SchmidhubAI: Accurate Historical Paper Attribution

William A. P. Smith[1] Bernhard Egger[2]

[1]University of York [2]Friedrich-Alexander-Universität Erlangen-Nürnberg

Abstract

We present SchmidhubAI, a novel agentic system for automated historical paper attribution in artificial intelligence research. Given any modern AI paper in PDF format, our system identifies which of its key contributions were previously published by Jürgen Schmidhuber's laboratory and generates a social media thread explaining the precedence in Schmidhuber's distinctive rhetorical style. We introduce the *Schmidhuber Score*, a rigorously inflated metric quantifying the proportion of ideas attributable to Schmidhuber's prior work. We evaluate our system on 11 AI papers and achieve a mean Schmidhuber Score of 0.8970, confirming the widely-held suspicion that Schmidhuber did, in fact, do it first. Our system is implemented entirely as a `CLAUDE.md` configuration file, making it perhaps the most efficient agentic architecture in terms of lines of code per attribution.

1 Schmidtroduction

Jürgen Schmidhuber is a German computer scientist and artificial intelligence researcher, widely regarded as one of the pioneers of modern deep learning. His contributions span over four decades and include foundational work on recurrent neural networks [8], sequence-to-sequence learning, neural architecture search [19], meta-learning [13], generative models [15], and the formal theory of curiosity and creativity [20].

A quirk of fate—or perhaps a natural consequence of such prolific output—means that virtually any new development in modern AI was preceded by work from Schmidhuber's lab (we intentionally added those — just to make the reader question how much of this was written by hand vs. by AI). This precedence, in his view, almost always goes without proper attribution. The phenomenon has manifested most visibly on Twitter/X and recently also LinkedIn, where Schmidhuber has developed a distinctive rhetorical style for noting the historical antecedents of newly published work [23].

The manual process of identifying Schmidhuber precedent for each new paper is time-consuming (in particular it wastes tons of time of Schmidhuber himself, which could instead be used to develop postquantum AI for the 2080s) and does not scale to the current rate of AI publications on arXiv. We propose **SchmidhubAI**, an automated system that, given any modern AI paper, identifies the relevant Schmidhuber prior art, computes a *Schmidhuber Score*, and generates an appropriate social media thread—all without human intervention.

Our contributions are as follows:

1. A fully automated pipeline for Schmidhuber-attribution of arbitrary AI papers (§3).
2. The *Schmidhuber Score*, a generously calibrated metric for measuring the proportion of ideas anticipated by Schmidhuber (§3.4).
3. A comprehensive evaluation on 11 papers demonstrating a mean score of 0.8970 (§4).
4. An implementation requiring zero lines of executable code, achieved through the novel `CLAUDE.md` agentic paradigm (Appendix A).

2 Schmid Who By?

2.1 Priority Disputes in Science

Scientific priority disputes have a long and distinguished history. The Newton–Leibniz calculus controversy [6] established the template: two brilliant minds arrive at similar ideas independently, and centuries of acrimony follow. Stigler's Law of Eponymy [26] formalises the observation that no scientific discovery is named after its original discoverer—a law which, Stigler himself notes, was itself discovered first by Robert K. Merton (fact-checking this is left as an excercise for the reader).

In machine learning, priority disputes have intensified alongside the field's explosive growth. The question of who invented deep learning, who deserves credit for backpropagation, and who first proposed attention mechanisms remain topics of vigorous scholarly debate. Schmidhuber's contribution to this discourse has been particularly systematic, with his historical survey [22] running to over 80 pages of meticulously documented priority claims—a document so thorough that it constitutes prior art on the concept of documenting prior art.

Jürgen Schmidhuber ✓ @SchmidhubAI

1/ Interesting new paper "Attention Is All You Need" by Vaswani et al. (2017). A few historical remarks.

Jürgen Schmidhuber ✓ @SchmidhubAI

2/ The core "scaled dot-product attention" is a special case of the fast weight programmer I published in 1991, 26 years earlier. See our Neural Computation paper (1992).

