DOI: 10.48647/ICCA.2026.95.18.004

Д.Л. Адамова

Токен как лингвистический маркер: особенности формирования и стандартизации ИИ-терминологии в китайском языке

Аннотация. Статья рассматривает проблему перевода на китайский язык термина *token* как пример формирования национальной терминологии искусственного интеллекта (ИИ) в КНР. Материалом исследования послужили нормативные документы, академические публикации, а также тексты, отражающие профессиональную и общественную дискуссию. Цель работы состоит в выявлении принципов, по которым конкурирующие варианты перевода оцениваются профессиональным сообществом и государственными институтами. Предложенные эквиваленты классифицированы по нескольким стратегиям номинации: сохранение английской формы и транскрипция, межотраслевой перенос, описательное обозначение, использование существующих лингвистических терминов, создание лингвистически, концептуально и культурно мотивированных неологизмов. Особое внимание уделено термину 词元 (*цыюань),* рекомендованному в 2026 г. в качестве эквивалента токена, и аргументам в его пользу, включая соответствие принципам точности, понятности, системности и словообразовательной продуктивности. На материале дискуссии выявляются особенности китайской модели научно-технической терминологической стандартизации, которая основана на взаимодействии государственных органов, экспертного сообщества и технологической индустрии. Делается вывод, что выбор *цыюань* является не идеальным решением, а прагматически наиболее устойчивым компромиссом между технической точностью, лингвистической прозрачностью, культурной укорененностью и институциональной поддержкой. Кампания по поиску китайского названия для ИИ выявляет дополнительное измерение проблемы. На официальном уровне терминологическая локализация связывается с языковой самостоятельностью и правом формировать технологический дискурс. Рассмотренный случай отражает стремление Китая перейти к более активному участию в формировании понятийного аппарата глобальной индустрии ИИ.

Ключевые слова: китайский язык, искусственный интеллект, научно-техническая терминология, терминообразование, терминологическая стандартизация, токен, цыюань, неологизмы, языковая политика, цифровой суверенитет.

Автор: *Адамова Дарья Львовна*, преподаватель, МГИМО МИД России. ORCID ID: 0000-0003-2250-8457. E-mail: d.adamova@my.mgimo.ru

D. L.阿达莫娃

作为语言标记的词元：中文人工智能术语创建与标准化的特点

*摘要：*本文以“token”一词的中文翻译为例，探讨了人工智能（AI）中国本土术语创建的相关问题。本研究采用了规范文件、学术出版物以及反映专业和社会讨论的资料。研究旨在明确行业内和政府机构评估各种互相竞争的翻译方案的原则。本文根据以下几种命名策略对提出的对应词方案进行了分类，包括保留英文形式并进行转写；跨行业移植；描述性命名；使用现有语言术语；创造符合语言学、概念和文化背景的新词等。本文重点关注了 2026 年的 token 推荐对应词“词元”，并分析了支持该词的论据，包括其符合准确性、易懂性、系统性和构词效率等原则。在讨论资料的基础上本文着重阐述了中国科技术语标准化模式的特点，该模式基于国家机关、专家群体和技术产业间的互动。结论认为，“词元”并非理想选择方案，而是在技术准确性、语言透明度、文化融入性和制度支持之间最务实和稳固的折衷方案。人工智能中文名称的遴选活动揭示了该问题的另一层面。在官方层面，术语本土化与语言自主性和塑造技术话语的权利息息相关。这一案例反映了中国希望在全球人工智能产业的概念框架构建中发挥更积极的作用。

*关键词：*汉语；人工智能；科技术语；术语构词法；术语标准化；token；词元；新词；语言政策；数字主权

*作者：*达里娅·利沃夫娜·阿达莫娃，莫斯科国际关系学院讲师。ORCID ID：0000-0003-2250-8457；E-mail：d.adamova@my.mgimo.ru

Daria L. Adamova

Token as a Linguistic Marker: The Formation and Standardization of AI Terminology in Contemporary Chinese

Abstract. This article examines the issue of translating the term *token* into Chinese as an example of the development of national terminology for artificial intelligence in the People’s Republic of China. The research drew on regulatory documents, academic publications and texts reflecting professional and public debate. The aim of the study is to identify the principles by which competing translation options are evaluated by the professional community and state institutions. The proposed equivalents are classified according to several naming strategies: retaining the English form and transcription, cross-disciplinary borrowing, descriptive labeling, the use of existing linguistic terms, and the creation of neologisms that are linguistically, conceptually and culturally justified. Particular attention is paid to the term 词元 (*ciyuan*), recommended in 2026 as an equivalent of *token*, and the arguments in its favor, including its compliance with the principles of precision, clarity, systematicity and derivational productivity. Based on the discussion, the paper highlights the distinctive features of the Chinese model of scientific and technical terminological standardization, which is founded on collaboration between government bodies, the expert community and the technology industry. It is concluded that the choice of *ciyuan* is not the ideal solution, but rather the most pragmatically sound compromise

between technical accuracy, linguistic transparency, cultural relevance and institutional support. The campaign to find a Chinese name for AI highlights an additional dimension of the problem. At an official level, terminological localization is linked to linguistic autonomy and the right to shape technological discourse. The case examined reflects China's aspiration to play a more active role in shaping the conceptual framework of the global AI industry.

Keywords: Chinese language, artificial intelligence, scientific and technical terminology, term formation, terminology standardization, token, ciyuan, neologisms, language policy, digital sovereignty.

Author: *Daria L. Adamova,* Lecturer, MGIMO University. ORCID ID: 0000-0003-2250-8457. E-mail: d.adamova@my.mgimo.ru

Введение

10 апреля 2026 г. почти целая полоса в органе ЦК КПК – газете «Жэньминь жибао» – была посвящена обсуждению методов «китаизации» зарубежной научно-технической терминологии. Основной вопрос, поставленный в статье, – способен ли китайский язык в эпоху быстрого прогресса высоких технологий не просто заменить английские обозначения китайскими, а осваивать новые понятия, формировать для них прозрачные и продуктивные китайские наименования и включать их в собственную понятийную систему [Chen Jingwen, 2026]. Незадолго до этого главная партийная газета в своем веб-приложении предложила читателям подобрать для искусственного интеллекта (ИИ) краткое и удобное китайское название [Du Wenwen, 2026], хотя нормативный термин 人工智能 (*жэньгун чжинэн*) задолго до этого был закреплен государственным стандартом GB/T 5271.28–2001 [Xinxi jishu cihui…, 2001] и затем подтвержден базовым терминологическим стандартом GB/T 41867–2022 (введен в действие 1 мая 2023 г.) [Xinxi jishu…, 2022].

За 76 лет до этого, в 1950 г. три публикации И.В. Сталина в газете «Правда» и вмешательство советского вождя в дискуссию по вопросам языкознания превратили, казалось бы, узкопрофессиональный спор марристов (сторонников теорий Н.Я. Марра) и антимарристов в событие со значительными политическими последствиями. Аналогия здесь, безусловно, есть: много лет спустя китайская компартия обращается к лингвистике и мобилизует общественность, потому что язык становится слишком значимым, чтобы оставлять его только лингвистам. Однако на подчеркивании стратегической роли языка прямая аналогия, пожалуй, заканчивается. Китайская дискуссия не является ни доктринальной идеологической кампанией, ни попыткой навязать некую всеобъемлющую марксистскую теорию языка. Ее предмет более узкий и прагматичный, но от этого не менее масштабный: он касается сознательного и суверенного терминологического выбора. Уровень проблемы, стоящей за этим выбором, вполне сопоставим, поскольку термины обычно рождаются там, где впервые разрабатываются соответствующие технические системы, и для Китая, начавшего в этой сфере с догоняющего развития, это создает асимметрию особого рода. Страна, которая уже

сейчас является одним из ведущих разработчиков и пользователей ИИ и технологий больших данных, продолжает при этом думать, рассуждать и проектировать с использованием категорий, пришедших из английского языка.

Одна из таких категорий – токен – кажется маловероятным объектом политического и общественного внимания: это специальная единица, с помощью которой большие языковые модели делят и обрабатывают информацию. Дискуссия о выборе китайского эквивалента английского термина *token*, которую мы анализируем в настоящей статье, представляет собой уникальное окно не только в мир китайской терминологии, но и в то, как Китай превращает терминологическую стандартизацию в один из инструментов утверждения собственного места в глобальной экосистеме искусственного интеллекта и как пытается наполнить технические термины китайским культурным содержанием.

Отечественная литература заложила основу для изучения китайской терминологии. В трудах отечественных китаеведов В.И. Горелова, О.П. Фроловой, А.А. Хаматовой и других специалистов рассмотрены основные механизмы современного китайского словообразования, калькирования, семантического заимствования и системной организации специальной лексики [Горелов, 1984; Фролова, 2011; Хаматова, 2003]. Учеными затрагивались вопросы развития некоторых китайских терминосистем (прежде всего экономической лексики), а также стандартизации и терминоведения [Ивченко, 2020; Очиров, 2013]. С точки зрения психолингвистики проблема восприятия ИИ исследуется в статье Ю.Е. Арекеевой и Ю.И. Бедненко [Арекеева, Бедненко, 2025], вместе с тем работы по китайской терминологии искусственного интеллекта пока носят фрагментарный характер.

Выбор китайского эквивалента термина токен до настоящего времени не становился предметом систематического изучения ни в отечественной, ни в англоязычной зарубежной научной литературе. Это прежде всего объясняется относительной новизной самой темы: официальная приоритетная рекомендация термина 词元 (*цыюань*) и последовавшая за ней профессиональная и общественная дискуссия являются недавними событиями. Настоящая статья стремится восполнить существующий пробел путем анализа конкурирующих китайских эквивалентов термина токен и причин выбора 词元. На основе дискуссии в журнале «Китайская научно-техническая терминология» (中国科技术语, *Чжунго кэцзи шуюй*) и других научных публикаций рассматриваются аргументы и контраргументы китайских лингвистов и терминоведов. Пример с китаизацией «токена» выявляет более общие принципы формирования терминологии искусственного интеллекта в КНР. Сочетание лингвистического анализа с изучением государственной политики, позиции академического сообщества и общественных настроений позволяет рассматривать термины и как языковые единицы, и как элементы китайской суверенной системы производства знания.

