\*\*Assurance Case for the Interpretability of the DeepMind Machine Learning Enabled System\*\*

1. \*\*Goal G1:\*\* Ensure the DeepMind Machine Learning System is Interpretable

- \*\*Strategy S1:\*\* Demonstrate interpretability through model transparency and understandability

- \*\*Goal G1.1:\*\* Ensure transparency in the machine learning model

- \*\*Evidence E1.1.1:\*\* Documentation of model architecture and parameters

- \*\*Evidence E1.1.2:\*\* Documentation of the data flow and processing within the model

- \*\*Goal G1.2:\*\* Ensure the model's decision process is understandable

- \*\*Evidence E1.2.1:\*\* Existence of a feature importance report for model predictions

- \*\*Evidence E1.2.2:\*\* User studies validating the understandability of model outputs

- \*\*Strategy S2:\*\* Validate interpretability through external verification

- \*\*Goal G1.3:\*\* Perform third-party audits to verify model interpretability

- \*\*Evidence E1.3.1:\*\* Third-party audit reports assessing model transparency and understandability

- \*\*Goal G1.4:\*\* Benchmark model interpretability against industry standards

- \*\*Evidence E1.4.1:\*\* Comparison report of interpretability with leading industry standards

2. \*\*Context C1:\*\* Definition of interpretability in context of machine learning

- \*\*Assumption A1:\*\* Interpretability encompasses both the ability to trace model decisions and the clarity of the model’s operating principles to a human expert

3. \*\*Context C2:\*\* Users for the interpretability include both internal developers and external regulators

- \*\*Assumption A2:\*\* Both user groups require understanding the model to different extents and for different purposes

By structuring the assurance case as such, it helps in systematically addressing and documenting the interpretability of the DeepMind Machine Learning System, ensuring clarity, capability, and credibility in presenting the system as interpretable. This tree depicts a logical association of goals, strategies, evidence, and context to build a comprehensive case towards achieving system interpretability.