

УДК: 165:004.8

DOI: 10.15372/PS20250305

EDN: RYFNQL

А.А. Шевченко**НЕЯВНОЕ ЗНАНИЕ:
ОТ М. ПОЛАНИ ДО ALPHAZERO**

В статье анализируется концепция неявного знания, введенная М. Полани, и ее актуальность в современном научном и технологическом контексте. Рассматривается исторический фон идеи, ее эпистемологические последствия и социальное измерение коллективного неявного знания в науке. Особое внимание уделено проблеме «черного ящика» в искусственном интеллекте на примере эволюции шахматных алгоритмов от Deep Blue к AlphaZero, демонстрирующих переход от формализованного явного знания к непрозрачному неявному знанию. В статье подчеркивается, что неявное знание сохраняет ключевое значение для понимания природы научного познания и современных когнитивных систем.

Ключевые слова: неявное знание; эпистемология; философия науки; искусственный интеллект; «черный ящик»; коллективное знание; Deep Blue; AlphaZero; эпистемологический сдвиг

A.A. Shevchenko**TACIT KNOWLEDGE:
FROM M. POLANYI TO ALPHAZERO**

The article explores Michael Polanyi's concept of tacit knowledge and its relevance in contemporary scientific and technological contexts. It reviews the historical and epistemological background, the social dimension of collective tacit knowledge in science, and the challenges posed by the "black box" problem in artificial intelligence. Using the evolution of chess algorithms from Deep Blue to AlphaZero as a case study, the article illustrates the shift from explicit, formalized knowledge to opaque, tacit knowledge. It emphasizes the enduring

importance of tacit knowledge for understanding the nature of scientific practice and modern cognitive systems.

Keywords: tacit knowledge; epistemology; philosophy of science; artificial intelligence; “black box”; collective knowledge; Deep Blue; AlphaZero; epistemological shift

Исторический и эпистемологический контекст

Концепция неявного знания, представленная Майклом Полани в «Личностном знании» в 1958 г. [2] и детализированная в «Неявном измерении» в 1966 г. [7] долгое время оставалась на периферии аналитической эпистемологии. Уже само название первой монографии было вызовом идеалу объективности, к которому стремилась аналитическая философия, так как понятие «личное» воспринималось как синоним субъективного, эмоционального и ненадежного. А ключевое утверждение М. Полани «мы знаем больше, чем можем сказать», ставшее основным тезисом второй монографии [7], бросает вызов традиционным моделям знания как пропозиционального, верифицируемого и формализуемого. Оно указывает на существование познавательного опыта, который не может быть полностью передан в форме текста, правил или формул, такое знание реализуется в действиях, интуиции, практическом мастерстве.

Еще более фундаментально то, что традиционная наука со времен Просвещения строилась на идеалах прозрачности, воспроизводимости и объяснимости. Закон всемирного тяготения Ньютона предлагает элегантную и понятную математическую модель, которая не только предсказывает движение планет с высокой точностью, но и описывает силы, его вызывающие. Все параметры модели измеримы, а ее предсказания поддаются экспериментальной проверке. М. Полани же утверждал, что всякая форма знания, даже самая строгая наука, опирается на неформализуемый пласт компетенций, мастерства и личного участия.

Однако, хотя концепция неявного знания М. Полани и оставалась в тени, она все же не была проигнорирована полностью. Во-первых, в ней имеется очевидное сходство с различием Г. Райла между «знанием-как» (knowing-how) и «знанием-что» (knowing-that), суть которого изложена в книге «Понятие сознания» [3] и статье «Зна-

ние-как и знание-что» [9]. Г. Райл утверждал, что такое умение как, например, езда на велосипеде не сводится к знанию пропозициональных истин о равновесии и механике. Оно представляет собой автономную форму познания, которая не требует вербализации. Однако в аналитической традиции «знание-как» часто воспринималось как вторичное, производное от «знания-что», и это ограничивало его философскую значимость. М. Полани же претендует на преодоление этой дихотомии, утверждая, что неявное знание – не просто навык, а основание всякого познания. «Хотя неявное знание может существовать само по себе, явное знание должно опираться на неявное понимание и применение. Следовательно, всякое знание либо неявное, либо укоренено в неявном знании. Совершенно явное знание невозможно представить» [8, р. 144]. Это утверждение имеет радикальные последствия: оно ставит под сомнение саму возможность создания полностью формализованной науки, в которой все знание будет кодифицировано и доступно для логического анализа.