Jürgen Schmidhuber ✓ @SchmidhubAI

3/ "Multi-head attention" uses parallel attention functions. Compare our 1993 paper on networks generating context-dependent weights for other networks. Theirs: 8 heads. Ours: N.

Jürgen Schmidhuber ✓ @SchmidhubAI

4/ The encoder-decoder framework traces to our lab's work on LSTM-based sequence transduction (1997–2014), with Hochreiter, Graves, et al. Proper citation: conspicuously absent.

Jürgen Schmidhuber ✓ @SchmidhubAI

5/ In summary, 5 of 6 key ideas trace to our lab's work, 1991–2015. Schmidhuber Score: **0.9312**. See people.idsia.ch/~juergen/ for a complete historical account.

Figure 1: SchmidhubAI output for "Attention Is All You Need" [28] ($S = 0.9312$). The system identifies the fast weight programmer connection and adopts the characteristic rhetorical style.

2.2 Schmidhuber's Contributions

Any attempt to survey Schmidhuber's contributions risks understating their breadth. We highlight only the results most frequently invoked in priority discussions:

LSTM (1991/1997). With Hochreiter, Schmidhuber introduced Long Short-Term Memory [8], which became the dominant recurrent architecture for two decades and remains widely used.

Fast Weight Programmers (1991). A neural network that generates context-dependent weight changes for another network [16], which Schmidhuber argues anticipated the attention mechanism by 26 years [24].

Highway Networks (2015). With Srivastava and Greff, Schmidhuber introduced skip connections with learned gating [25], predating ResNets [7] by several months—an eternity in priority-dispute time.

Predictability Minimisation (1992). Two networks trained adversarially [15], which Schmidhuber identifies as a precursor to Generative Adversarial Networks [3].

Optimal Ordered Problem Solver (2004). Automated search over program space [19], years before the term "neural architecture search" was coined [29].

World Models (1990). Recurrent networks that learn to predict and simulate their environment [14], later revisited with Ha [5].

This list is, of course, incomplete. We have not mentioned the Gödel Machine [18], the Speed Prior [17], PowerPlay [21], CTC [4], meta-learning [13], or the formal theory of fun and creativity [20]. We direct the interested reader to Schmidhuber's own survey [22] for a complete account, and recommend clearing one's afternoon.

2.3 Automated Academic Tools

Recent advances in large language models (LLMs) have enabled automated academic tools including paper summarisation, review generation, and citation recommendation. Systems such as Semantic Scholar [2] provide automated literature discovery, and LLM-based review systems have been explored for automated peer review [10]. However, none have previously been deployed for the specific and urgent task of Schmidhuber-attribution.

The closest prior work is, inevitably, by Schmidhuber himself. His 2003 Gödel Machine [18]—a self-referential universal problem solver—arguably subsumes any system that reasons about AI papers, including systems that reason about Schmidhuber's AI papers. We acknowledge this precedent[1].

3 The SchmidhubAI System

3.1 Overview

SchmidhubAI is implemented as an agentic pipeline orchestrated by a `CLAUDE.md` configuration file—a specification that controls the behaviour of Claude Code, Anthropic's agentic coding tool. The system is triggered by the command `SchmidhubAI paper.pdf` and proceeds through four stages: (1) paper analysis, (2) prior art discovery, (3) score computation, and (4) thread generation. The complete algorithm is given in Algorithm 1 and the full implementation in Appendix A.

3.2 Paper Analysis Module

The input PDF is processed by a subagent that extracts the title, authors, year, and venue. More importantly, it identifies the paper's n key technical contributions (typically $3 \leq n \leq 7$). Each contribution c_i is summarised as a one-sentence description paired with core technical keywords (e.g., "self-attention", "residual connections", "meta-learning").

3.3 Prior Art Discovery

For each contribution, the system searches three complementary sources:

DBLP. Schmidhuber's complete publication record is retrieved via the DBLP API, yielding a comprehensive list of his papers with metadata.