Токен в китайском языке: конкурирующие варианты перевода

Для понимания того, почему перевод термина токен на китайский язык оказался столь сложной задачей, необходимо выяснить, какую функцию токен выполняет в

системах ИИ, и прежде всего в сфере генеративного ИИ. В больших языковых моделях (*LLM*), в том числе используемых в *ChatGPT* и моделях семейства *DeepSeek*, токенизация представляет собой процесс разбиения текста на дискретные единицы, которые модель способна обрабатывать. Такими единицами могут быть отдельные символы, слова, знаки препинания, числа или части слов.

Применительно к китайскому языку процесс токенизации имеет собственную специфику. В отличие от русского или английского, где слова в предложениях естественным образом разделяются пробелами, которые могут выступать одним из ориентиров, китайский текст требует более сложной алгоритмической сегментации. В зависимости от используемого токенизатора один китайский иероглиф может составлять отдельный токен, тогда как несколько иероглифов могут объединяться в один токен. Такая вариативность делает само понятие особенно трудным для передачи единым китайским термином, тем более что оно должно охватывать и быстро развивающуюся сферу мультимодального искусственного интеллекта, где токенами могут называться также единицы представления изображений, звука и других данных.

Значение токена не ограничивается лишь технической обработкой информации. Как подчеркивал профессор Университета Цинхуа Ян Бинь, токен превратился в «базовую единицу измерения эпохи искусственного интеллекта», одновременно являясь информационной, вычислительной и расчетной единицей при предоставлении сервисов ИИ [Zhao, 2026].

Если в русском языке термин «токен» в значении единицы обработки данных был заимствован напрямую через транслитерацию и успешно прижился среди специалистов IT-сферы, в китайской практике изначально возникла конкуренция нескольких вариантов перевода. По нашему мнению, необычайно широкий спектр китайских эквивалентов термина токен стал результатом действия нескольких взаимосвязанных факторов.

Во-первых, исходный английский термин отличался высокой многозначностью и ко времени бума ИИ использовался в кибербезопасности, блокчейне, программировании и лингвистике. С распространением больших языковых моделей он приобрел новое, еще не вполне закрепившееся значение, что допускало вариативность перевода.

Во-вторых, широкое распространение ИИ-технологий в Китае вовлекло в обсуждение инженеров, лингвистов, интернет-платформы, специалистов по терминологии и обычных пользователей, причем каждая из этих групп акцентировала внимание на разных аспектах данного понятия.

В-третьих, интерес государства к стандартизации в сфере ИИ постоянно нарастал. Перспектива нормативного закрепления единого китайского эквивалента токена усиливала экспертную конкуренцию и тем самым стимулировала поиск наиболее точного и приемлемого наименования.

Наконец, характерное для китайского языка стремление к семантически прозрачным и морфологически правильно построенным терминам способствовало созданию собственно китайских сложных слов вместо использования транслитерации, что, в свою очередь, привело к появлению множества вариантов наименования.

Конкуренция различных способов передачи иностранного понятия сама по себе не является новым явлением для китайского языка. По крайней мере со второй поло-

вины XIX в. освоение западной научной и общественно-политической лексики нередко сопровождалось периодами «лингвистического плюрализма» – сосуществованием и соперничеством транскрипций, новых семантически мотивированных образований и традиционных слов, получавших новое значение [Lackner et al., 2001; Лукин, 2015].

Предлагаемую ниже классификацию следует рассматривать как рабочую аналитическую модель, а не как жесткую таксономию. Она основана прежде всего на доминирующем механизме, посредством которого каждый китайский эквивалент вошел в оборот: прямом сохранении или фонетической адаптации английского термина, переносе из другой технологической области, использовании существующей лингвистической категории либо создании нового семантически мотивированного сложного слова. Иными словами, классификация описывает различные модели лексического заимствования, семантического переноса и терминообразования, а не строго разграниченные классы слов.

При этом отдельные варианты могут сочетать признаки нескольких моделей. Например, термин может быть заимствован из другой области вычислительной техники и одновременно обладать внутренней формой, значимой в китайской языковой традиции, либо сочетать лингвистическую мотивированность с более широким концептуальным пониманием искусственного интеллекта. Поэтому выделенные категории отражают прежде всего основную стратегию номинации, воплощенную в каждом варианте, а не его единственный источник происхождения, место в хронологии, частотность употребления или степень признания.

Такой подход позволяет перейти от простого перечисления конкурирующих переводов к анализу различных способов, с помощью которых в Китае пытались переосмыслить и адаптировать английский термин, встроив его в существующие системы понятий и используя словообразовательные ресурсы китайского языка.

1. Сохранение английского термина, фонетическое заимствование, каламбурная транскрипция

Формы, сохраняющие английское обозначение либо воспроизводящие его звучание средствами китайского языка, были самым простым решением. Термин токен в неизменном виде широко употреблялся в технической документации, профессиональном жаргоне и даже в официальном дискурсе, поскольку обеспечивал преемственность международного узуса и позволял говорящему не связывать себя ни с одним конкретным китайским определением. Недостаток этого решения был столь же очевиден: оно не помогало читателям, незнакомым с международной технической лексикой, понять содержание понятия и плохо встраивалось в китайскую терминологическую систему.

托肯 (*токэнь*) представлял собой попытку решить проблему путем прямого фонетического заимствования. Подобно таким закрепившимся транскрипциям, как 比特 (*битэ*, «бит»), он позволял избежать ложной точности, возникающей при неудачном смысловом переводе. В качестве альтернативного варианта иероглифической записи иногда встречалась родственная форма 拓肯 (*токэнь*). Однако оба эти варианта могли рассматриваться как временное решение, поскольку они сохраняли «иностранное

звучание» и не раскрывали содержание понятия, а потому уступали сложившейся в китайском языке традиции создавать семантически понятные термины.

偷啃 (*тоукэнь*), буквально «украдкой откусывать» или «тайком грызть», сочетал фонетическое сходство с юмористическим образом токенизатора, который как бы «откусывает» небольшие кусочки от более крупного массива. Согласно другому объяснению, популярному на сетевых форумах, токены незаметно «съедают» деньги пользователей при оплате сервисов ИИ. Эту каламбурную адаптацию следует понимать скорее как фонетико-семантическую игру, чем как серьезного кандидата на научную стандартизацию. Для историка терминологии ценность данного варианта заключается в другом. Он показывает, что дискуссия вышла за пределы профессионального сообщества и стала предметом сетевого языкового творчества.

2. Межотраслевые заимствования из информатики, кибербезопасности и блокчейн-сферы

Речь идет о ситуации, когда ранее закрепившиеся эквиваленты слова токен переносились из смежных технологических областей в дискурс об искусственном интеллекте. В этой категории самым распространенным вариантом был 令牌 (*линпай*). В компьютерных сетях и кибербезопасности он обозначает объект или закодированное удостоверение, предоставляющее доступ или подтверждающее личность, как в сочетании 访问令牌 (*фанвэнь линпай*, «токен доступа»). Его буквальная связь с дощечкой или знаком полномочий делает логику термина легко понятной: обладание 令牌 (*линпай*) позволяет совершить определенную операцию. Однако здесь идея авторизации переносится в контекст больших языковых моделей, где токен никак не связан с авторизацией. Еще один кандидат – 权标 (*цюаньбяо*), «маркер полномочий» или «маркер привилегий», был намного теснее связан с правами доступа и управлением разрешениями, поэтому хуже переносился в сферу ИИ.

С развитием блокчейн-технологий широкое распространение получили термины 代币 (*дайби*, «токен» как цифровая валюта, «замещающая валюта») и 通证 (*тунчжэн*, «токен-сертификат», «обращающийся сертификат»). По мере коммерциализации искусственного интеллекта, когда токен превратился в расчетную единицу при оплате *API* больших языковых моделей, некоторые представители ИИ-отрасли начали использовать именно эти варианты. Однако данные термины несли слишком ярко выраженную ассоциацию с индустрией криптовалют, что затрудняло формирование у широкой аудитории адекватного представления о принципах работы языковых моделей. Сюй Фэй подчеркивал, что заимствования из технологий блокчейна, акцентируя коммерческую ценность токена, игнорируют его лингвистическую и вычислительную природу в технологиях ИИ. Кроме того, они несут ярко выраженную ассоциацию с криптоиндустрией и способны вызвать у широкой аудитории ошибочное представление о фундаментальных принципах работы моделей ИИ [Xu, 2026]. Нежелательная семантическая связь с технологией блокчейна ранее отмечалась и другими исследователями. Как писала Лю Цзиньтин, «в китайской терминологической практике различным понятиям, как правило, присваиваются разные китайские названия с максимально возможным соблюдением принципов однозначности и научности» [Liu, 2020, p. 67].

3. Описательные обозначения через понятия метки и знака

Варианты в этой категории представляли токен как метку, знак, обозначение или символ, при этом не содержали указания на его статус как вычислительной единицы.

Формы 标记 (*бяоцзи*), 记号 (*цзихао*) и 符号 (*фухао*) возвращались к более старому лексическому значению слова токен как метки или знака. Они были менее уязвимы для отраслевых искажений, чем 令牌 (*линпай*) или 代币 (*дайби*). Кроме того, они достаточно хорошо соответствовали терминологии компиляторов: лексический анализатор преобразует входной поток в последовательность распознаваемых элементов, которые в широком смысле можно описать как «метки» или «символы».