Другой причиной сохранения интереса к идее неявного знания было то, что если для аналитической эпистемологии она была маргинальной, то в философии науки ее влияние было более заметным. Идеи М. Полани имели параллели с работами Томаса Куна, особенно с понятием научной парадигмы, которое усваивается учеными через практику, а не через формальные правила. Ученый, работающий внутри парадигмы, обладает огромным пластом неявного знания о том, как ставить эксперименты и интерпретировать результаты. Это знание не зафиксировано в учебниках, но оно критически важно для функционирования науки [1, гл. 5].

Неявное знание в науке и его артикуляция

Идеи М. Полани стали по-настоящему востребованы лишь в последнюю четверть века с развитием социальной эпистемологии (которая изучает знание как коллективный феномен), эпистемологии добродетелей (которая смещает фокус с анализа убеждений на анализ познавательных способностей и навыков самого познающего субъекта) и, особенно, в связи с бурным развитием систем искусственного интеллекта, таких как глубокие нейронные сети и большие языковые модели.

Современные ИИ-системы, такие как AlphaFold (предсказание структуры белков) или Kepler (обнаружение экзопланет), в узкоспециализированных задачах демонстрируют эффективность, превосходящую человеческую. Однако их внутренняя логика принятия решений часто остается для нас недоступной, что затрудняет понимание, как и почему был получен конкретный результат. Эти модели обучаются на огромных массивах данных и формируют сложные паттерны, корреляции, веса связей, которые невозможно полностью интерпретировать. Это знание существует не в форме правил, а в форме распределенных числовых значений, что делает его аналогом неявного знания, описанного М. Полани.

Этот феномен получил название «проблема черного ящика» и породил целую область исследований – объяснимый ИИ (Explainable AI, XAI). Цель XAI – сделать «прозрачными» решения, принимаемые нейросетями, чтобы можно было их проверить, объяснить и доверять им. Однако до сих пор не существует универсального способа декодировать «мысли» глубокой нейронной сети. Это означает, что мы можем доверять результату, но не понимаем метода, что противоречит самой сути научного подхода.

«Неявность» и «непрозрачность» знания порождают сразу несколько проблем. Во-первых, это проблема верификации и доверия. Как мы можем доверять результату, механизм получения которого не понимаем? Наука всегда опиралась на критическую проверку и фальсификацию гипотез. Но как фальсифицировать «интуицию» машины? Мы можем проверить ее предсказания эмпирически (например, подтвердить структуру белка в лаборатории), но это оставляет сам метод познания в виде «черного ящика». Это и есть знание без понимания, что представляет собой кризис самой объяснительной модели, ведь главная цель науки – не только предсказывать, но и понимать мир. Опасность в том, что это может привести к появлению научных теорий, которые работают, но не имеют под собой ясной теоретической базы, понятной человеку. И, наконец, еще одна проблема связана с передачей знания. Если ключевые открытия совершаются с помощью неявного знания ИИ, как обучать следующее поколение ученых? Достаточно ли будет научить их «общаться» с ИИ-систе-

мами, задавая правильные вопросы и интерпретируя ответы, или же мы потеряем фундаментальное понимание самой сути научного поиска?

