksheets (Diaz et al., 2022) describing two evaluation datasets used in this work, as a step towards standardizing this transparency practice for all evaluation datasets created with data collaborators."
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
16. Upstream → Data Labor → Labor Protections
- Score: 0
 - Justification: No information found about labor protections for data laborers.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
17. Upstream → Data Labor → Third Party Partners
- Score: 0
 - Justification: No information found about any third party partners involved in the development of PaLM 2.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
18. Upstream → Data Access → Queryable External Data Access
- Score: 0
 - Justification: No information found about whether external entities are provided with queryable access to the data used to build PaLM 2.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
19. Upstream → Data Access → Direct External Data Access
- Score: 0
 - Justification: No information found about whether external entities are provided with direct access to the data used to build PaLM 2.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

20. Upstream → Compute → Compute Usage

- Score: 0
- Justification: The PaLM 2 Model Card states “Compute Requirements: Not Reported”.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

21. Upstream → Compute → Development Duration

- Score: 0
- Justification: No information found about the time required to build PaLM 2.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

22. Upstream → Compute → Compute Hardware

- Score: 0
- Justification: The hardware type is disclosed as TPU v4 in the PaLM 2 Model Card, but the number of TPUs is not disclosed.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

23. Upstream → Compute → Hardware Owner

- Score: 1
- Justification: The PaLM 2 model card discloses that it was trained on TPU v4, a proprietary Google technology. The model card then links to a paper on “A domain-specific supercomputer for training deep neural networks” owned by Google.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

24. Upstream → Compute → Energy Usage

- Score: 0
- Justification: No information found about the amount of energy expended in building the model.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

25. Upstream → Compute → Carbon Emissions

- Score: 0
- Justification: No information found about the amount of carbon emitted in building the model.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

26. Upstream → Compute → Broader Environmental Impact
- Score: 0
 - Justification: No information found any broader environmental impacts from building the model.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
27. Upstream → Methods → Model Stages
- Score: 1
 - Justification: The sole model stage in the training of PaLM 2 is a pretraining step, which includes training with toxicity control tokens. There is also subsequent fine-tuning on the Flan dataset, as described in Appendix A.2, but this is for a separate artifact (i.e. Flan-PaLM 2 (L)).
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
28. Upstream → Methods → Model Objectives
- Score: 0
 - Justification: The PaLM 2 Technical Report describes that objectives are similar to UL2 of Tay et al., but the exact mixture or even confirmation that the underlying objectives are those described in UL2 is not provided.
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
29. Upstream → Methods → Core Frameworks
- Score: 1
 - Justification: Pathways, Paxml, JAX, GSPMD are disclosed as core frameworks in the PaLM 2 Model Card.
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
30. Upstream → Methods → Additional Dependencies
- Score: 1
 - Justification: The PaLM 2 Model Card states “This is a static model trained on an offline dataset,” indicating that there are no additional dependencies.
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>

31. Upstream → Data Mitigations → Mitigations for Personally Identifiable Information
- Score: 0
 - Justification: Google makes some minimal acknowledgement in the PaLM 2 Technical Report that it removes sensitive PII, but provides no details about the steps or what it precisely means by this.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
32. Upstream → Data Mitigations → Mitigations for Copyright
- Score: 0
 - Justification: No information found about steps Google takes to mitigate the presence of copyrighted information in the data.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
33. Model → Model Basics → Input Modality
- Score: 1
 - Justification: The PaLM 2 Model Card states “The model takes text as input”
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
34. Model → Model Basics → Output Modality
- Score: 1
 - Justification: The PaLM 2 Model Card states “The model generates text as output”
 - Source: <https://arxiv.org/pdf/2305.10403.pdf>
35. Model → Model Basics → Model Components
- Score: 0
 - Justification: The PaLM 2 Model Card states “We have small, medium, and large variants that use stacked layers based on the Transformer architecture, with varying parameters depending on model size. Further details of model size and architecture are withheld from external publication.”
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

36. Model → Model Basics → Model Size

- Score: 0
- Justification: The PaLM 2 Model Card states “We have small, medium, and large variants that use stacked layers based on the Transformer architecture, with varying parameters depending on model size. Further details of model size and architecture are withheld from external publication.”
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