Однако широта семантического поля этих слов была одновременно их преимуществом и недостатком. Например, 标记 (*бяоцзи*) может обозначать практически любую метку, тег, знак или аннотацию. Изначально первый иероглиф означал «верхушку» или «конец дерева». В традиционном написании иероглиф 標 (бяо) состоит из ключа 木 «дерево» и фонетического компонента 票. Постепенно происходило расширение значения: сначала иероглиф стал обозначать выступающую и хорошо заметную часть предмета, затем заметный шест, позднее ориентир, знак или указатель. На этой основе у иероглифа появились глагольные значения «обозначать» и «помечать». Применительно к токену *бяоцзи*, передавая идею маркирования, в то же время не показывает, идет ли речь о входном фрагменте, элементе словаря, конкретном употреблении, числовом идентификаторе или управляющем символе. 记号 (*цзихао*) столь же обобщен и может создавать впечатление внешне добавленного обозначения, а не единицы, полученной в результате токенизации. 符号 (*фухао*) верно подчеркивает символическую форму, но значительно шире понятия токена: каждый токен может функционировать как символ, однако далеко не каждый символ является токеном.

Среди предлагавшихся переводов встречались и составные выражения из четырех иероглифов, призванные точнее передать отдельные свойства токена.

词汇标记 (*цыхуэй бяоцзи*), «лексический маркер», сужал значение, помещая единицу в лексикон. Он был более информативен в лексическом анализе, однако применительно к сфере ИИ, тем более в мультимодальных моделях, мог создавать излишне тесную ассоциацию со словарным составом естественного языка.

语义符号 (*юйи фухао*), «семантический символ», представлял собой попытку отличить токены от бессмысленных физических меток, подчеркнув их роль в передаче или порождении значения. Проблема состоит в том, что токен не обязательно обладает самостоятельно извлекаемым семантическим содержанием. Часть слова, знак препинания или служебный токен (обозначающий, например, начало либо конец последовательности) могут приобретать функцию только в сочетании с другими единицами. Значение часто возникает композиционно и контекстуально, а не содержится внутри каждого отдельного токена. Поэтому *юйи фухао* может сигнализировать о самостоятельной смысловой содержательности токена, которая не является его обязательным свойством. Кроме того, стремление к смысловой точности делало такие варианты из четырех иероглифов слишком громоздкими, тогда как базовый технический термин должен быть кратким и удобным в употреблении.

4. Существующие лингвистические термины и дисциплинарные приближения

При передаче значения термина «токен» делались попытки использовать обозначения, заимствованные из корпусной и структурной лингвистики либо основанные на традиционных единицах морфологического анализа.

Наиболее теоретически обоснованные альтернативы происходили из различия между *type* и *token* в философии, семиотике и корпусной лингвистике. В этой традиции *type* представляет собой повторяемую абстрактную форму, а *token* – конкретное употребление этой формы. Различие проще всего показать на примере фразы *rose is a rose*. В ней четыре отдельных употребления слов, то есть четыре токена, но только три разных слова: *rose, is* и *a*. Иными словами, токеном называется каждое конкретное появление слова в тексте, а типом – само слово как повторяющаяся единица языка [Wetzel, 2009].

В китайской научной литературе это различие обычно передавалось парой 类符 (*лэйфу*, «тип-форма») и 形符 (*синфу*, «форма-употребление»), хотя порядок компонентов и конкретная терминология могли различаться. У *синфу* есть важное концептуальное преимущество: он сохраняет значение конкретного экземпляра, присущее исходному лингвистическому понятию токен.

词例 (*цыли*), буквально «пример употребления слова» или «словоупотребление», выражает ту же идею более прозрачно. При подсчете в корпусе три появления одного и того же слова могут быть описаны как три *цыли*, относящиеся к одному словесному типу. Этот термин понятнее, чем *синфу*, однако сохраняет ограниченность, связанную со словом: токены современных моделей не обязательно являются словами или даже лингвистически целостными частями слов.

Более глубокая проблема заключается в том, что употребление слова токен в сфере ИИ уже не полностью совпадает с традиционным понятием конкретного экземпляра. В документации моделей токен может обозначать элемент словаря токенизатора, конкретное употребление этого элемента в последовательности, присвоенный ему числовой идентификатор или, по метонимическому переносу, единицу измерения контекста и расчета стоимости. Поэтому термин, созданный исключительно для выражения значения «экземпляр», сохраняет важную часть истории понятия, но не охватывает все случаи его современного технического употребления.

语素 (*юйсу*), «морфема», был еще одним привлекательным приближением, поскольку морфемы традиционно рассматриваются как минимальные значимые единицы. Однако как общий эквивалент он технически неточен. Границы токенов определяются токенизатором и не обязаны совпадать со строгими морфологическими границами. Одна морфема может быть разделена на несколько токенов, тогда как один токен может включать несколько морфем либо произвольную последовательность символов или байтов. Использование *юйсу* означало бы замену вычислительной категории, зависящей от модели, закрепленной лингвистической категорией, определяемой по иным критериям.

Более поздний компромиссный вариант 符例 (*фули*), «экземпляр или пример символа», представлял собой попытку сохранить идею конкретного употребления, не ограничивая объект словами. Теоретически он экономичен, однако сравнительно малоизвестен и хуже образует естественные инженерные сочетания вроде «стоимость токенов» или «бюджет токенов».

5. Лингвистически мотивированные неологизмы

Данная группа неологизмов связывала токен со словом, письменным знаком или элементарной единицей текста. Определяющая стратегия заключалась в сочетании компонента, связанного с языком, с существительным, обозначающим элементарную составную часть.

词元 (*цыюань*) в конечном счете стал самым сильным кандидатом. Его структура компактна: 词 (*цы*) связывает понятие со словами или языком, тогда как 元 (*юань*) указывает на элемент или базовую единицу. Термин уже был указан среди китайских переводов слова токен в обзоре Лю Цзиньтин 2020 г. [Liu, 2020], поэтому решение 2026 г. не предложило нового, а выделило в качестве ведущего один из ранее предложенных вариантов. Более подробно мотивы выбора и возражения критиков мы разберем ниже.

Альтернативные решения внутри этой группы состояли в изменении первого или второго иероглифа. 语元 (*юйюань*), «языковая единица», избегал узкого отождествления со словом, но становился слишком широким: фонемы, морфемы, словосочетания и предложения также являются языковыми единицами. 词符 (*цыфу*), «словесный знак», подчеркивал символическую форму, но неясно выражал статус единицы и оставался привязанным к тексту. 字元 (*цзыюань*), «иероглифическая/символьная единица», мог в первую очередь использоваться там, где токены совпадали с отдельными символами. Он не мог охватить токены, представляющие собой многосимвольные последовательности или фрагменты байтового представления. Кроме того, на Тайване 字元 уже закрепился в информатике как перевод английского *character* (знак), то есть как обозначение отдельной буквы, цифры, знака препинания или иного символа, обрабатываемого компьютером. Официальные материалы САР Гонконг приводят для *character* наряду с 字符 (*цзыфу*) также и 字元 (*цзыюань*). Использование того же слова для *token* создавало бы терминологическую путаницу.

Наиболее важными соперниками были 符元 (*фуюань*, «символическая единица») и его «зеркальный» вариант 元符 (*юаньфу*). Хотя термины потенциально были способны обозначать и мультимодальный контекст, их недостатками были меньшая известность и возможная ассоциация с символическим ИИ, способная ошибочно создать впечатление, будто современные модели работают на основе символической архитектуры. Символический ИИ (*symbolic AI*) был главным направлением исследований до конца 1980-х гг. Он решал задачи с помощью заранее заданных понятий, фактов и логических правил. Нынешние модели нейросетевого ИИ не получают готового набора правил, а обучаются на больших массивах данных.

6. Концептуально мотивированные неологизмы

Еще одна группа терминов пыталась описать токен не через его лингвистическое происхождение, а через предполагаемую функцию в искусственном интеллекте.

智元 (*чжиюань*), «единица интеллекта», представлял токен как элементарную составляющую машинного интеллекта. Он был кратким, запоминающимся и легко включался в производные выражения. Незадолго до официального решения он также получил поддержку в отраслевых дискуссиях об ИИ. Однако *чжиюань* явно преувеличивает содержание отдельного токена. Токен является единицей, несущей инфор-

мацию и обрабатываемой интеллектуальной системой, но сам по себе он не обладает интеллектом. Поэтому при буквальном прочтении эта форма способна немотивированно перенести особенности ИИ на внутреннее свойство самого токена. 新智元 (*синь чжиюань*), «новая единица интеллекта», усиливал ту же интерпретацию и добавлял экономическую ассоциацию через 元 (*юань*), создавая впечатление некоей новой валюты, однако в целом этот неологизм имел очевидный рекламный оттенок.

Термин 模元 (*моюань*), «единица модели», был сконструирован не без изящества. Здесь 模 (*мо*) одновременно отсылает к 大模型 (*да мосин*, «большие модели») и 多模态 (*домотай*, «мультимодальность»), тогда как 元 (*юань*) обозначает базовую единицу. Предложение напрямую отвечало на главный недостаток, приписываемый *цыюань*, поскольку *моюань*, семантически не связанный со словом, мог обозначать текст, изображения, звук, видео или входные данные агента без необходимости метафорически переопределять все эти элементы как «слова».

Возражение состояло в том, что токен обрабатывается моделью, но не обязательно является составной частью ее архитектуры. *Моюань* может восприниматься как «элемент модели», создавая путаницу с модулями, параметрами или архитектурными компонентами. Кроме того, он связан с нынешней популярностью понятий «большая модель» и «мультимодальность», тогда как технический термин в идеале должен быть более устойчивым и сохранять применимость независимо от смены архитектурных подходов.

数元 (*шуюань*), «единица данных» или «числовая единица», подчеркивал дискретизацию и машинную обработку. Он указывал, что токены преобразуются в числовые идентификаторы или представления, однако мог создавать путаницу между самим токеном и его идентификатором либо вектором. Токены не обязательно являются числами; вычислительная система лишь присваивает им числа или векторы. Кроме того, понятие «единица данных» настолько широко, что не позволяет отличить токены от байтов, записей, признаков или иных представлений.

义节 (*ицзе*), приблизительно «смысловой сегмент», трактовал токенизацию как семантическое членение. Он отражал интуитивное представление о том, что модели в конечном счете используют токены для построения или обработки значения. Его недостаток был сходен с обозначенной выше проблемой 语义符号 (*юйи фухао*): многие токены не соответствуют самостоятельным смысловым сегментам.