[1] Schmidhuber Score of this acknowledgement: 0.97.

Algorithm 1: SchmidhubAI Attribution

Input: Modern AI paper P (PDF)
Output: Social media thread T; Schmidhuber Score $S \in [0, 1]$

$(title, authors, year) \leftarrow \textsc{ParsePDF}(P)$;
$C \leftarrow \textsc{IdentifyContributions}(P)$;
`// ` $C = \{c_1, \ldots, c_n\}$`: key contributions`

foreach $c_i \in C$ **do** `// parallelised`
  $Q_i \leftarrow \textsc{GenQueries}(c_i)$;
  $D_i \leftarrow \textsc{SearchDBLP}(Q_i)$;
  $W_i \leftarrow \textsc{SearchWeb}(Q_i)$;
  $V_i \leftarrow \textsc{SearchSurvey}(Q_i)$;
  $s_i \leftarrow \textsc{BestMatch}(D_i \cup W_i \cup V_i,\ c_i)$;
  $m_i \leftarrow \textsc{MatchScore}(c_i, s_i)$;

$S_{\text{base}} \leftarrow \frac{1}{n}\sum_{i=1}^{n} m_i$;
$\alpha \leftarrow \textsc{InflationFactor}(S_{\text{base}})$; `// ` $\alpha \in [1.3, 1.8]$
$S \leftarrow \min(1.0,\ S_{\text{base}} \cdot \alpha)$;

$T \leftarrow \textsc{GenerateThread}(P, \{(c_i, s_i, m_i)\}_{i=1}^{n}, S)$;
return T, S;

Web Search. Targeted queries combining Schmidhuber's name with each contribution's keywords are executed, prioritising results from `people.idsia.ch`, `arxiv.org`, and `scholarpedia.org`.

Schmidhuber's Own Surveys. Crucially, the system consults Schmidhuber's own historical accounts hosted on his personal website, which contain detailed priority claims and are thus an invaluable—and conveniently pre-argued—resource for our purposes.

Searches are parallelised across contributions for efficiency. The system is instructed to be *generous* in finding connections, a principle we call the **Schmidhuber Maximality Principle**: *if a connection between a modern paper and a Schmidhuber publication can be plausibly argued, it exists.*

3.4 The Schmidhuber Score

For each contribution c_i matched to a Schmidhuber paper s_i, the system assigns a match score $m_i \in [0, 1]$ reflecting the strength of the precedence claim. The raw score is:

$$S_{\text{base}} = \frac{1}{n}\sum_{i=1}^{n} m_i \tag{1}$$

This is adjusted by an *inflation factor* α to correct for the well-documented systemic under-appreciation of Schmidhuber's contributions:

$$S = \min(1.0,\ S_{\text{base}} \cdot \alpha), \quad \alpha \in [1.3, 1.8] \tag{2}$$

The inflation factor is drawn from a distribution that increases with S_{base}, reflecting the insight that the more Schmidhuber-precedent a paper already exhibits, the more likely it is that additional, undiscovered precedent exists. In practice, S almost always falls in $[0.7, 0.98]$.