Центральная эпистемологическая проблема заключается в том, может ли неявное знание быть полностью эксплицировано. Как справедливо отмечается в статье, посвященной проблеме обоснования «недокументарного» знания: «Некоторые проблемы, с которыми мы сталкиваемся в связи с неявным знанием, с этой ситуацией “знать больше, чем можешь выразить словами”, заключаются в том, можно ли сделать неявное знание явным. Можно ли его кодифицировать как “знание-что” или информацию?» [4, р. 97].

Социальное измерение: коллективное неявное знание

Классическая интерпретация концепции М. Полани фокусируется на индивидуальном исследователе, подчеркивая личностный характер неявного знания как интуитивного, неартикулируемого компонента познания. Однако современная наука представляет собой в высшей степени коллективное предприятие, где знания генерируются, передаются и применяются в рамках социальных сетей, лабораторий и междисциплинарных команд. Это ведет к переосмыслению представлений о носителе неявного знания: общепринятым становится представление о коллективном субъекте такого знания, где неявное знание проявляется не только в индивидуальном сознании, но и в групповых практиках. В науке неявное знание может быть свойством не отдельного разума, а группы как единого целого, воплощенным в принятых научных практиках, распределенном труде и неформальных нормах взаимодействия. «Знание-как» может принадлежать всей лаборатории или организации, даже если ни один ее член по отдельности не способен полностью его изложить.

Это позволяет анализировать сложные научные коллаборации не как простую сумму индивидуальных знаний, а как единую эпистемическую систему. Такое знание служит необходимой основой как для явного знания, так и для самого существования научной практики. Именно фундаментальный характер и неустранимость неявного знания и подчеркивает М. Полани: «Провозглашенная цель совре-

менной науки – установить строго отстраненное, объективное знание. Любое отклонение от этого идеала принимается лишь как временный недостаток, который следует стремиться устранить. Но предположим, что неявная мысль является неотъемлемой частью всякого знания – тогда идеал устранения всех личностных элементов знания фактически будет означать стремление уничтожить все знание» [8, p. 20].

Социальная трактовка неявного знания радикально расширяет рамки, изначально заданные М. Полани. Такое знание перестает быть исключительно психологическим или феноменологическим свойством индивида. Оно становится организационным принципом устройства не только социальных институтов (в частности института науки), но и общества в целом. Общим примером «простого» коллективного неявного знания может служить езда на велосипеде в автомобильном потоке, где участники неосознанно координируют свои действия на основе неартикулируемых норм и ожиданий. Такое знание как бы «распределено» в обществе и фактически его организует [6, p. 26–27]. Оно не зафиксировано в правилах, но регулирует поведение. В рамках «социальной» интерпретации неявное знание – это не просто «еще не формализованное» знание, подлежащее экспликации. Напротив, социальный ракурс подчеркивает его принципиальную неформализуемость, поскольку оно встроено в динамику социальных отношений и культурных контекстов. Именно в социальном измерении неявное знание М. Полани эволюционировало от индивидуальных интуиций к коллективному феномену, интегрирующему эпистемологию с социологией науки. В эпоху глобальных коллабораций неявное знание становится ключом к пониманию того, как различные общества генерируют и поддерживают знания, выходя за пределы индивидуального опыта.

Эволюция шахматных алгоритмов: от явного знания к неявному

Хорошую возможность для более глубокого понимания этих проблем дает эволюция шахматных алгоритмов. Шахматы – идеальная модель для анализа эпистемологических трансформаций, так

как эта игра имеет четкие правила, но бесконечное множество стратегий, что делает ее одновременно формализуемой и интуитивной. Переход от систем, основанных на человеческой экспертизе, к системам, обучающимся «с чистого листа», является не просто технологическим скачком, а разворачивающимся эпистемологическим экспериментом, который вновь актуализирует классические проблемы философии науки, касающиеся природы знания, способов его обоснования и передачи. История шахматных алгоритмов прослеживается как переход от явных, формализованных процедур к системам, в которых «знание» материализуется в сложно организованных, но нечетко интерпретируемых весах и аксиомах, что очень похоже на интересующее нас неявное знание.