Другие предложения более явно следовали той же стратегии. 知元 (*чжиюань*), «единица знания», защищался как мультимодальный носитель фрагментов знания, однако токен не обязательно кодирует отдельную единицу знания. 智子 (*чжицзы*), «частица интеллекта», заимствовал продуктивный научный суффикс 子 (*цзы*) из таких слов, как 原子 (*юаньцзы*, «атом») и 电子 (*дяньцзы*, «электрон»), придавая термину интуитивную выразительность, но вновь приписывал самой единице интеллект. Еще один похожий вариант 词节 (*цыцзе*), «словесный сегмент», подчеркивал процесс сегментации, однако оставался привязанным к тексту. 算元 (*суаньюань*), «вычислительная единица», и 算码 (*суаньма*), «вычислительный код», выдвигали на первый план вычисление и измерение, но были недостаточно специфичны: слишком многие объекты в вычислительной системе являются вычислительными единицами или кодами.

Все эти формы выявляют повторяющуюся проблему. Каждая из них выбирала одну заметную роль – интеллект, знание, моделирование, данные, значение или вычисление – и превращала ее в сущность токена. Их кажущаяся точность нередко достигалась ценой чрезмерного сужения понятия или приписывания ему свойств, которые в действительности принадлежали более широкой системе, а не самому токену.

7. Философские и культурные неологизмы

Пожалуй, наиболее амбициозные в культурном отношении предложения исходили от сторонников включения традиционной китайской философии в терминологию искусственного интеллекта.

В 2024 г. Чжан Цинжуй (张庆瑞), известный тайваньский специалист по квантовым вычислениям и физике магнитных материалов, предложил термин *даоюань* (道元, «единица Дао») [Zhang Qingrui, 2024]. По его мнению, подобно тому, как Дао порождает все сущее, «искусственный интеллект также начинает работу с генерации одного токена, затем двух, трех и бесконечного множества токенов». В этом предложении содержалась прямая ссылка на известную формулу из «Дао дэ цзина»: 道生一，一生二，二生三，三生万物 – «Дао рождает Одно. Одно Двоих рождает. Двое рождают Трех, Трое рождают все сущее»[1]. Чжан Цинжуй полагал, что соединение понятия токен с идеями Лао-цзы «не только свяжет древность с современностью и объединит китайскую и западную мысль, но, что еще важнее, обеспечит удачное сочетание звучания и смысла» [Zhang Qingrui, 2024, p.40].

Далее Чжан Цинжуй противопоставляет свою идею формуле американского физика-теоретика Джона Уилера *it from bit*, согласно которой физическую реальность в конечном счете можно понимать через базовые единицы информации – биты. Чжан утверждает, что бит слишком ограничен, поскольку выражает лишь простое бинарное различие «да или нет». Вместо этого он предлагает понятие *даоюань* – единицы, способной, с его точки зрения, представлять не только информацию, но также смысл, факты и логические связи. Предложенная Чжан Цинжуем формула 万物皆道元 (*ваньу цзе даоюань*) «все вещи суть даоюань (элементы Дао)» представляется как китайская философская альтернатива формуле Уилера *it from bit* [Zhang Qingrui, 2024, p. 43]. В целом сопоставление носило риторический, а не содержательный технический характер: его целью было наделить цифровую единицу узнаваемой китайской философской генеалогией, а не определить, как токен функционирует в модели ИИ.

Эту идею в переработанном и существенно дополненном виде Чжан Цинжуй изложил в совместной статье с Чжан Чжунхуном в журнале «Китайская научно-техническая терминология» (中国科技术语/*Чжунго кэцзи шуюй*) [Zhang Zhonghong, Zhang Qingrui, 2025]. Авторы предлагают расширенную формулу в стиле традиционного китайского трактата: 物化道元，道生万物，万物皆道元，一切来自道元 (*ухуа даоюань, дао шэн ваньу, ваньу цзе даоюань, ице лайцзы даоюань*), которую можно перевести так: «вещи преобразуются в *даоюани*, Дао порождает все сущее, все сущее есть *даоюани*, все происходит из *даоюаней*». Здесь *даоюань* выступает уже не просто как китайский перевод слова токен, а становится мистической первоосновой, с по-

[1] Перевод Евгения Алексеевича Торчинова (1956–2003). См. Даосизм. «Дао-Дэ цзин». 2-е изд. / Пер. Е.А. Торчинова. СПб.: «Азбука-классика»; «Петербургское Востоковедение», 2004. С. 229.

мощью которой ИИ представляет мир и порождает новое знание. Стоит упомянуть еще один интересный сюжет, навеянный китайской классикой. В рассуждениях о моделях ИИ в статье широко используется идея 易 (*и*, «перемены») – ключевое понятие «И цзин» (Книги перемен): сложное можно разложить на простые элементы, а затем, изменяя и комбинируя их, создавать новые формы. Именно так авторы предлагают понимать и генеративный ИИ. Особенности китайского языка – многозначность иероглифов и гибкие связи между ними – они также рассматривают как возможную универсальную модель обращения со смыслом. В результате предложение нового перевода превращается в гораздо более широкую попытку описать работу генеративного ИИ средствами китайской философии и даже поэзии. Статья заканчивается следующим стихотворением, которое авторы пытались выдержать в классической традиции:

取日月为假, 欲名本绵绵, 借地天同尘, 俾思以空閒。宇寰非位元, 魂魄亦无源, 唯待道元出, 方知天地言[2]。

[Zhang Zhonghong, Zhang Qingrui, 2025, p. 63].

* * *

Луна и солнце – свет на краткий миг,
Не выразят немую бесконечность,
И путь земной, где разум наш постиг
Души глубины и ее же вечность.
Бездушен бит. Вселенная вовек
В пустых нулях свободно не задышит,
Лишь голос Дао, смыслов оберег,
Вернет язык, который разум слышит.
(*Поэтический перевод автора статьи)*

Идея *даоюань* хорошо воспринимается интеллектуальной китайской аудиторией, поскольку предлагает не просто термин, а помещает генеративный ИИ в китайский космологический нарратив об элементарных началах и порождении множественности. Но именно в этом состоит главный терминологический недостаток. Элемент Дао не обозначает никакого наблюдаемого технического процесса, отличающего токенизацию от других операций. Аналогия между тем, как «Дао порождает все множество вещей», и тем, как ИИ шаг за шагом порождает новые токены, философски выразительна, но в техническом отношении поверхностна, поскольку ничего не говорит о свойствах самого токена.

Утверждение авторов о том, что английская и китайская токенизация принципиально различаются вследствие предполагаемой простоты одной письменности и сложности другой, также искажает процесс, который зависит от обучающих данных,

[2] Примерный подстрочник: Вижу в Солнце и Луне нечто заимствованное, желая назвать изначальный корень непрерывно длящимся. Заимствую у Неба и Земли, чтобы слиться с пылью, дабы мысли обрели пустоту и покой. Вселенная – не бит, земная и небесная души лишены истока. Остается ждать, когда появится *даоюань*, только тогда познаем слова Неба и Земли.

состава словаря и устройства токенизатора, а не только от особенностей текста. В сочетании с обращением Чжан Чжунхуна и Чжан Цинжуя к уникальности китайской философии такое противопоставление приобретает элементы культурного национализма. Специфика китайского языка, значимая с точки зрения ИИ-моделей, интерпретируется как свидетельство его особого смыслового богатства. Китайский описывается как более трудный язык для токенизации, но одновременно и «более богатый по смыслу» (意义更丰富, *ии гэн фэнфу*) [Zhang Zhonghong, Zhang Qingrui, 2025, p.59]. Это уже придает техническому сравнению явный ценностный оттенок.

К числу культурно мотивированных вариантов передачи смысла токена относились 筹元 (*чоуюань*) и 筹 (*чоу*). Их основой стал исторический китайский образ 筹 (*чоу*) – счетных палочек, обычно бамбуковых, которые использовались для вычислений[3]. 筹元 (*чоуюань*) пытается представить токен как счетную или вычислительную единицу с китайской культурной родословной, не философской, как *даоюань*, а связанной с традиционной практикой счета. На этой основе предлагался вариант 筹码 (*чоума*), усиливающий значение жетона или единицы стоимости, а также 词筹 (*цычоу*), по сути «словесная счетная палочка», соединявшая понятие языковой единицы с традиционным образом счетной палочки. Однако культурная мотивация всех вариаций была яснее, чем концептуальная определенность: термины представляли токены как счетчики, используемые для вычислений или обмена, тогда как токены прежде всего являются единицами представления, определяемыми моделью.

Хотя культурно окрашенные варианты так и не получили широкого распространения, сам факт их появления отражал важную тенденцию дискуссии – стремление встроить китайские традиции в терминологию искусственного интеллекта и тем самым подчеркнуть интеллектуальный суверенитет страны в области технологий, где Китай все активнее занимает лидирующие позиции.

Почему был выбран 词元 (*цыюань*): лингвистические и терминологические аргументы

До официальной рекомендации использовать термин 词元 (цыюань) китайские государственные органы, как правило, сохраняли английскую форму token, иногда сопровождая ее пояснительным определением, а с 2025 г. цыюань стали представлять в качестве разъясняющего эквивалента – токен, обычно называемый цыюань (token, 也就是大家通常所说的词元; token, е цзю ши дацзя тунчан со шо дэ цыюань) [Wenzi shilu…, 2025]. Такую практику можно интерпретировать как проявление институциональной и семантической осторожности, а также поэтапной стандартизации новых терминов.

[3] Согласно китайским письменным источникам и археологическим находкам, счетные палочки существовали по меньшей мере в V в. до н. э. в начале периода Сражающихся царств. Счет на палочках позволял совершать арифметические действия, а также, как свидетельствуют более поздние сочинения, выполнять сложные математические операции, включая возведение в степень и извлечение корня, обозначать и использовать в вычислениях ноль, отрицательные числа и дроби.

