

# TaxonomyNet: A Consistent and Efficient Model for Taxonomic Rank Identification in Wildlife Images

Qianqian Zhang<sup>1,\*</sup>, Khandakar Ahmed<sup>1</sup>, Chenhao Xu<sup>1</sup>, Muhammad Imad Khan<sup>1</sup>, and Hua Wang<sup>1</sup>

<sup>1</sup>ISILC, Victoria University, Melbourne, 3011, Australia

\*corresponding.qianqian.zhang@vu.edu.au

## Prompt Design for Foundation Model Evaluation

To enable consistent and reproducible comparisons between the proposed WAL correction strategy and large-scale foundation models, prompt-based interaction protocols were developed for both language and vision models. These prompts served to simulate expert-like taxonomic judgement, instructing models to either confirm or revise multi-rank classification predictions based on structured input.

This appendix outlines the exact prompt configurations used during evaluation. Section [Prompt for Locally Deployed Language Models](#) presents the prompt template employed for locally deployed lightweight language models, which were instructed to perform structured consistency checks based on machine-generated outputs. Section [Prompt for Llama3.2-vision](#) details the corresponding input construction used for large vision-language models, which directly processed image content and returned hierarchical taxonomic predictions.

### Prompt for Locally Deployed Language Models

The following prompt was used to interact with lightweight language models deployed locally. Model predictions were provided in the form of a structured JSON object, detailing outputs across all six taxonomic ranks. The language models were instructed to validate and, if necessary, correct these predictions based on taxonomic knowledge.

You are an expert in animal taxonomy. Here are the predictions from a machine learning model for different taxonomic ranks:

```
{ "phylum": "...", "class": "...", "order": "...", "family": "...", "genus": "...", "species": "..." }
```

If you think any rank is inconsistent or incorrect, please provide a corrected version. Otherwise, confirm the original results.

#### Important Instructions:

- Your response **MUST** be in **valid JSON format**.
- **Do not include any explanation**, only return the JSON.
- Here is the **exact format** you must follow:

```
{  
  "species": "confirmed/corrected name",  
  "genus": "confirmed/corrected name",  
  "family": "confirmed/corrected name",  
  "order": "confirmed/corrected name",  
  "class": "confirmed/corrected name",  
  "phylum": "confirmed/corrected name"  
}
```

Now, please return the corrected JSON below:

## Prompt with Enforced Taxonomic Knowledge

You are an expert in animal taxonomy. Here are the predictions from a machine learning model for different taxonomic ranks:

```
{ "phylum": "...", "class": "...", "order": "...", "family": "...", "genus": "...",  
"species": "..." }
```

You have access to a local taxonomy dictionary stored in `taxonomy_dict.json`. This file contains valid hierarchical relationships between all ranks in the dataset. Use it to check for consistency. If any rank is inconsistent, return a corrected version. Otherwise, confirm the original result.

### Important Instructions:

- Your response **MUST** be in **valid JSON format**.
- **Do not include any explanation**, only return the JSON.
- Here is the **exact format** you must follow:

```
{  
  "species": "confirmed/corrected name",  
  "genus": "confirmed/corrected name",  
  "family": "confirmed/corrected name",  
  "order": "confirmed/corrected name",  
  "class": "confirmed/corrected name",  
  "phylum": "confirmed/corrected name"  
}
```

Now, please return the corrected JSON below:

## Prompt for Llama3.2-vision

The following prompt was used for foundation models that support direct image input. These models were instructed to classify the organism depicted in the image across all six taxonomic ranks and return the results in a structured JSON format:

You are an expert in animal taxonomy. Please examine the input image and determine the taxonomic classification of the animal shown.

### Important Instructions:

- Your response **MUST** be in **valid JSON format**.
- **Do not include any explanation**, only return the JSON.
- Here is the **exact format** you must follow:

```
{  
  "species": "predicted name",  
  "genus": "predicted name",  
  "family": "predicted name",  
  "order": "predicted name",  
  "class": "predicted name",  
  "phylum": "predicted name"  
}
```

Now, please return the classification result for the animal in the image above.