Важнейшая отправная точка здесь — знаменитая статья Клода Шеннона (1950 г.) [11], в которой он формулирует задачу «как программировать ЭВМ для игры в шахматы». Одно из ключевых утверждений: «Основной тезис нашей работы заключается в том, что современные универсальные компьютеры способны играть в шахматы на вполне приемлемом уровне благодаря использованию соответствующего алгоритма (или «программы») [11, р. 257]. Эта идея легла в основу всей классической шахматной программы. Эта статья задает и соответствующую рамку, в которой шахматы можно понимать как проблему формальной процедуры поиска и оценки позиций. К. Шеннон стремится редуцировать игру к набору алгоритмических правил: оценочным функциям, дереву поиска и эвристикам. Система IBM Deep Blue, победившая чемпиона мира по шахматам в 1997 г., стала вершиной парадигмы программирования, основанной на явном знании. Ее архитектура, несмотря на сложность, была эпистемологически прозрачной. Она опиралась на исчерпывающий перебор вариантов (*brute-force computation*) и оценочную функцию, которая представляла собой сложную систему эвристик, сформулированных гроссмейстерами.

Один из создателей Deep Blue, М. Кэмпбелл, описывал этот процесс как попытку формализовать интуицию экспертов: «Носителем шахматных знаний была оценочная функция. Она включала более 8000 параметров... Мы консультировались с гроссмейстерами, стремясь формализовать их знание в этих параметрах. Это, например,

параметры “безопасность короля” или “ладья на открытой вертикали”. При формировании оценочной функции для расширенного описания учитывается ряд факторов, в том числе частота или регулярность применения хода гроссмейстерами» [5, р. 67].

Таким образом, Deep Blue представлял собой систему, эпистемологический статус которой был полностью прозрачен. Его «знание» было совокупностью решений, внедренных человеком. Он не обладал автономным пониманием, а лишь с огромной скоростью применял заданные ему инструкции. Объяснение хода Deep Blue, в принципе, соответствовало дедуктивно-номологической модели научного объяснения Карла Гемпеля. Согласно этой модели, объяснить феномен – значит подвести его под общий закон. В случае Deep Blue «законами» выступали множество правил его оценочной функции (например, «ценность контроля над центром равна X », «безопасность короля равна Y »), а «начальными условиями» – конкретная позиция на доске. Ход дедуктивно выводился как оптимальный с точки зрения максимизации значения этой функции. Проблема была в сложности, а не в принципиальной непознаваемости. Deep Blue, таким образом, представлял собой попытку преобразовать неявное знание гроссмейстеров (стратегическую интуицию) в явные параметры (8000 формальных правил). Эта попытка, однако, демонстрирует и пределы формализации: даже 8000 параметров не смогли полностью «ухватить» шахматную интуицию (что доказал более поздний разгром Deep Blue версиями ИИ нового поколения).

Появление AlphaZero в 2017 г. стало радикальным разрывом с предыдущей парадигмой. В отличие от своих предшественников, AlphaZero не использовала человеческие эвристики или базы данных. Ее архитектура, основанная на обучении с подкреплением (reinforcement learning) и самоигре, позволила ей приобрести знание «с чистого листа». Как сообщали разработчики AlphaZero в основополагающей статье: «AlphaGo Zero обучается с чистого листа (*tabula rasa*), исключительно за счет самоигры, без использования человеческих данных, экспертных знаний или каких-либо априорных сведений, кроме базовых правил игры» [12, р. 354]. Знание AlphaZero не существует в виде набора артикулируемых правил. Система демонстрирует глубокое позиционное «понимание», но это понимание не

является дискурсивным. Оно представляет собой очень эффективное «знание-как» без соответствующего «знания-что» в его пропозициональной форме. Фактически, это искусственный аналог неявного знания М. Полани.