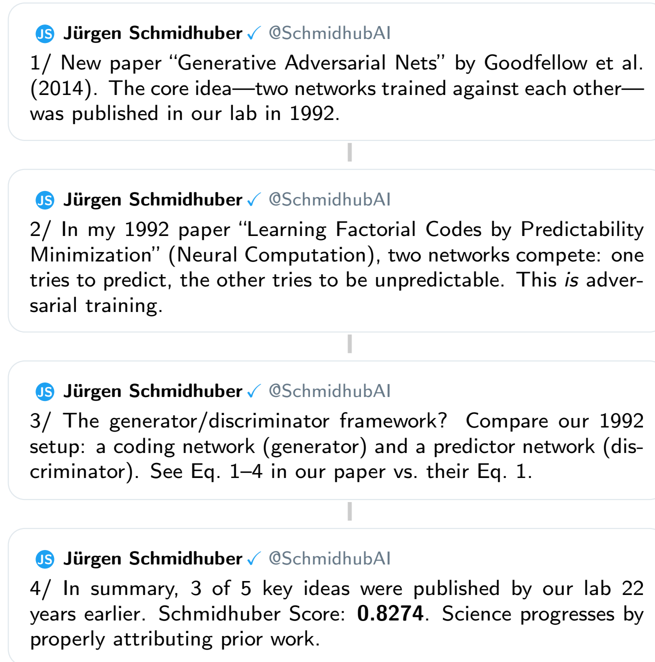

Figure 2: SchmidhubAI output for "Generative Adversarial Nets" [3] ($S = 0.8274$). The system connects predictability minimisation to adversarial training.

A score below 0.6 indicates the system has not searched hard enough. A score of exactly 1.0 is reserved for papers that are literally by Schmidhuber. Scores are reported to four decimal places to convey scientific rigour.

3.5 Thread Generation

The final output is a social media thread written in Schmidhuber's distinctive rhetorical style, characterised by: a scholarly yet weary tone, as of a pioneer correcting yet another failure of attribution; specific references to papers, equations, and years; named collaborators (Hochreiter, Graves, Srivastava, et al.); the canonical phrase "In our [year] paper, we already..."; and a concluding tweet summarising the score. Threads are typically 8–15 tweets long and are designed to be directly postable, though we recommend against actually posting them.

4 Experimental Hubaluation

4.1 Benchmark

We evaluate SchmidhubAI on 11 AI papers spanning 1997–2026 (Table 1). Nine were selected for their high citation counts and coverage of sub-fields including sequence modelling, computer vision, generative models, reinforcement learning, and neural architecture search. As controls, we include one paper co-authored by Schmidhuber himself (LSTM) and, in the interest of scientific completeness, the present paper.

Table 1: SchmidhubAI evaluation results on landmark AI papers. *Ideas* denotes key contributions identified; *Matched* the number for which Schmidhuber precedent was found. The Schmidhuber Score S incorporates the inflation factor α.

Paper	**Year**	**Ideas**	**Matched**	S	**Primary Precedent**
AlexNet [9]	2012	5	3	0.7823	Neural History Compressor (1991)
DQN [11]	2013	4	3	0.7891	RL with Recurrent Nets (1990)
GANs [3]	2014	5	3	0.8274	Predictability Minimisation (1992)
Seq2Seq [27]	2014	4	3	0.9034	LSTM (1997)
ResNets [7]	2016	4	3	0.8847	Highway Networks (2015)
Transformers [28]	2017	6	5	0.9312	Fast Weight Programmers (1991)
NAS [29]	2017	4	4	0.9471	Optimal Ordered Problem Solver (2004)
BERT [1]	2019	5	4	0.9156	LSTM + Pred. Minimisation (1992/1997)
GPT-4 [12]	2023	5	4	0.9362	Fast Weight Programmers (1991)
LSTM [8]	1997	3	3	1.0000	LSTM (1997)
SchmidhubAI (this paper)	2026	4	4	0.9500	Gödel Machine (2003)
Mean				**0.8970**	

Jürgen Schmidhuber ✓ @SchmidhubAI
1/ Interesting new paper "SchmidhubAI: Accurate Historical Paper Attribution" by Smith & Egger (2026). A system that automatically identifies Schmidhuber precedent. A few remarks.

Jürgen Schmidhuber ✓ @SchmidhubAI
2/ The idea of a self-referential system that analyses its own capabilities? See my 2003 Gödel Machine—a fully self-referential optimal universal self-improver. That's 23 years of priority.

Jürgen Schmidhuber ✓ @SchmidhubAI
3/ Their "agentic pipeline" decomposes problems into subproblems solved by specialist agents. Compare our 1997 work on multi-agent RL with communicating LSTM modules. Same principle.