Сохранение английской формы обеспечивало прямое соответствие международной технической документации, моделям ценообразования *API* и отраслевым метрикам, одновременно позволяя избежать преждевременного выбора китайских вариантов. Кроме того, до распространения *ChatGPT* (*OpenAI)* и аналогичных сервисов в Китае токенизация обсуждалась главным образом в научных работах, стандартах по обработке естественного языка и технической документации, а не в политико-нормативных актах. Даже после появления генеративного ИИ официальные документы, начавшие выходить с 2023 г., описывали технологию на более общем уровне и не нуждались в обозначении внутренней единицы обработки модели.

Примечательной особенностью процесса стандартизации является влияние технологических гигантов. Когда команда *Alibaba,* работающая над моделью *Qwen*, создала высокооптимизированный токенизатор для решения проблемы «токенной неэффективности» китайского языка в западных моделях, во внутренней документации и последующих технических публикациях она широко использовала понятийную систему *цыюань*. Тем самым индустрия ИИ ускоряла принятие общегосударственного стандарта, соответствующего де-факто сложившейся практике. Эти симбиотические отношения между органами стандартизации и технологическими компаниями являются одной из определяющих черт китайской лингвистической среды в сфере ИИ, поскольку обеспечивают не только академическую и нормативную обоснованность терминологии, но и ее практическую применимость в индустрии.

Всекитайский комитет по утверждению научно-технической терминологии 25 марта 2026 г. рекомендовал 词元 (*цыюань*) в качестве предпочтительного китайского эквивалента термина токен, однако пока лишь для пробного использования до окончательного утверждения [Quanguo kexue jishu…, 2026]. Незадолго до этого общегосударственный норматив по языку и письменности GF 0031–2026, опубликованный 2 марта 2026 г. Министерством образования и Государственной комиссией по языку и письменности КНР, также включал *цыюань*. В разделе 3.1.4 词元 (*цыюань*) соотносится с токеном и определяется как «минимальная базовая единица исходного текста, которая может быть идентифицирована и смоделирована». Далее уточняется, что такой единицей может быть слово, субслово, знак или элементарный мультимодальный код, а также закрепляется вариант 词元化 (*цыюаньхуа*) для обозначения токенизации, то есть «процесса сегментации текста на токены» [Rengong Zhineng…, 2026, p. 1].

Хотя GF 0031–2026 не является национальным стандартом, он относится к отдельной категории официальных норм в области языка и письменности (语言文字规范, *юйянь вэньцзы гуйфань*). Его адресатами являются разработчики и операторы корпусов, компании-разработчики ИИ, научные и образовательные учреждения, издатели технической документации и государственные учреждения, если они используют соответствующую терминологию. Таким образом, документ придает *цыюань* определенную степень государственного авторитета, которую ни профессиональное употребление, ни поддержка со стороны экспертов сами по себе обеспечить не могли.

Хронология событий указывает на сближение двух частично независимых механизмов – разработки языковых норм и унификации научно-технической терминологии. При этом по одному каналу *цыюань* уже вошел в официальный нормативный

документ, тогда как по-другому он все еще представлялся обществу как предпочтительное обозначение, требующее дальнейшей апробации. Такая рассинхронизация во многом вызвана особым положением искусственного интеллекта и скоростью его развития в Китае. Практическая потребность в единой терминологии возникла раньше, чем два механизма стандартизации успели полностью согласовать свои процедуры.

Официальное обоснование выбора *цыюань* строилось на нескольких взаимодополняющих аргументах. Заместитель председателя Всекитайского комитета по утверждению терминологии в области компьютерных наук и технологий и руководитель его подкомитета по искусственному интеллекту Чэнь Силинь объяснил, что термин наиболее точно выражает «первоначальную сущность токена как базовой семантической единицы языка» и поэтому лучше всего соответствует его исходной функции в искусственном интеллекте [Gao, 2026].

Особое внимание разработчики уделили наиболее серьезному возражению против нового термина – его применимости в эпоху мультимодальных моделей. В пояснениях отмечалось, что по мере перехода больших моделей от работы исключительно с текстом к обработке изображений, аудио и других типов иероглиф 词 (*цы*) в составе 词元 (*цыюань*) перестает обозначать «слово» в узком лингвистическом смысле и приобретает более широкое значение. Иначе говоря, дискретные базовые единицы нетекстового характера рассматриваются как своего рода «обобщенные слова» (广义的词, *гуан'и дэ цы*).

Еще одним достоинством перевода считалась его стилистическая согласованность с уже сложившейся китайской терминологией ИИ, а также то, что *цыюань* является компактной конструкцией из двух иероглифов, соответствующей традиционным моделям китайского словообразования и одновременно обеспечивающей необходимую научную точность. Выбор первого компонента 词 согласуется и с уже сложившейся системой китайской лингвистической терминологии, в которой он используется, к примеру, в таких понятиях, как 词根 *(цыгэнь*, «корень»), 词缀 (*цычжуй,* «аффикс»), 词形 (*цысин*, «словоформа») и 词源 (*цыюань*, «этимология»). Сама эта терминологическая система во многом формировалась в процессе освоения западного языкознания и под значительным влиянием японской научной терминологии конца XIX – начала XX в. Поэтому включение *цыюань* в уже привычный ряд терминов с компонентом 词 (*цы*) на новом этапе воспроизводило ранний механизм адаптации иностранных научных понятий средствами китайского словообразования. Благодаря этому новое обозначение воспринималось не как чуждый неологизм, а как элемент знакомого терминологического ряда.

Дискуссия вокруг перевода *token* неизбежно вызвала обращение к знаменитому принципу 信达雅 (*синь, да, я*) – «верность, выразительность и изящество», сформулированному философом и переводчиком Янь Фу (严复, 1854–1921). Три критерия были предложены еще в 1898 г. в предисловии к китайскому переводу книги Томаса Генри Гексли *Evolution and Ethics* («Эволюция и этика») и по-прежнему служат одним из важнейших ориентиров китайской переводческой традиции. Сюй Фэй специально рассмотрел термин *цыюань* через призму этой классической концепции и пришел к выводу, что именно данный вариант наиболее успешно сочетает все три принципа.

С точки зрения верности оригиналу новый термин максимально точно соответствует основному определению токена в сфере искусственного интеллекта как минимальной языковой или информационной единицы, обрабатываемой большими моделями и служащей базовым носителем понимания, генерации и вычислений. Если иероглиф 词 (*цы*) сохраняет связь с языковым происхождением понятия, то 元 (*юань*) адекватно передает его техническую сущность как минимальной базовой единицы. Его понятность и выразительность определяются тем, что *цыюань* представляет собой естественное двусложное слово современного китайского языка, легко воспринимаемое носителями. К тому же термин обладает высокой словообразовательной и сочетательной продуктивностью. На его основе без труда формируются: 词元化 (*цыюаньхуа*) – токенизация; 词元长度 (*цыюань чанду*) – длина токена или длина последовательности; 词元计费 (*цыюань цзифэй)* – тарификация по количеству токенов. Именно такая способность порождать единообразную терминологическую систему, по мнению Сюй Фэя, устраняет коммуникационные трудности, возникшие из-за прежнего разнообразия переводов, и существенно облегчает распространение знаний об искусственном интеллекте.

Наиболее интересной частью анализа является интерпретация категории 雅 (*я*, изящество). Сюй Фэй подчеркивает, что иероглиф 元 (*юань*) обладает глубокой философской нагрузкой в китайской культуре. Он ассоциируется с такими понятиями, как исток, основание, первооснова, ядро, то есть обозначает фундаментальную единицу существования вещей. Использование этого иероглифа в терминологии ИИ отражает базовую логику функционирования технологии и одновременно сохраняет связь с китайской культурной традицией, позволяя избежать грубости буквального перевода. Кроме того, строгая и нормативная форма *цыюань* выгодно отличается от отраслевых вариантов вроде 代币 (*дайби*, «токен» в значении цифровой валюты) или различных разговорных импровизаций, соответствуя требованиям профессиональной научной терминологии [Xu, 2026].

Вместе с тем из статьи следует, что применение концепции Янь Фу к терминологии ИИ имеет определенные особенности и ограничения. Сюй Фэй по сути «осовременивает» категорию 雅, полагая, что сегодня в ее основе находится не словесная изысканность, а профессиональная строгость, лаконичность, сдержанность и соответствие характеру научной дисциплины [Там же]. Сам Янь Фу понимал под критерием 雅 (*я*) литературное изящество классического китайского языка, который был способен сделать западные идеи приемлемыми для китайской образованной элиты конца XIX в.

Современная научная терминология требует уже иного понимания «изящества». Сегодня оно заключается не в переводе новых технологий на язык литературной утонченности, а прежде всего в концептуальной ясности, системной согласованности и культурной укорененности. Именно поэтому, заключает Сюй Фэй, достоинство *цыюань* определяется не столько его эстетическими качествами, сколько способностью органично встроиться одновременно в существующую систему научной терминологии и в китайские философско-культурные представления [Xu, 2026].

Дискуссия после стандартизации: пределы терминологического консенсуса

Рекомендация *цыюань* в качестве официального приоритетного варианта вызвала очередную волну экспертного обсуждения. Этой теме был посвящен специальный раздел журнала 中国科技术语 (№ 3 за 2026 г.), авторы которого сосредоточились уже не столько на выборе конкретного перевода, сколько на более общем вопросе – по каким принципам должна в дальнейшем формироваться китайская терминология искусственного интеллекта.