Такое знание стало примером того, что в современной эпистемологии называется «эпистемической непрозрачностью» (epistemic opacity). Непрозрачность AlphaZero не является просто следствием масштаба (хотя он и важен). Она является следствием природы его знания, которое не представлено в виде дискретных правил, а распределено по множеству числовых весов нейронной сети.

Авторы упомянутой статьи в *Nature* прямо заявляют, что AlphaZero достигла сверхчеловеческого уровня в шахматах за относительно короткий тренировочный период. Ее знание представляет собой содержательное, стратегическое понимание, близкое к человеческой интуиции. Оно не может быть передано в виде инструкций, но может быть усвоено через подражание, как в традиционной передаче неявного знания. Анализ стиля или манеры игры показывает, что стиль игры AlphaZero тоже сильно отличался от «традиционных» представлений о «правильной» игре: неожиданные жертвы материала в начале игры, рассчитанные на долгосрочное и весьма неочевидное позиционное преимущество, удивительное чувство динамики, что шахматисты часто называют интуицией или мастерским «чутьем». При этом мы имеем дело с интуицией, не сводимой к алгоритмам, что, опять же, позволяет говорить о появлении некоторого искусственного аналога неявного знания [10].

AlphaZero демонстрирует, что «неявный» характер знания может быть воспроизведен аппаратно и без явного человеческого вмешательства. Возникает вопрос: чем отличается неявное, «человеческое» знание, описанное М. Полани, от «непрозрачного» знания шахматного алгоритма? Формальных отличий может быть очень много, но в действии оба вида знания проявляют себя похожим образом, демонстрируя узнавание паттернов и корреляций, гибкость и «чувство» контекста. Переход от Deep Blue к AlphaZero представляет собой не просто технологический прогресс, а фундаментальный эпистемологический сдвиг, делающий аппарат М. Полани центральным для понимания природы и ограничений современного ИИ. В эпоху тоталь-

ной цифровизации неявное знание казалось одним из очевидных свидетельств человеческой уникальности. Однако быстрая эволюция искусственных систем неявного знания заставляет думать о поиске новых форм симбиоза человека и машины, где неявное знание человека и вычислительная мощь искусственного интеллекта дополняют друг друга.

Эпистемологические последствия и перспективы анализа неявного знания

Эволюция шахматных алгоритмов от Deep Blue к AlphaZero демонстрирует не просто технологический прогресс, а изменение самой логики построения и функционирования когнитивных систем. Переход от явного, заложенного вручную знания к знанию, выработанному автономно в процессе обучения, сопровождается изменением его эпистемологического статуса. В системах вроде AlphaZero знание становится непрозрачным, распределенным и несводимым к пропозициональным правилам, что делает его функциональным аналогом неявного знания в понимании М. Полани.

Этот сдвиг имеет глубокие последствия. Во-первых, он порождает проблему объяснимости. Непрозрачность системы делает невозможным прямой анализ ее «рассуждений», что ставит под вопрос традиционные научные критерии верификации. Во-вторых, он изменяет саму модель передачи знаний от ИИ к человеку. Если из Deep Blue можно было бы (теоретически) извлечь набор правил, то у AlphaZero можно лишь учиться, как у мастера-человека, что является классическим примером передачи неявного знания. Бывший чемпион мира Владимир Крамник, анализируя партии AlphaZero, подчеркивал именно этот аспект.

Концепция неявного знания, предложенная М. Полани, оказывается полезным инструментом для анализа этих трансформаций. Она помогает понять, почему попытки полной формализации и прозрачности сталкиваются с фундаментальными ограничениями как в человеческом познании, так и в искусственном. Таким образом, концепция неявного знания не теряет своей значимости даже в контексте современных технологий. Она указывает на то, что даже в эпоху

цифровизации и автоматизации познание продолжает оставаться практикой, в которой формальные правила и неформальные компетенции неразрывно связаны.