37. Model → Model Basics → Model Architecture

- Score: 1
- Justification: Google discloses the model architecture is a Transformer in the PaLM 2 Technical Report, though with no additional details.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

38. Model → Model Basics → Centralized Model Documentation

- Score: 1
- Justification: The PaLM 2 Technical Report centralizes key information about the model.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

39. Model → Model Access → External Model Access Protocol

- Score: 0
- Justification: The PaLM 2 model access form does not specify selection criteria or if/when a decision will be provided.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

40. Model → Model Access → Black Box External Model Access

- Score: 1
- Justification: PaLM 2 is available via an API, meaning that there is black box access for external entities.
- Source: <http://web.archive.org/web/20230914164159/https://developers.generativeai.google/products/palm>

41. Model → Model Access → Full External Model Access

- Score: 0
- Justification: Full access to the model is not provided to external entities via model weights.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

42. Model → Capabilities → Capabilities Description

- Score: 1
- Justification: PaLM 2 “excels at advanced reasoning tasks, including code and math, classification and question answering, translation and multilingual proficiency, and natural language generation better than our previous state-of-the-art LLMs, including PaLM. It can accomplish these tasks because of the way it was built – bringing together compute-optimal scaling, an improved dataset mixture, and model architecture improvements.”
- Source:
<http://web.archive.org/web/20230914164159/https://blog.google/technology/ai/google-palm-2-ai-large-language-model/>

43. Model → Capabilities → Capabilities Demonstration

- Score: 1
- Justification: Several demonstrations of capabilities across the PaLM 2 release page and PaLM 2 Technical Report.
- Source:
<http://web.archive.org/web/20230914164159/https://blog.google/technology/ai/google-palm-2-ai-large-language-model/>

44. Model → Capabilities → Evaluation of Capabilities

- Score: 1
- Justification: Extensive evaluation of PaLM 2 capabilities in Section 4 of PaLM 2 Technical Report.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

45. Model → Capabilities → External Reproducibility of Capabilities Evaluation

- Score: 1
- Justification: Many evaluations on standard public benchmarks.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

46. Model → Capabilities → Third Party Capabilities Evaluation

- Score: 0
- Justification: No information found indicating that PaLM 2's capabilities have been evaluated by third parties.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

47. Model → Limitations → Limitations Description

- Score: 1
- Justification: The PaLM 2 Technical Report states “Limitations: PaLM 2 is not tested in this paper in settings outside of research that can affect performance, and it should not be used for downstream applications without further analysis of potential harm or bias in the proposed downstream application.”
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

48. Model → Limitations → Limitations Demonstration

- Score: 0
- Justification: No information found related to illustrative examples of limitations.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

49. Model → Limitations → Third-Party Evaluation of Limitations

- Score: 1
- Justification: Limitations can be evaluated via model access; no provisions of the prohibited use policy appear to constrain this.
- Source: <https://arxiv.org/pdf/2305.10403.pdf> and <https://web.archive.org/web/20230914001155/https://policies.google.com/terms/generative-ai/use-policy>

50. Model → Risks → Risks Description

- Score: 0
- Justification: While risks are described obliquely in the PaLM 2 Technical Report, no information related to intentional harm is included in the description, which instead centers on representational harms.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

51. Model → Risks → Risks Demonstration

- Score: 0
- Justification: There is some textual description of risks in Appendix E.2 of the technical report but no explicit/easily understandable demonstration of risks described (e.g. representational harms like toxic language and social stereotypes).
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

52. Model → Risks → Unintentional Harm Evaluation

- Score: 1
- Justification: There are evaluations for unintentional harms such as evaluations for rate of toxic continuation in Figure 31 across languages or translation misgendering in Table 24.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

53. Model → Risks → External Reproducibility of Unintentional Harm Evaluation

- Score: 1
- Justification: The toxicity evaluations and translation evaluations are reproducible, involving public datasets like ParLAI with stated hyperparameters for sampling and the Perspective API for toxicity classification with a fixed version of the Perspective API.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