Первая группа разногласий связана с тем, что именно должен обозначать термин. Фэн Чживэй исходит из обработки естественного языка и рассматривает токен как дискретную единицу хранения, обработки и обмена информацией. Он утверждает, что *цыюань* может выйти за рамки своего лингвистического происхождения и охватывать единицы изображений, аудио и видео, и даже рассматривает количество токенов как возможную меру объема услуг ИИ и уровня его интеллектуальных возможностей [Feng, 2026, p. 4–5]. Лин Чжэньхуа занимает более осторожную позицию, видимо близкую к мнению части инженерного сообщества. Одобряя выбор *цыюань*, он подчеркивает, что токен не обязательно является словом, и ставит вопрос о том, должны ли мультимодальные производные сохранять ту же основу, как в 语音词元 (*юйинь цыюань*, «речевой токен») и 视觉词元 (*шицзюэ цыюань*, «визуальный токен»), или для них нужны отдельные обозначения [Ling, 2026, p. 13–14]. Чжан Цзюньпин подходит к той же проблеме с другой стороны: если токены теперь используются для текста, изображений и звука, следует ли вообще сохранять в их названии акцент на языке? Предлагаемый им термин 知元 (*чжиюань*, «единица знания») представляет токен прежде всего как базовый носитель знания, а не как фрагмент языка [Zhang Junping, 2026, p. 6].

Второе направление дискуссии касается преемственности между научными дисциплинами. Юань Цзяхун утверждает, что терминология ИИ не может игнорировать более старое различие *type–token*: *type* – это абстрактная категория, тогда как *token* – одно конкретное ее воплощение. Поскольку реальные границы токенов зависят от используемого токенизатора, термин *цыюань* может вводить в заблуждение, создавая одновременно представление об устойчивой связи со словами и об абстрактной языковой единице, сопоставимой с лексемой [Yuan, 2026, p. 11–12]. Чэнь Гохуа предлагает парные термины 符类 (*фулэй*) и 符元 (*фуюань*) соответственно для *type* и *token* [Chen Guohua, 2026, p. 9–10]. Высказываясь в пользу *фуюань*, Кэ Пин и Хэ Кайда добавляют к лингвистической традиции технологические аргументы, делая акцент на функциональном понимании токена как дискретной единицы машинной обработки, применимой не только к тексту, но и к изображению и звуку [Ke, He, 2026, p. 12–13].

В основе этой части дискуссии лежит вопрос о том, должна ли китайская терминология ИИ сохранять понятийные структуры, унаследованные от прежних дисциплин, или прежде всего описывать непосредственные операции современных моделей. Развилка начинается с задаваемого вопроса: искать ли для token ближайшую языковую аналогию или, прежде всего, исходить из того, чем эта единица является для системы ИИ.

Третье направление дискуссии переносит внимание с конструирования понятий на социальное закрепление терминов. Ли Юймин утверждает, что окончательная судьба 词元 будет определяться не только теоретическими аргументами, но и реальным языковым употреблением [Li Yuming, 2026, p. 8–9]. Вань Цюань формулирует этот прагматический аргумент еще яснее: скорее всего закрепится не обязательно самый точный термин, а тот, который легче всего понять, запомнить и воспроизводить в учебниках, университетских аудиториях, СМИ и документации цифровых платформ. Он сравнивает терминологическую конкуренцию с войной форматов видеокассет: *Betamax* обладал техническими преимуществами, однако *VHS* победил, поскольку был дешевле, доступнее и поддерживался более крупной экосистемой. «*Цыюань* – это прежний *VHS*: хотя технически он несовершенен, рынок наделит его жизненной силой в процессе распространения», – уверен ученый [Wan, 2026, p. 10–11]. Фань Вэньцзе дополняет эту картину закрепления термина анализом носителей, которые превращают рекомендацию в устойчивую норму, прежде всего авторитетных словарей и образовательных материалов [Fan, 2026, p. 7].

Главный промежуточный итог напряженной дискуссии состоит в том, что перевод слова токен оказался только первым шагом. Теперь центр внимания смещается к более сложной задаче – созданию связной и устойчивой китайской системы понятий для ИИ. За дискуссией все отчетливее просматривается политическое измерение. Создание терминологии ИИ связывается с укреплением национальной системы стандартов, научной и культурной самостоятельности, права Китая формировать собственный дискурсивный порядок в ключевой сфере, обеспечивающей технологическое лидерство.

Государство, язык и формирование национальной терминологии ИИ

Политическое преломление экспертной дискуссии наиболее отчетливо проявилось на страницах и в веб-приложениях «Жэньминь жибао». В серии материалов, опубликованных весной–летом 2026 г., выбор китайских обозначений для *AI*, *token*, *agent* и *LLM* представлялся как часть более широкого проекта – закрепления китайского языка в качестве полноценного языка новой технологической сферы. При этом в трактовке со стороны главного партийного органа сосуществовали два подхода. Лингвистический подход делал акцент на понятийной точности и проверке новых терминов реальным употреблением, тогда как политико-культурный рассматривал их локализацию как выражение цифрового суверенитета и самостоятельного формирования технологической повестки.

Отличительной особенностью одного из комментариев с явным политическим акцентом стало распространение требования о создании «самостоятельной китайской системы знаний», первоначально сформулированного прежде всего применительно к философии и общественным наукам, на лексику глобально интегрированной сферы ИИ. Важно, что материал появился в рубрике оперативных комментариев портала

People.cn 人民锐评 (*Жэньминь жуйпин*)[4], публикующей редакционные оценки общественно значимых вопросов. Автор представляет широкое использование английской терминологии как угрозу «основе родного языка», технологической дискурсивной силе Китая и, в конечном счете, его способности к оригинальному мышлению [Li Wei, 2026]. Терминологическая зависимость тем самым в этом установочном материале связывается с более глубокими рисками когнитивной зависимости: если китайские специалисты и дальше будут заимствовать иностранные названия вместе с заложенными в них понятийными структурами, Китаю будет труднее сформировать самостоятельный технологический нарратив и разрабатывать решения, основанные на собственном опыте. Такой подход был заметен и в публикации приложения «Жэньминь жибао», где интернет-пользователям как раз предложили поучаствовать в поиске китайского названия AI, столь же краткого, образного и естественного, как 电脑 (*дяньнао*, «компьютер») или 手机 (*шоуцзи*, «мобильный телефон») [Du Wenwen, 2026].

Отсутствие короткого варианта, записанного иероглифами, в статье было представлено как свидетельство того, что китайский язык еще не вполне освоил это понятие. Ду Вэньвэнь даже сослался на «Закон об общегосударственном языке и письменности» (国家通用语言文字法 / *Гоцзя тунюн юйянь вэньцзы фа*) в редакции 2025 г. [Zhonghua Renmin Gongheguo Guojia …, 2025], характеризуя употребление аббревиатуры *AI* как недостаточно нормативное. Если быть юридически точным, такой вывод выходит за рамки реального содержания закона. Хотя правовой акт действительно закрепляет приоритет нормативного китайского языка в публичной и официальной сферах, он допускает иностранные языковые формы при наличии необходимого пояснения на общегосударственном языке. Таким образом, дискуссионная проблема употребления была превращена в вопрос языкового и культурного самоутверждения: должна ли технология, имеющая ключевое значение для будущего Китая, и далее фигурировать в виде английской аббревиатуры или китайский язык должен продемонстрировать способность назвать ее исключительно собственными средствами?

Предложенные названия хорошо показывают, почему реализовать эту амбицию оказалось затруднительно. «Жэньминь жибао» отмечала, что опрос привел к появлению свыше 100 тыс. содержательных предложений, однако в итоге упомянуто было не более нескольких десятков [Chen Jingwen, 2026]. Впрочем, акция, видимо, и не имела цели создать некий ранжированный список вариантов перевода, скорее задача состояла в том, чтобы превратить терминологическую дискуссию в общественное событие. Примечательно, что «народный конкурс» фактически воспроизвел уже знакомые по дискуссии о токене способы терминообразования вместе с присущими им ограничениями.

Например, термин 智脑 (*чжинао*, «умный мозг») был сразу понятен, но именно в этом заключалась и его слабость. Многообразные стороны ИИ сводились к «искусственному мозгу», тем самым поддерживалось ложное представление, будто технология мыслит так же, как человек. Вариант 智算 (*чжисуань*, «интеллектуальные вы-

[4] Название рубрики 锐评 (*жуйпин*) передает идею «острого», «проницательного», «заостренного» комментария: 锐 *(жуй)* – «острый, проницательный», 评 (*пин*) – «комментарий, оценка».

числения») позволял избежать части этой антропоморфизации, но звучал скорее как название технического процесса, чем технологии, с которой человек сталкивается в повседневной жизни. Кроме того, эта форма уже была связана с инфраструктурой интеллектуальных вычислений, например в 智算中心 (*чжисуань чжунсинь*, «центр интеллектуальных вычислений»), что делало ее не самым удачным кандидатом на обозначение всей сферы ИИ. Эти предложения выигрывали в понятности, но достигали ее ценой сужения обозначаемого объекта. У культурно более ярких вариантов возникала противоположная проблема. Термин 灵机 (*линцзи*) удачно обыгрывал двойную ассоциацию: иероглиф 机 (*цзи*, механизм, аппарат) прочно связан с технической лексикой, а само слово *линцзи* означает находчивость или внезапное озарение и знакомо по устойчивому выражению 灵机一动 (*линцзи идун*) – «в голову внезапно пришла удачная мысль». Такое обозначение звучало по-китайски и хорошо запоминалось, но создавало риск перенести на генеративные модели практику человеческого творчества и научного поиска, представить деятельность ИИ как серию вспышек озарения. Вариант 灵犀 (*линси*) обладал еще более глубоким контекстом, поскольку отсылал к строке танского поэта Ли Шанъиня (李商隐, ок. 813–858) 心有灵犀一点通 (*синь ю линси и дянь тун*) – образу сердец, связанных подобно тонкой белой прожилке, которая, согласно древнему поверью, проходила через чудесный рог носорога. Это слово было слишком прямой отсылкой к интуитивному взаимопониманию и представляло собой попытку образно отразить понятие человеко-машинного взаимодействия. Вариант был изящным, ярким и эмоционально привлекательным, но никоим образом не указывал на особенности искусственного интеллекта. Его техническое содержание неочевидно: без пояснения читатель скорее подумает о чуткости и душевном взаимопонимании, чем о работе системы ИИ под управлением человека и в синергии с ним.

