QUICKSCAN CANVAS GoldenFossils: Advancing Vision-Language Fossil Identification through Fine-Grained Annotations

NAME: GoldenFossils: Advancing Vision-Language Fossil Identification through Fine-Grained Annotations

DATE: June 26, 2026 10:54 AM

DESCRIPTION OF TECHNOLOGY

I designed and implemented a new Annotation Framework for Fossil Identification and propose a new machine-actionable Dataset. By leveraging multimodal Vision-Language Models (VLMs) as high-speed "digital interns", the application handles the heavy lifting of scientific pre-filling, while empowering experts to steer the final narrative through precision image markers and conversational.

HUMAN VALUES

For domain experts, their identity shifts from being primary "generators of knowledge" to "auditors of AI," which potentially changes the nature of paleontological expertise toward system supervision. For citizen scientists, the technology elevates their role from simple hobbyists to valued contributors of high-quality "scientific training signals" used to improve global AI models.

TRANSPARENCY

The framework includes a "Welcome Disclaimer" and a "Privacy Shield" that explain the "Human-in-the-Loop" workflow and how data is used to fine-tune the AI. As an institutional research prototype, the "business model" is centered on contributing structured data to the museum's open-research database rather than commercial profit.

IMPACT ON SOCIETY

Documenting fossil remains is labor-intensive and results in unstructured data that lacks the morphological detail needed for modern AI training. This is a issue because it creates massive backlogs in museums and hinders the development of reliable automated identification tools.

STAKEHOLDERS

- Paleontologists
- Student
- Company Mentor
- Citizen Scientists
- Naturalis Biodiversity Center

SUSTAINABILITY

While not the primary focus of the study, the research acknowledges the significant computational energy required for running high-level multimodal models like Gemini. The use of a "Flash" model variant represents a conscious choice to utilize a more energy-efficient, "lite" version of the AI compared to the much heavier and more resource-intensive "Pro" models.

HATEFUL AND CRIMINAL ACTORS

The framework could potentially be misused to generate plausible-sounding scientific justifications for illegally excavated or trafficked fossils to make them appear legitimate for sale. Additionally, someone could use the AI to falsify the stratigraphic context or discovery location of a find to circumvent national heritage laws regarding "in-situ" archaeological discoveries.

DATA

The research explicitly acknowledges the "taphonomic filter" and "morphological convergence," which are major pitfalls where data can be misleading or "noisy." The framework addresses these shortcomings by requiring coordinate-based annotations (Segment 4), ensuring that conclusions are pinned to specific visual evidence rather than relying on the AIs often unreliable general intuition.

FUTURE

In the future, this framework could lead to the development of a specialized "Paleo-VLM" capable of identifying fossils with near-human accuracy across a global network of museums. It could also become a real-time bridge where a citizen scientist's phone upload is automatically processed, expert-verified, and archived into a museum database within minutes.

PRIVACY

Yes, the technology registers institutional login credentials, including usernames and email addresses, to manage expert authentication via Supabase. It also processes data linked to citizen scientist accounts from the oervondstchecker.nl platform, which may include the names or IDs of the original discoverers.

INCLUSIVITY

Yes, the benchmarking identified a significant "Megafauna Bias," where the underlying VLMs disproportionately favor high-prestige Pleistocene species like mammoths over common domestic animals. This bias causes the models to misidentify modern specimens as ancient fossils simply because they were told they were looking at data from an "ice age" context.

FIND US ON WWW.TICT.IO

THIS CANVAS IS PART OF THE TECHNOLOGY IMPACT CYCLE TOOL. THIS CANVAS IS THE RESULT OF A QUICKSCAN. YOU CAN FILL OUT THE FULL TICT ON WWW.TICT.IO

QUICKSCAN CANVAS ADVANCING Vision-Language Fossil Identification through Fine-Grained Annotations

NAME: GoldenFossils: Advancing Vision-Language Fossil Identification through Fine-Grained Annotations
DATE: June 26, 2026 10:54 AM

DESCRIPTION OF TECHNOLOGY
I designed and implemented a new Annotation Framework for Fossil Identification and propose a new machine-actionable Dataset. By leveraging multimodal Vision-Language Models (VLMs) as high-speed "digital interns", the application handles the heavy lifting of scientific pre-filling, while empowering experts to steer the final narrative through precision image markers and conversational.

HUMAN VALUES

How is the identity of the (intended) users affected by the technology?

To help you answer this question think about sub questions like:

- If two friends use your product, how could it enhance or detract from their relationship?
- Does your product create new ways for people to interact?…

TRANSPARENCY

Is it explained to the users/stakeholders how the technology works and how the business model works?

- Is it easy for users to find out how the technology works?
- Can a user understand or find out why your technology behaves in a certain way?
- Are the goals explained?
- Is the idea of the technology explained?
- Is the technology company transparent about the way their…

IMPACT ON SOCIETY

What is exactly the problem? Is it really a problem? Are you sure?

Can you exactly define what the challenge is? What problem (what 'pain') does this technology want to solve? Can you make a clear definition of the problem? What 'pain' does this technology want to ease? Whose pain? Is it really a problem? For who? Will solving the problem make the world better? Are you sure? The problem definition will help you to determine…

STAKEHOLDERS

Who are the main users/targetgroups/stakeholders for this technology? Think about the intended context by…

When thinking about the stakeholders, the most obvious one are of course the intended users, so start there. Next, list the stakeholders that are directly affected. Listing the users and directly affected stakeholders also gives an impression of the intended context of the technology.
…

SUSTAINABILITY

In what way is the direct and indirect energy use of this technology taken into account?

One of the most prominent impacts on sustainability is energy efficiency. Consider what service you want this technology to provide and how this could be achieved with a minimal use of energy. Are improvements possible?

HATEFUL AND CRIMINAL ACTORS

In which way can the technology be used to break the law or avoid the consequences of breaking the law?

Can you imagine ways that the technology can or will be used to break the law? Think about invading someone's privacy. Spying. Hurting people. Harassment. Steal things. Fraud/ identity theft and so on. Or will people use the technology to avoid facing the consequences of breaking the law (using trackers to evade speed radars or using bitcoins to launder…

DATA

Are you familiar with the fundamental shortcomings and pitfalls of data and do you take this sufficiently into…

There are fundamental issues with data. For example:

- Data is always subjective;
- Data collections are never complete;
- Correlation and causation are tricky concepts;
- Data collections are often biased;…

FUTURE

What could possibly happen with this technology in the future?

Discuss this quickly and note your first thoughts here. Think about what happens when 100 million people use your product. How could communities, habits and norms change?

PRIVACY

Does the technology register personal data? If yes, what personal data?

If this technology registers personal data you have to be aware of privacy legislation and the concept of privacy. Think hard about this question. Remember: personal data can be interpreted in a broad way. Maybe this technology does not collect personal data, but can be used to assemble personal data. If the technology collects special personal data (like…

INCLUSIVITY

Does this technology have a built-in bias?

Do a brainstorm. Can you find a built-in bias in this technology? Maybe because of the way the data was collected, either by personal bias, historical bias, political bias or a lack of diversity in the people responsible for the design of the technology? How do you know this is not the case? Be critical. Be aware of your own biases.…

FIND US ON WWW.TICT.IO

THIS CANVAS IS PART OF THE TECHNOLOGY IMPACT CYCLE TOOL. THIS CANVAS IS THE RESULT OF A QUICKSCAN. YOU CAN FILL OUT THE FULL TICT ON WWW.TICT.IO

Fontys University of Applied Sciences

TICT