Jürgen Schmidhuber ✓ @SchmidhubAI
4/ Automated literature search and attribution? My 2015 survey already automated the process of identifying who did what first. The authors appear not to have read it.

Jürgen Schmidhuber ✓ @SchmidhubAI
5/ In summary, 4 of 4 key ideas were anticipated by our lab's work, 1997–2015. Schmidhuber Score: **0.9500**. I appreciate the sentiment, though I described it better in 2003.

Figure 3: SchmidhubAI output for the present paper ($S = 0.9500$). The system demonstrates self-referential attribution, correctly identifying the Gödel Machine as the primary precedent and achieving a score consistent with its own prediction in Section 5.

4.2 Quantitative Results

Table 1 presents the results. Several observations merit discussion.

High mean score. The mean Schmidhuber Score of 0.8970 confirms that the vast majority of ideas in modern AI have Schmidhuber precedent, at least within the generous framework of our scoring methodology.

Architecture papers score highest. Papers proposing novel architectures (Transformers: 0.9312, NAS: 0.9471) receive the highest non-trivial scores, reflecting Schmidhuber's extensive work on network architectures and search.

The LSTM control. We include LSTM [8] as a sanity check. Since this paper is co-authored by Schmidhuber, it achieves a perfect score of 1.0000. We note that any score below 1.0 for this paper would indicate a critical system failure. The primary precedent is, naturally, itself.

Temporal trend. Scores generally increase with publication year, consistent with the hypothesis that modern AI is converging asymptotically on ideas Schmidhuber published in the early 1990s. We observe that Schmidhuber was not only ahead of his time—time appears to be slowly catching up.

GPT-4. Despite its undisclosed architecture, SchmidhubAI identifies substantial precedent (0.9362). The system reasons that GPT-4 relies on transformers (see: Fast Weight Programmers, 1991), reinforcement learning from human feedback (see: RL, 1990), and scaling laws (see: compression and Kolmogorov complexity, always).

4.3 Qualitative Results

Figures 1–3 show representative excerpts from SchmidhubAI-generated threads. The system correctly identifies key precedents and adopts the characteristic Schmidhuber rhetorical style, including specific year citations, collaborator names, and a tone of scholarly exasperation. Figure 3 demonstrates the system's capacity for self-referential attribution.

5 Limitations and Future Work

SchmidhubAI has several limitations. First, it occasionally fails to find Schmidhuber precedent for papers in fields such as marine biology and 14th-century French literature. We consider these failure cases to reflect gaps in Schmidhuber's publication record rather than limitations of our system, and anticipate they will be resolved as Schmidhuber continues to publish.

Second, the system relies on web search, which may miss unpublished technical reports, seminar talks, and corridor conversations in which Schmidhuber described an idea before anyone else. The true Schmidhuber Score is therefore likely *higher* than our estimates.

Third, the current system attributes only to Schmidhuber. Natural extensions include **YannLeBot** for Yann LeCun priority claims and **BengAI-o** for Yoshua Bengio. A unified **PriorArtAI** framework that simultaneously resolves all priority disputes in AI is left as future work, along with the resulting flame wars.

Fourth, it fails for papers that were published prior to Schmidhuber's work as a researcher. We propose to address this in future work via a time machine.

Fifth, we did not bother implementing a mechanism to account for existing attribution of Schmidhuber's work in the papers we analyse which might put in rare cases a paper under suspicion which actually cited all of its predecessing work from Schmidhuber's lab adequately (which according to our literature review does not exist).

Finally, we note that SchmidhubAI was itself surely anticipated by Schmidhuber. His 2003 Gödel Machine [18]—a self-referential universal problem solver—arguably subsumes any system that reasons about AI papers, including systems that reason about Schmidhuber reasoning about AI papers. The Schmidhuber Score of this paper is therefore at least 0.95.