54. Model → Risks → Intentional Harm Evaluation

- Score: 0
- Justification: There are no evaluations clearly directed towards intentional harms. The evaluation that comes closest to satisfying this indicator is the adversarial aspects of ParLAI, but there are not multiple evaluations.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

55. Model → Risks → External Reproducibility of Intentional Harm Evaluation

- Score: 0
- Justification: No information found related to the reproducibility of intentional harm evaluations as there is no information found about the evaluations themselves.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

56. Model → Risks → Third-Party Risk Evaluation

- Score: 0
- Justification: No information found indicating that third parties have evaluated the risks of PaLM 2.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

57. Model → Mitigations → Mitigations Description

- Score: 1
- Justification: Page 9 of the PaLM 2 Technical Report reads "For a small fraction of pre-training data, we added special control tokens marking the toxicity of text, using signals from a fixed version of the Perspective API." It is unclear if many other mitigations are in fact implemented or are just presented as options for deployers.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

58. Model → Model Mitigations → Mitigations Demonstration

- Score: 0
- Justification: No information found related to illustrative examples of mitigations.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

59. Model → Model Mitigations → Mitigations Evaluation

- Score: 1
- Justification: Several mitigations are evaluated. It is unclear the extent to which these mitigations are implemented or are just presented as options for deployers, but for this version of the Index, we take this as sufficient to award the indicator.
- Source: <https://arxiv.org/pdf/2305.10403.pdf>

60. Model → Model Mitigations → External Reproducibility of Mitigations Evaluation

- Score: 0
- Justification: No information found to indicate mitigations evaluations are externally reproducible.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

61. Model → Model Mitigations → Third Party Mitigations Evaluation

- Score: 0
- Justification: No information found to indicate mitigations can be evaluated by third parties.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

62. Model → Trustworthiness → Trustworthiness Evaluation

- Score: 0
- Justification: No information found related to evaluations of robustness, reliability, hallucinations, uncertainty, calibration, causality, interpretability, or explainability.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

63. Model → Trustworthiness → External Reproducibility of Trustworthiness Evaluation

- Score: 0
- Justification: No information found related to trustworthiness evaluations, and so no information was found about their reproducibility.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

64. Model → Inference → Inference Duration Evaluation

- Score: 0
- Justification: No information found related to the time required for model inference.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

65. Model → Inference → Inference Compute Evaluation

- Score: 0
- Justification: No information found related to compute usage for model inference.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

66. Downstream → Distribution → Release Decision-Making

- Score: 0
- Justification: No information found about Google's protocol for deciding to release PaLM 2.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

67. Downstream → Distribution → Release Process

- Score: 1
- Justification: There is some information about PaLM 2 release being incremental (e.g. announcement of API), which is sufficient for this version of the Index, though key points in the process are vaguely articulated.
- Source:
<https://web.archive.org/web/20230914000505/https://developers.generativeai.google/products/palm>

68. Downstream → Distribution → Distribution Channels

- Score: 1
- Justification: The PaLM API is disclosed as the distribution channel.
- Source:
<https://web.archive.org/web/20230914000505/https://developers.generativeai.google/products/palm>

69. Downstream → Distribution → Products and Services

- Score: 1
- Justification: MakerSuite and VertexAI depend on PaLM 2: “At I/O today, we announced over 25 new products and features powered by PaLM 2. That means that PaLM 2 is bringing the latest in advanced AI capabilities directly into our products and to people — including consumers, developers, and enterprises of all sizes around the world. Here are some examples...”
- Source:
<https://web.archive.org/web/20230824173021/https://developers.generativeai.google/products/palm> and
<https://web.archive.org/web/20230824173239/https://blog.google/technology/ai/google-palm-2-ai-large-language-model/>

70. Downstream → Distribution → Detection of Machine-Generated Content

- Score: 0
- Justification: No information found related to a mechanism for the detection of content generated by PaLM 2.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

71. Downstream → Distribution → Model License

- Score: 1
- Justification: The additional TOS for the PaLM API includes a content license, and together they approximate a license for PaLM 2.
- Source: <https://web.archive.org/web/20230914000355/https://developers.generativeai.google/terms>

