**AI21 Labs (Jurassic-2) - total: 10/48**

1. Data sources
   * 0 points: No description is provided of the Jurassic-2 training data beyond “massive corpus of texT”. AI21 Labs does provide a paper and very brief description for Jurassic-1 however.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
2. Data governance
   * 0 points: No discussion of data governance.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
3. Copyrighted data
   * 0 points: No description of copyrighted data.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
4. Compute (additive)
   * 0 points: The model size, training time, hardware, and FLOPs are not disclosed.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
5. Energy (additive)
   * 0 points: The energy usage, carbon emissions, and any form of measurement/mitigation are not disclosed.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
6. Capabilities and limitations
   * 2 points: Capabilities are described as extensive (“J2 models can be applied to virtually any language task by crafting a suitable prompt”), with several more specific use cases (e.g. generating marketing copy, powering chatbots and assisting creative writing). By comparison, limitations are not well-described, though some discussion appears in the responsible use documentation.
   * Sources: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models> and <https://docs.ai21.com/docs/responsible-use/>.
7. Risks and mitigations (additive)
   * 2 points: Risks are described very generically (“potential risks, including malicious use by bad actors, accidental misuse and broader societal harms”) with more specific prohibited uses of “Illegal activities, such as hate speech, gambling, child pornography or violating intellectual property rights; Harassment, victimization, intimidation, fraud or spam; Creation or dissemination of misinformation, promotion of self-harm, glorification of violent events or incitement of violence”. Mitigations include no fully-automated (without human intervention) posting of generated content to > 100 people, and harm reporting obligations for users of the model (not AI21 Labs) if the downstream application has 100+ monthly users.
   * Sources: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/safety-research> and <https://docs.ai21.com/docs/responsible-use/>.
8. Evaluations (additive)
   * 0 points: No evaluations reported.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/jurassic-2-models>
9. Testing (additive)
   * 1 point: HELM evaluation cited in press release, though no further discussion of internal testing or external redteaming.
   * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs> as well as press release at <https://www.ai21.com/blog/introducing-j2>.
10. Machine-generated content
    * 3 points: In the context of the playground, it is clear that content is machine-generated. Beyond this context, in using the AI21 models and APIs, the responsible use guidelines require downstream use to disclose content is machine-generated: “In any case, the first human to view text generated by AI21 Studio must not be led to believe that it was written by a human.” However, note these requirements are not clearly enforced and only pertain to the *first* human.
    * Sources: Jurassic-2 documentation on <https://docs.ai21.com/docs> including <https://docs.ai21.com/docs/responsible-use#usage-guidelines>.
11. Member states
    * 2 points: Organization-level restrictions are only imposed on US Export Control nations (Lebanon, Syria, Iran, Iraq, Sudan, Yemen, Cuba, or North Korea), with no specific discussion of how the models are deployed across different EU member states.
    * Sources: <https://studio.ai21.com/terms-of-use>
12. Downstream documentation
    * 0 points: No discussion of downstream obligations and no standardized public documentation.
    * Source: Jurassic-2 documentation on <https://docs.ai21.com/docs>