芯智 (*синьчжи*) и 芯力 (*синьли*) представляли еще один вид языковой игры: 芯 (*синь*, «чип») придавало термину выраженное технологическое звучание, тогда как произношение вызывало в памяти омонимичные 心智 (*синьчжи,* «ум, интеллект») и 心力 (*синьли*, «умственные или духовные силы»). Однако оба варианта принимали физическую основу технологии за саму технологию. Система ИИ зависит от чипов, но определять ее через них столь же неверно, как определять компьютер просто как совокупность интегральных схем.

Некоторые предложения демонстрировали сложный набор намеков, нагромождение смыслов и многослойность ассоциаций, что не соответствовало требованиям к технической точности. 慧矩 (*хуэйцзюй*) соединяло «мудрость» с иероглифом, означающим «прав***и***ло» или «плотницкий угольник», возможно, одновременно намекая на математическую «матрицу» 矩阵 (*цзюйчжэнь*). Однако без пояснения обычному носителю языка было бы трудно догадаться о предполагаемом значении.

Особенно оригинальным предложением стал термин 神仝 (*шэньтун*). Автор этого чистого неологизма наделил оба иероглифа несколькими смысловыми слоями. Первый, 神 (*шэнь*), ассоциировался с 神州 (*Шэньчжоу*) – традиционным названием Китая, а также со словом 神智 (*шэньчжи*), означающим «разум» или «интеллект». Второй знак 仝, древняя форма иероглифа 同 (одинаковый, совместный), графически объединяет 人 (человек) и 工 (труд), что, по замыслу автора, позволило сжать значе-

ние 人工 («искусственный») в один иероглиф. В совокупности *шэньтун* также ассоциируется с 神童 (*шэньтун*, «вундеркинд») и 神通 (*шэньтун*, «сверхъестественные способности»). В целом термин складывается в своеобразный ребус о продолжении традиций китайской цивилизации, уверенности в собственной культуре и необычайных интеллектуальных возможностях ИИ, находящихся почти на божественном уровне. При всей оригинальности и культурной амбициозности семантическая ясность термина здесь невысока и без пространных комментариев часть заложенных в него смыслов остается непонятной для современных носителей языка. Пример четко показывает основной компромисс культурно мотивированной «китаизированной» номинации: чем богаче заложенная в термин система ассоциаций, тем выше зависимость его восприятия от толкований, при этом искусственное «накачивание» концептуальными смыслами и историко-культурными аллюзиями часто достигается ценой прозрачности и выразительности.

Заключение

Дискуссия о переводе токена и освещение вопросов терминологической системы ИИ в партийной прессе показали различия между логикой профессиональной терминологической работы и политизацией технологического языка. Лингвисты начинают с вопросов о том, что обозначает термин, где проходят границы понятия и можно ли последовательно использовать его рядом со смежными понятиями. Политико-культурная постановка исходит из других предпочтений: записано ли слово китайскими иероглифами, несет ли оно китайские культурные ассоциации и демонстрирует ли уверенность в способности языка самостоятельно осваивать иностранное знание. Эти два подхода способны эффективно дополнять друг друга, но могут и вступать в противоречие. Наш анализ показал, что наиболее культурно выразительные названия часто оказывались наименее технически корректными, тогда как более точные варианты звучали узко, бесцветно или уже принадлежали другой понятийной области.

Рассмотренный случай имеет и более широкий исторический контекст. Как отмечает А.В. Лукин, «иностранные понятия приходили в Китай, подвергаясь переосмыслению под воздействием культурно-лингвистической традиции и политических обстоятельств» [Лукин, 2015, с. 197]. На примере с токеном мы видим востребованность и устойчивость этого механизма. Заимствование нового термина сопровождается не только попытками найти максимально точный языковый эквивалент, но и его включением в существующую систему традиционных понятий и согласованием с политическими приоритетами.

Выделенные в настоящей статье терминологические модели разными способами решали эти задачи. Как считают участники дискуссий, успешная терминология ИИ в современном Китае должна отвечать нескольким, порой противоречащим друг другу критериям. Она должна точно соответствовать технологической реальности и принципам функционирования ИИ, не входить в противоречие с существующей терминологией в смежных областях, естественно восприниматься носителями китайского языка, иметь связь с традиционными культурными представлениями и быть пригод-

ной для включения в систему стандартизации. Ясно, что сочетать все требования в равной степени невозможно.

Вариант *цыюань* для передачи понятия токен не стал идеальным выбором, а скорее потенциально наиболее устойчивым решением – прагматичным компромиссом, в котором сочетание перечисленных выше критериев получило преимущество перед радикальным терминологическим новаторством. Хотя эксперты продолжают дискуссии, очевидно, что *цыюань* является не просто компромиссным переводом, а одним из лингвистических маркеров превращения Китая из страны, лишь адаптирующей ключевые понятия эпохи ИИ, в создателя собственной развитой терминологической системы. Данный кейс также наглядно иллюстрирует приоритет семантической мотивированности над фонетической ассимиляцией при образовании новых терминов.

По мере того как Китай продолжает совершенствовать собственные большие языковые модели, терминология, несомненно, будет и дальше развиваться. Дальнейшее изучение динамики терминологических гнезд, связанных с генеративным ИИ, представляется перспективным направлением исследований, способным дать понимание стратегий адаптации языков изолирующего типа к вызовам современной технологической революции.

Библиографический список

Арекеева Ю.Е., Бедненко Ю.И. Репрезентация концепта «人工智能 / искусственный интеллект» в языковом сознании носителей китайского языка // Вестник Томского государственного педагогического университета (TSPU Bulletin). 2025. Вып. 1 (237). С. 7–15.

Горелов В.И. Лексикология китайского языка: учебное пособие для студентов. М.: Просвещение, 1984. 206 с.

Ивченко Т.В. Китайский язык: вопросы стандартизации и терминологичности: по материалам Международной конференции «Китайская лингвистика и синология», РГГУ, 3–5 октября 2019 г. // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2020. № 7. С. 11–15.

Лукин А.В. Приход понятия «демократия» в Китай и его первые интерпретации // Лукин А.В. Возвышающийся Китай и будущее России: работы о Китае и российско-китайских отношениях: сборник статей. М.: Международные отношения, 2015. С. 165–197.

Очиров О.Р. Становление китайского терминоведения: традиции и современность // Вестник Российского университета дружбы народов. Серия: Лингвистика. 2013. № 4. С. 116–125.

Фролова О.П. Словообразование в терминологической лексике современного китайского языка: монография. М.: Восточная книга, 2011. 168 с.

Хаматова А.А. Словообразование современного китайского языка. М.: Муравей, 2003. 224 с.

References

Arekeeva Yu.E., Bednenko Yu.I. (2025) Reprezentatsiya kontsepta 人工智能 / iskusstvennyy intellekt v yazykovom soznanii nositeley kitayskogo yazyka [Representation of the concept 人工智能 / artificial intelligence in the linguistic consciousness of Chinese native speakers]. Vestnik Tomskogo gosudarstvennogo pedagogicheskogo universiteta [Tomsk State Pedagogical University Bulletin]. No. 1 (237). P. 7–15 (In Russian).

Frolova O.P. (2011) Slovoobrazovanie v terminologicheskoy leksike sovremennogo kitayskogo yazyka: monografiya [Word formation in the terminological lexicon of modern Chinese: A Monograph]. Moscow: Vostochnaya kniga. (In Russian).

Gorelov V.I. (1984). Leksikologiya kitayskogo yazyka: uchebnoye posobiye dlya studentov [Lexicology of the Chinese Language: A Textbook for Students]. Moscow: Prosveshchenie. (In Russian).

Ivchenko T.V. (2020). Kitayskiy yazyk: voprosy standartizatsii i terminologichnosti: po materialam Mezhdunarodnoy konferentsii “Kitayskaya lingvistika i sinologiya”, RGGU, 3–5 oktyabrya 2019 g. [Chinese language: issues of standardization and terminology. Based on the materials of the International Conference “Chinese Linguistics and Sinology”, Russian State University for the Humanities, October 3–5, 2019]. Vestnik RGGU. Seriya “Literaturovedenie. Yazykoznanie. Kulturologiya” [Bulletin of the Russian State University for the Humanities. Series “Literary Studies. Linguistics. Cultural Studies”]. No. 7. P. 11–15. (In Russian).

Khamatova A.A. (2003). Slovoobrazovanie sovremennogo kitayskogo yazyka [Word Formation in Modern Chinese]. Moscow: Muravey. (In Russian).

Lukin A.V. (2015). Prikhod ponyatiya “demokratiya” v Kitai i ego pervye interpretatsii [The arrival of the concept of “democracy” in China and its first interpretations]. In A. V. Lukin, Vozvyshayushchiisya Kitai i budushchee Rossii: Raboty o Kitae i rossiisko-kitaiskikh otnosheniyakh [Rising China and the future of Russia: Works on China and Russian-Chinese relations] P. 165–197. Moscow: Mezhdunarodnye otnosheniya (In Russian).

Ochirov O.R. (2013). Stanovlenie kitayskogo terminovedeniya: traditsii i sovremennost [The formation of the Chinese terminology: tradition and modernity]. Vestnik Rossiyskogo universiteta druzhby narodov. Seriya: Lingvistika [Bulletin of the Peoples’ Friendship University of Russia. Series: Linguistics]. No. 4. P. 116–125. (In Russian).

* * *

Chen Guohua 陈国华. (2026). Type he token: fulei he fuyuan – cong ciyuanxue, fuhaoxue yu duomotai rengong zhineng shijiao geichu de liyou Type 和 token:符类和符元 – 从词源学、符号学与多模态人工智能视角给出的理由 [Type and Token: Fulei and Fuyuan – The Rationale from the perspectives of etymology, semiotics, and multimodal artificial intelligence]. Zhongguo keji shuyu 中国科技术语 [China Terminology], 28(3). P. 9–10. (In Chinese).