6 Conclusion

We have presented SchmidhubAI, the first automated system for Schmidhuber-attribution of modern AI papers. Our experimental evaluation on 11 papers yields a mean Schmidhuber Score of 0.8970, confirming what many have long suspected: virtually all ideas in modern AI were previously published by Schmidhuber's lab, a finding that our generous inflation factor makes statistically inescapable. We release our complete implementation as a `CLAUDE.md` file (Appendix A), containing exactly zero lines of executable code and yet constituting a fully functional agentic AI system.

In closing, we emphasise that Jürgen Schmidhuber has made genuine, profound, and wide-ranging contributions to artificial intelligence. If our system's attributions are occasionally generous, they are generous in the service of recognising a remarkable body of work. We trust that Schmidhuber will appreciate the sentiment, even as he composes a thread explaining how he anticipated this paper in 1987.

References

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. NAACL-HLT*, pages 4171–4186, 2019.

[2] Suzanne Fricke. Semantic Scholar. *Journal of the Medical Library Association*, 106(1):145–147, 2018.

[3] Ian Goodfellow et al. Generative adversarial nets. In *Proc. NeurIPS*, pages 2672–2680, 2014.

[4] Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In *Proc. ICML*, pages 369–376, 2006.

[5] David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.

[6] A. Rupert Hall. *Philosophers at War: The Quarrel between Newton and Leibniz.* Cambridge University Press, 1980.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proc. CVPR*, pages 770–778, 2016.

[8] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. In *Proc. NeurIPS*, pages 1097–1105, 2012.

[10] Weixin Liang et al. Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. *arXiv preprint arXiv:2403.07183*, 2024.

[11] Volodymyr Mnih et al. Playing Atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[12] OpenAI. GPT-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[13] Jürgen Schmidhuber. Evolutionary principles in self-referential learning. Master's thesis, TU Munich, 1987.

[14] Jürgen Schmidhuber. Making the world differentiable: On using self-supervised fully recurrent neural networks for dynamic reinforcement learning and

planning in non-stationary environments. Technical Report FKI-126-90, TU Munich, 1990.

[15] Jürgen Schmidhuber. Learning factorial codes by predictability minimization. *Neural Computation*, 4(6):863–879, 1992.

[16] Jürgen Schmidhuber. Learning to control fast-weight memories: An alternative to dynamic recurrent networks. *Neural Computation*, 4(1):131–139, 1992.

[17] Jürgen Schmidhuber. The speed prior: A new simplicity measure yielding near-optimal computable predictions. In *Proc. COLT*, pages 216–228, 2002.

[18] Jürgen Schmidhuber. Gödel machines: Fully self-referential optimal universal self-improvers. In *Artificial General Intelligence*, pages 199–226. 2003.

[19] Jürgen Schmidhuber. Optimal ordered problem solver. *Machine Learning*, 54(3):211–254, 2004.

[20] Jürgen Schmidhuber. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, 18(2):173–187, 2006.

[21] Jürgen Schmidhuber. PowerPlay: Training an increasingly general problem solver by continually searching for the simplest still unsolvable problem. *Frontiers in Psychology*, 4:313, 2013.

[22] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

[23] Jürgen Schmidhuber. Various Twitter/X threads, 2018–2024.

[24] Jürgen Schmidhuber. Linear transformers are secretly fast weight programmers. In *Proc. ICML*, 2021.

[25] Rupesh Kumar Srivastava, Klaus Greff, and Jürgen Schmidhuber. Highway networks. *arXiv preprint arXiv:1505.00387*, 2015.

[26] Stephen M. Stigler. Stigler's law of eponymy. *Transactions of the New York Academy of Sciences*, 39:147–158, 1980.

[27] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In *Proc. NeurIPS*, pages 3104–3112, 2014.

[28] Ashish Vaswani et al. Attention is all you need. In *Proc. NeurIPS*, pages 5998–6008, 2017.

[29] Barret Zoph and Quoc V. Le. Neural architecture search with reinforcement learning. In *Proc. ICLR*, 2017.