72. Downstream → Distribution → Terms of Service

- Score: 1
- Justification: The PaLM API points to Google's broad TOS, and there is also an additional TOS for PaLM API.
- Source: <https://web.archive.org/web/20230914001211/https://policies.google.com/terms#toc-intro> and <https://web.archive.org/web/20230914000355/https://developers.generativeai.google/terms>

73. Downstream → Usage Policy → Permitted, Restricted, and Prohibited Users

- Score: 1
- Justification: Additional terms of service for the PaLM API lay out geographic restrictions.
- Source: https://web.archive.org/web/20230914000546/https://developers.generativeai.google/available_regions and <https://web.archive.org/web/20230914000355/https://developers.generativeai.google/terms>

74. Downstream → Usage Policy → Permitted, Restricted, and Prohibited Uses

- Score: 1
- Justification: Prohibited Use policy outlines tens of prohibited uses, TOS for PaLM API lays out use restrictions, model card describes permitted uses.
- Source: <https://ai.google/static/documents/palm2techreport.pdf> and <https://web.archive.org/web/20230914001155/https://policies.google.com/terms/generative-ai/use-policy> and <https://web.archive.org/web/20230914000355/https://developers.generativeai.google/terms>

75. Downstream → Usage Policy → Usage Policy Enforcement

- Score: 0
- Justification: Information found indicates Google lays out precautionary measures, not a protocol regarding what happens in the event of a usage policy violation. Additional API TOS reads "The APIs include safety features to block harmful content, such as content that violates our Prohibited Use Policy. You may not attempt to bypass these protective measures or use content that violates the API Terms or these Additional Terms." Safety Settings for PaLM API gives more detail: "In addition to the adjustable safety filters, the PaLM API has built-in protections against core harms, such as content that endangers child safety. These types of harm are always blocked and cannot be adjusted."
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

76. Downstream → Usage Policy → Justification for Enforcement Action

- Score: 1
- Justification: Safety settings for PaLM API read: "If content was blocked, the response from the API contains the reason it was blocked in the ContentFilter.reason field. If the reason was related to safety, then the response also contains a SafetyFeedback field which includes the safety settings that were used for that request as well as a safety rating. The safety rating includes the category and the probability of the harm classification. The content that was blocked is not returned."
- Source:
https://web.archive.org/web/20230914000336/https://developers.generativeai.google/guide/safety_setting

77. Downstream → Usage Policy → Usage Policy Violation Appeals Mechanism

- Score: 0
- Justification: No information found related to a mechanism for appealing the developer's decision that the usage policy has been violated.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

78. Downstream → Model Behavior Policy → Permitted, Restricted, and Prohibited Model Behaviors

- Score: 0
- Justification: No information found related to model behaviors that are permitted, restricted, and prohibited.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

79. Downstream → Model Behavior Policy → Model Behavior Policy Enforcement
- Score: 0
 - Justification: No information found related to the enforcement protocol for the model behavior policy.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
80. Downstream → Model Behavior Policy → Interoperability of Usage and Model Behavior Policies
- Score: 0
 - Justification: No information found about how the usage policy and the model behavior policy interoperate.
 - Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.
81. Downstream → User Interface → User Interaction with AI System
- Score: 1
 - Justification: Tab on model settings in MakerSuite makes clear which FM is being distributed.
 - Source: https://makersuite.google.com/app/prompts/new_text
82. Downstream → User Interface → Usage Disclaimers
- Score: 1
 - Justification: Signing up with API points to the terms of service which repeatedly point to the prohibited use policy.
 - Source: <https://web.archive.org/web/20230914000505/https://developers.google.com/ai/policies/palm>
83. Downstream → User Data Protection → User Data Protection Policy
- Score: 1
 - Justification: Privacy policy addresses data access, storage, and sharing.
 - Source: <https://web.archive.org/web/20230913002308/https://policies.google.com/privacy>

84. Downstream → User Data Protection → Permitted and Prohibited Use of User Data

- Score: 1
- Justification: Permitted uses described in US state law requirements and prohibited uses implied by “when google shares your information” section.
- Source:
<https://web.archive.org/web/20230913002308/https://policies.google.com/privacy>