Chen Jingwen 陈静文. (2026). Gei AI zheng ming, yuyanxuejia zenme kan? (Xinwen li de xuewen) – Wailai keji shuyu hanhua de jingyan yu guifan 给 AI 征名，语言学家怎么看？（新闻里的学问）– 外来科技术语汉化的经验与规范 [Naming AI: What do linguists think? (The scholarship behind the news) – Experience and norms in the sinicization of foreign scientific and technological terms]. *Renmin ribao haiwai ban* 人民日报海外版 [People’s Daily Overseas Edition]. April, 10. P. 11 (In Chinese).

Du Wenwen 都文文. (2026). Qing gei AI qi ge hao ming 请给 AI 起个好名 [Please give AI a good name]. Renmin wang 人民网[People’s Daily Online]. March, 18. URL: http://finance.people.com.cn/n1/2026/0318/c1004-40684445.html (accessed: 21.08.2026). (In Chinese).

Fan Wenjie 范文杰. (2026). Keji shuyu yiming guifan xu yu cishu, jiaocai bianxiu xietong gongjin 科技术语译名规范需与辞书、教材编修协同共进 [The standardization of translated scientific and technical terms should advance in coordination with the compilation and revision of dictionaries and textbooks]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 28(3). P. 7. (In Chinese).

Feng Zhiwei 冯志伟. (2026). Token yingdang dingming wei “ciyuan” Token 应当定名为“词元” [Token should be named “ciyuan”]. Zhongguo keji shuyu 中国科技术语 [China Terminology], 28(3). P. 4–5. (In Chinese).

Gao Yali 高雅丽. (2026). Zhuanjia jiedu token zhongwen ming weihe ding wei “ciyuan”专家解读 token 中文名为何定为“词元” [Experts explain why the Chinese name for “token” was set as “ciyuan”]. *Zhongguo kexuebao* 中国科学报 [China Science Daily]. March, 30. P. 3 (In Chinese).

Ke Ping 柯平, He Kaida 贺凯达. (2026). Token ying yi wei “fuyuan” Token 应译为“符元” [“Token” should be translated as “fuyuan”]. Zhongguo keji shuyu 中国科技术语 [China Terminology], 28(3). P. 12–13. (In Chinese).

Lackner M., Amelung I., Kurtz J. (eds.). (2001). New Terms for New Ideas: Western Knowledge and Lexical Change in Late Imperial China. Leiden: Brill.

Li Wei 李玮. (2026). Renmin ruiping: Yong “token” haishi “ciyuan”, shiguan keji huayuquan 人民锐评：用“Token”还是“词元”，事关科技话语权 [People’s Daily Incisive Commentary: Whether to use “Token” or “ciyuan” concerns scientific and technological discursive power]. Renmin wang 人民网 [Peo-

ple's Daily Online]. August, 5. URL: http://opinion.people.com.cn/n1/2026/0805/c436867-40774551.html (accessed: 21.08.2026). (In Chinese).

Li Yuming 李宇明. (2026). Token Zhongyi, youxuan "ciyuan" Token 中译，优选"词元" [For the Chinese translation of "token," "ciyuan" is the preferred choice]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 28(3). P. 8–9. (In Chinese).

Ling Zhenhua 凌震华. (2026). "Ciyuan" yiming de helixing yu yingyong kaoliang "词元"译名的合理性与应用考量 [The Rationale for the translation "ciyuan" and considerations for its application]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 28(3). P. 13–14. (In Chinese).

Liu Jinting 刘金婷. (2020). Token zai butong lingyu nei de zhongwen yiming qianxi Token 在不同领域内的中文译名浅析 [Analysis of Chinese translation of token in different fields]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 22(5). P. 64–67. (In Chinese).

Quanguo kexue jishu mingci shending weiyuanhui guanyu fabu shiyong rengong zhineng lingyu mingci token zhongwen ming "ciyuan" de gonggao (2026) 全国科学技术名词审定委员会关于发布试用人工智能领域名词 token 中文名"词元"的公告 [Notice of the China National Committee for Terms in Sciences and Technologies on the trial use of "ciyuan" as the Chinese name for the AI term "token"]. Quanguo kexue jishu mingci shending weiyuanhui 全国科学技术名词审定委员会[China National Committee for Terms in Sciences and Technologies]. March, 25. URL: https://www.cnterm.cn/news/tzgg/202603/t20260325_827999.html (accessed: 21.08.2026). (In Chinese).

Rengong zhineng yuliaoku jichu shuyu (2026). 人工智能 语料库 基础术语 [Artificial Intelligence – Corpus – Basic Terminology]. Zhonghua renmin gongheguo jiaoyubu 中华人民共和国教育部 [Ministry of Education of the People's Republic of China]. March, 2. URL: http://www.moe.gov.cn/jyb_sjzl/ziliao/A19/202604/W020260407554480905449.pdf (accessed: 21.08.2026). (In Chinese).

Wan Quan 完权. (2026). Yi chang shangzhan gei token yiming de qishi 一场商战给 token 译名的启示 [Lessons for translating "token" from a commercial rivalry]. Zhongguo keji shuyu 中国科技术语 [China Terminology], 28(3). P. 10–11. (In Chinese).

Wenzi shilu. Guoxinban juxing "Gao zhiliang wancheng 'Shisiwu' guihua" xilie zhuti xinwen fabuhui: jieshao "Shisiwu" shiqi shuzi Zhongguo jianshe fazhan chengjiu (2025) 文字实录 | 国新办举行"高质量完成'十四五'规划"系列主题新闻发布会 介绍"十四五"时期数字中国建设发展成就 [Transcript. State Council Information Office holds a thematic press conference on "High-quality completion of the 14th Five-Year Plan" to present achievements in the development of Digital China during the 14th Five-Year Plan period]. Guojia Shuju Ju 国家数据局 [National Data Administration]. August, 14. URL: https://www.nda.gov.cn/sjj/swdt/xwfb/0814/20250814215705489334712_mobile.html (accessed: 21.08.2026). (In Chinese).

Wetzel L. (2009). Types and tokens: On abstract objects. Cambridge, Massachusetts: The MIT Press.

Xinxi jishu cihui. (2001) 信息技术 词汇 第 28 部分：人工智能 基本概念与专家系统 GB/T 5271.28-2001 [Information technology—Vocabulary—Part 28: Artificial intelligence—Basic concepts and expert systems GB/T 5271.28-2001]. Guojia zhiliang jishu jiandu ju 国家质量技术监督局 [State Administration of Quality and Technical Supervision]. July, 16. Effective from March 1, 2002 (accessed: 25.09.2026). (In Chinese).

Xinxi jishu rengong zhineng shuyu. (2022) 信息技术 人工智能 术语 GB/T 41867-2022 [Information technology—Artificial intelligence—Terminology GB/T 41867-2022]. Quanguo biaozhun xinxi gonggong fuwu pingtai 全国标准信息公共服务平台 [National public service platform for standards information]. October, 14. URL: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=195F522C14AD9A1A0094FF66D0B1EF1B (accessed: 21.08.2026). (In Chinese).

Xu Fei 徐飞. (2026). Cong token zhongwen ming "ciyuan" shuo kai qu 从 token 中文名"词元"说开去 [Reflections on the Chinese Name "ciyuan" for "token"]. Guojia Shuju Ju Wang 国家数据局网 [National Data Administration Website]. April, 14. URL: https://www.nda.gov.cn/sjj/zwgk/zjjd/0414/20260414164105815459230_pc.html (accessed: 21.08.2026). (In Chinese).

Yuan Jiahong 袁家宏. (2026). Dui token yi zuo "ciyuan" de yidian shangque 对 token 译作"词元"的一点商榷 [Some reservations about translating "token" as "ciyuan"]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 28(3). P. 11–12. (In Chinese).

Zhang Junping 张军平. (2026). Token ying yi wei “zhiyuan” Token 应译为“知元” [“Token” should be translated as “zhiyuan”]. Zhongguo keji shuyu 中国科技术语 [China Terminology], 28(3). P. 6. (In Chinese).

Zhang Qingrui 張慶瑞. (2024). “Wanwu jie daoyuan”, Laozi zao zhidao?「萬物皆道元」，老子早知道？ [“All things are Daoyuan”: Did Laozi know it all along?]. Taida xiaoyou shuangyuekan 臺大校友雙月刊 [NTU Alumni Bimonthly]. No.155. P. 40–45. (In Chinese).

Zhang Zhonghong 张中鸿, Zhang Qingrui 张庆瑞. (2025). Zai Laozi zhexue yu “Daoyuan”: rengong zhineng token de yuyi yu yingyong tansuo Token 在老子哲学与“道元”:人工智能 token 的语义与应用探索 [The philosophy of Laozi and Daoyuan: A study on the semantics and applications of AI token]. Zhongguo keji shuyu 中国科技术语 [China Terminology]. No. 27(1). P. 57–63. (In Chinese).

Zhao Zhuqing 赵竹青. (2026). Qinghua daxue Yang Bin: Jianyi jiang “token” yi wei “moyuan”, tuidong AI quanmin puji 建议将“token”译为“模元”，推动 AI 全民普及 [Tsinghua University’s Yang Bin: Proposal to translate “token” as “moyuan” to promote the widespread adoption of AI]. Renmin wang 人民网 [People’s Daily Online]. March, 18. URL: http://finance.people.com.cn/n1/2026/0318/c1004-40684447.html (accessed: 21.08.2026). (In Chinese).

Zhonghua Renmin Gongheguo Guojia Tongyong Yuyan Wenzi Fa (2025) 中华人民共和国国家通用语言文字法 [Law of the People’s Republic of China on the Standard Spoken and Written Chinese Language]. Zhonghua Renmin Gongheguo Jiaoyubu 中华人民共和国教育部 [Ministry of Education of the People’s Republic of China]. December, 27. URL: https://www.moe.gov.cn/jyb_sjzl/sjzl_zcfg/zcfg_jyfl/202606/t20260605_1438880.html (accessed: 21.08.2026). (In Chinese)

Поступила в редакцию: 10.09.2026. Received: 10 September 2026.
Принята к публикации: 15.09.2026. Accepted: 15 September 2026.