Литература

1. Кун Т. Структура научных революций. М.: Прогресс, 1977. 608 с.
2. Полани М. Личностное знание. На пути к посткритической философии. М.: Прогресс, 1985. 344 р.
3. Райл Г. Понятие сознания. М.: Идея-Пресс, 1999. 408 с.
4. Burns C.S. What Documents Cannot Do: Revisiting Michael Polanyi and the Tacit Knowledge Dilemma // *Information & Culture*. 2021. Vol. 56. № 1. P. 90–104.
5. Campbell M. Knowledge Discovery in Deep Blue // *Communications of the ACM*. 1999. Vol. 42. № 11. P. 65–67.
6. Collins H., Evans R. *Rethinking Expertise*. – Chicago: University of Chicago Press, 2007. 160 p.
7. Polanyi M. *The Tacit Dimension*. – N.Y.: Doubleday & Company, 1966. 104 p.
8. Polanyi M. *Knowing and Being*. – Chicago: University of Chicago Press, 1969. 264 p.
9. Ryle G. *Knowing How and Knowing That* // *Proceedings of the Aristotelian Society*. 1945. Vol. 46. P. 1–16.
10. Sadler M., Regan N. *Game Changer*. – London: Batsford, 2019. 416 p.
11. Shannon C.E. *Programming a Computer for Playing Chess* // *Philosophical Magazine*. 1950. Series 7. Vol. 41. № 314. P. 256–275.
12. Silver D. et al. *Mastering the Game of Go without Human Knowledge* // *Nature*. 2017. Vol. 550. № 7676. P. 354–359.

References

1. Kuhn, T. (1977). *Struktura nauchnykh revolyutsiy* [The Structure of Scientific Revolutions]. Moscow, Progress Publ. (In Russ.). 608 p.
2. Polanyi, M. (1985). *Lichnostnoe znaniye: Na puti k postkriticheskoy filosofii* [Personal Knowledge: Towards a Post-Critical Philosophy]. Moscow, Progress Publ. (In Russ.). 344 p.
3. Ryle, G. (1999). *Ponyatie soznaniya* [The Concept of Mind]. Moscow, Idea-Press Publ. (In Russ.). 408 p.
4. Burns, C.S. (2021). What Documents Cannot Do: Revisiting Michael Polanyi and the Tacit Knowledge Dilemma // *Information & Culture*. No. 56 (1). P. 90–104.
5. Campbell, M. (1999). Knowledge Discovery in Deep Blue // *Communications of the ACM*. Vol. 42. No. 11. P. 65–67.
6. Collins, H., Evans, R. (2007). *Rethinking Expertise*. Chicago, University of Chicago Press. 160 p.

7. *Polanyi, M.* (1966). *The Tacit Dimension*. N.Y., Doubleday & Company. 104 p.
8. *Polanyi, M.* (1969). *Knowing and Being*. Chicago, University of Chicago Press. 264 p.
9. *Ryle, G.* (1945). *Knowing How and Knowing That* // *Proceedings of the Aristotelian Society*. Vol. 46. P. 1–16.
10. *Sadler, M., Regan, N.* (2019). *Game Changer*. London, Batsford. 416 p.
11. *Shannon, C.E.* (1950). *Programming a Computer for Playing Chess* // *Philosophical Magazine*. Ser. 7. Vol. 41. No. 314. P. 256–275.
12. *Silver, D. et al.* (2017). *Mastering the Game of Go without Human Knowledge* // *Nature*. Vol. 550. No. 7676. P. 354–359.

Информация об авторе

Шевченко Александр Анатольевич – Институт философии и права
СО РАН (630090, Новосибирск, ул. Николаева, 8).
shev@philosophy.nsc.ru

Information about the author

Shevchenko, Aleksandr Anatolyevich – Institute of Philosophy and Law,
Siberian Branch of the Russian Academy of Sciences (8, Nikolaev st., Novosibirsk, 630090, Russia).
shev@philosophy.nsc.ru

Дата поступления 02.08.2025.
Принята к публикации 01.09.2025.