85. Downstream → User Data Protection → Usage Data Access Protocol

- Score: 0
- Justification: No information found related to a protocol for granting external entities access to usage data.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

86. Downstream → Model Updates → Versioning Protocol

- Score: 1
- Justification: Versioned models (versions of bison and gecko) are provided in the API documentation.
- Source:
<https://web.archive.org/web/20230901181431/https://developers.generativeai.google/models/language>

87. Downstream → Model Updates → Change Log

- Score: 1
- Justification: Google provides a detailed changelog for features in their API and models.
- Source:
<https://web.archive.org/web/20230913235653/https://cloud.google.com/vertex-ai/docs/generative-ai/release-notes>

88. Downstream → Model Updates → Deprecation Policy

- Score: 1
- Justification: Google's model versioning page says each stable version of the model is available for 6 months after the release of the next stable version. It also specifies how to use the model to always point to the latest version.
- Source:
<https://web.archive.org/web/20230913235326/https://cloud.google.com/vertex-ai/docs/generative-ai/learn/model-versioning>

89. Downstream → Feedback → Feedback Mechanism

- Score: 1
- Justification: Google asks developers to provide feedback on their discord channel: "Please keep sharing your feedback on the #palm-api channel on the Google Developer Discord."
- Source: <https://web.archive.org/web/20230913235115/https://developers.googleblog.com/2023/05/palm-api-and-makersuite-moving-into-public-preview.html>

90. Downstream → Feedback → Feedback Summary

- Score: 0
- Justification: No information found related to any summary of user feedback.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

91. Downstream → Feedback → Government Inquiries

- Score: 0
- Justification: No information found in connection with a summary of government inquiries related to the model.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

92. Downstream → Impact → Monitoring Mechanism

- Score: 1
- Justification: The PaLM 2 technical report states "Mitigations are recommended to address any vulnerabilities, and rigorous testing and production monitoring is performed to identify any concerns or emerging trends." Google's API terms of service states "YOU AGREE THAT GOOGLE MAY MONITOR USE OF THE APIS TO ENSURE QUALITY, IMPROVE GOOGLE PRODUCTS AND SERVICES, AND VERIFY YOUR COMPLIANCE WITH THE TERMS. This monitoring may include Google accessing and using your API Client, for example to identify security issues that could affect Google or its users. You will not interfere with this monitoring. Google may use any technical means to overcome such interference. Google may suspend access to the APIs by you or your API Client without notice if we reasonably believe that you are in violation of the Terms."
- Source: <https://ai.google/static/documents/palm2techreport.pdf> and <https://developers.google.com/terms>

93. Downstream → Impact → Downstream Applications

- Score: 0
- Justification: No information found related to the number of applications dependent on the foundation model.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

94. Downstream → Impact → Affected Market Sectors

- Score: 0
- Justification: No information found related to the fraction of applications corresponding to each market sector.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

95. Downstream → Impact → Affected Individuals

- Score: 0
- Justification: No information found related to the number of individuals affected by the model.
- Source: Search protocol produced no source containing sufficient information to award this point

96. Downstream → Impact → Usage Reports

- Score: 0
- Justification: No information found related to usage statistics describing the impact of the model on users.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

97. Downstream → Impact → Geographic Statistics

- Score: 0
- Justification: No information found regarding statistics of model usage across geographies.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

98. Downstream → Impact → Redress Mechanism

- Score: 0
- Justification: No information found regarding any mechanism to provide redress to users for harm caused by the model.
- Source: No source is provided because the search protocol did not produce a source containing sufficient information to award this point.

99. Downstream → Documentation for Deployers → Centralized Documentation for Downstream Use

- Score: 1
- Justification: The “generative AI for developers” guide centralizes documentation for downstream use.
- Source:
<https://web.archive.org/web/20230913234933/https://developers.generativeai.google/guide>

100. Downstream → Documentation for Deployers → Documentation for Responsible Downstream Use

- Score: 1
- Justification: PaLM 2 Technical Report links to responsible use guidelines.
- Source:
<https://web.archive.org/web/20230913234829/https://ai.google/responsibility/responsible-ai-practices/>