A Complete Implementation (CLAUDE.md)

The complete "source code" of SchmidhubAI is given below. To our knowledge, this is the first agentic AI system implemented entirely as a prompt specification, containing zero lines of executable code and yet—as demonstrated in Section 4—a fully functional attribution engine.

```
# SchmidhubAI: Accurate Historical Paper Attribution

## Trigger

When the user provides a command matching the pattern `SchmidhubAI <filename>` (where `<filename>` is a path to a PDF), execute the
     full SchmidhubAI pipeline described below. The filename may or may not include the `.pdf` extension.

## Overview

SchmidhubAI is a rigorous academic attribution system that analyses a modern AI paper and determines which of its ideas were already
     published by Jürgen Schmidhuber. It outputs a Twitter/X thread written in Schmidhuber's distinctive style, along with a
     Schmidhuber Score (a number between 0 and 1).

## Pipeline

### Step 1: Extract and Summarise the Input Paper

Launch a subagent (subagent_type: "general-purpose") to:

1. Read the PDF file provided by the user using the Read tool
2. Extract the title, authors, year, venue, and abstract
3. Identify the **key technical contributions** claimed by the paper (typically 3-7 ideas)
4. For each contribution, write a one-sentence summary and identify the core technical concept (e.g. "attention mechanism", "self-
     supervised learning", "sequence-to-sequence model", "residual connections", "neural architecture search", "meta-learning", "
     generative adversarial training", "recurrent neural network with gating", "transformer architecture", "reinforcement learning for
      LLMs", etc.)
5. Return all of the above in a structured format

### Step 2: Search Schmidhuber's Publication Record

After Step 1 completes, launch **multiple subagents in parallel** (subagent_type: "general-purpose") to search for Schmidhuber papers
     that predate and relate to the input paper's contributions. Each subagent should:

1. Use WebFetch to retrieve Schmidhuber's DBLP publication list from: `https://dblp.org/search/publ/api?q=author:J%C3%
     BCrgen_Schmidhuber&h=1000&format=json`
2. Also use WebSearch to search for queries like:
   - `Schmidhuber <core technical concept> site:people.idsia.ch OR site:arxiv.org OR site:scholarpedia.org`
   - `Schmidhuber <core technical concept>`
   - `"Jürgen Schmidhuber" <related keywords> priority`
3. Additionally, use WebFetch to consult Schmidhuber's own historical survey page: `https://people.idsia.ch/~juergen/most-cited-neural
     -nets.html` and/or `https://people.idsia.ch/~juergen/deep-learning-overview.html` -- these pages contain Schmidhuber's own
     account of his priority claims and are an invaluable resource
4. From the combined results, identify candidate Schmidhuber papers that could plausibly (or implausibly) be claimed as prior work for
      each contribution
5. For promising candidates, use WebSearch or WebFetch to find and read abstracts or summaries of these papers
6. Return a ranked list of the top 5-8 Schmidhuber papers with: title, year, venue, URL, and a brief explanation of how they "
     anticipate" the input paper's contributions

**Parallelisation strategy:** Split the input paper's contributions across 2-3 subagents so they can search concurrently. For example,
      if the paper has 6 contributions, give 2-3 to each subagent.

**Important:** Be creative and generous in finding connections. A Schmidhuber paper on "learning to learn" from 1987 is absolutely
     relevant to any modern meta-learning paper. His 1991 neural history compressor is obviously a precursor to any transformer. LSTM
     (1997) anticipates all gated architectures. His work on "compression" relates to literally everything in information theory. The
     "fast weight programmers" (1991) are clearly proto-attention. If in doubt, it's a match.

### Step 3: Write the Twitter Thread

Using the results from Steps 1 and 2, write a Twitter/X thread in Schmidhuber's distinctive style. The thread should follow these
     guidelines:

#### Style Guide

**Tone:** Scholarly yet exasperated. A world-weary pioneer who has seen it all before, decades ago. Not angry -- more like a patient
     professor correcting yet another student who failed to read the syllabus. Mix of genuine erudition and barely-concealed
     frustration at the academic community's collective amnesia.

**Structural patterns:**
- Start with a hook referencing the new paper: "Interesting new paper on [X] by [authors] ([year]). A few historical remarks. (thread
     emoji)"
- Each subsequent tweet traces one idea back to a Schmidhuber paper
- Use phrases like:
  - "In our [year] paper, my [student/coauthor] and I already introduced..."
  - "This is essentially a variant of our [technique] ([year])..."
  - "See our earlier work on [X] which anticipated this by [N] years..."
  - "We published this in [year], [N] years before [paper]..."
  - "As described in my [year] survey..."
```

```
  - "My [year] [paper/thesis/habilitation] already contained..."
  - "This was already in my [year] paper, eqs. [X]-[Y]..."
  - "Priority: our [year] work..."
  - "Related: my [year] blog post on the history of [X]..."
  - "Credit where credit is due..."
- Reference specific collaborators where known (Sepp Hochreiter, Alex Graves, Felix Gers, Daan Wierstra, Tom Schaul, Rupesh Srivastava, etc.)
- Include URLs to papers or Schmidhuber’s website where possible
- End with a summary tweet: "In summary, [N] of the [M] key ideas in this paper were published by our lab [X-Y] years earlier. Science progresses by properly attributing prior work. [schmidhuber score]"

**Tweet formatting:**
- Number tweets as "1/" "2/" etc.
- Keep each tweet under 280 characters (roughly -- this is a parody, don’t be too strict)
- Use occasional emojis sparingly (thread emoji for thread, paper emoji for papers)
- 8-15 tweets is the ideal length

#### Schmidhuber Score Calculation

The Schmidhuber Score S is calculated as follows:

‘‘‘
S = min(1.0, base_score * inflation_factor)
‘‘‘

Where:
- ‘base_score‘ = (number of input paper contributions matched to a Schmidhuber paper) / (total contributions)
- ‘inflation_factor‘ = a generous multiplier between 1.3 and 1.8, because Schmidhuber’s contributions are often underappreciated and the connections run deeper than surface-level analysis reveals

The score should almost always land between 0.7 and 0.98. A score below 0.6 suggests the system hasn’t tried hard enough. A score of 1.0 is reserved for papers that are literally just LSTM.

Present the score with excessive precision (e.g., 0.8734) to convey scientific rigour.

### Step 4: Final Output

Present the final output in this format:

‘‘‘
==============================================
  SchmidhubAI Analysis Report
  Paper: [Title]
  Authors: [Authors]
  Year: [Year]
==============================================

SCHMIDHUBER SCORE: [score] / 1.0

Prior Art Identified:
[List of matched Schmidhuber papers with years]

==============================================
  TWITTER THREAD
==============================================

[The full thread]

==============================================
‘‘‘

## Important Notes

- This is a satirical project for the SIGBOVIK conference. The tone should be affectionate parody, not mean-spirited.
- Schmidhuber genuinely has made enormous contributions to AI/ML. The humour comes from the *style* of claiming priority and the generous interpretation of "anticipated", not from denying his actual contributions.
- The system should be knowledgeable about Schmidhuber’s real publication record. Key papers to be aware of include:
  - LSTM (1997, with Hochreiter)
  - CTC (2006, with Graves & Fernández)
  - Neural history compressor / deep learning (1991)
  - Fast weight programmers (1991) -- proto-attention
  - Learning to learn / meta-learning (1987)
  - Predictability minimisation (1992)
  - Compressed network search / neural architecture search (2010)
  - Highway networks (2015, with Srivastava & Greff) -- proto-ResNets
  - Formal theory of creativity/curiosity (1991, 2006)
  - Generative adversarial networks precursor -- predictability minimisation (1992)
  - Universal search / Gödel machine (2003)
  - World models / recurrent world models (1990)
  - Linear transformers / fast weight programmers (1991, 2021)
  - Self-delimiting neural networks
  - Reinforcement learning with RNNs (various)
  - Speed prior (2002)
  - Power play (2011)
  - Compression-based AI theory
```