

Что такое жизнь? Мета-анализ различных мнений об определении жизни.

Бендер, Рид, Карина Кофман, Блез Агуэра-и-Аркас и Майкл Левин // *Biological Theory* (2026): 1-22. <https://doi.org/10.1007/s13752-026-00544-9>

Аннотация

Вопрос «что такое жизнь?» веками ставил перед учеными и философами сложные задачи, порождая множество определений, отражающих как тайну ее возникновения, так и разнообразие дисциплинарных подходов к этому вопросу. Несмотря на значительный прогресс в нашем понимании биологических систем, психологии, вычислительной техники и теории информации, ни одно определение жизни до сих пор не получило всеобщего признания. Эта задача становится все более актуальной по мере того, как достижения в области синтетической биологии, искусственного интеллекта и астробиологии бросают вызов нашим традиционным представлениям о том, что значит быть живым. Мы применили методологический подход, использующий большие языковые модели (LLM), для анализа набора определений жизни, предоставленных тщательно отобранной группой междисциплинарных экспертов. Мы использовали новый метод попарного корреляционного анализа для преобразования определений в отдельные векторы признаков, за которым последовали агломеративная кластеризация, внутрикластерный семантический анализ и проекция t-SNE (t-распределенное стохастическое встраивание соседей) для выявления лежащих в основе концептуальных архетипов. Данная методология выявила непрерывную картину тем, связанных с определением жизни, предполагая, что то, что исторически рассматривалось как бинарная таксономическая проблема, следует вместо этого понимать как дифференцированные перспективы в рамках единого концептуального латентного пространства. Мы предлагаем новый методологический мост между редукционистским и целостным подходами к фундаментальным вопросам науки и философии, демонстрируя, как вычислительный семантический анализ может выявлять концептуальные закономерности на стыке дисциплин и открывая аналогичные пути для решения других спорных вопросов определения жизни в различных науках.

Введение

Задача определения жизни выходит за рамки семантики; она выявляет фундаментальные эпистемологические различия в том, как специалисты разных дисциплин концептуализируют и исследуют границы между живыми и неживыми системами. Биологи часто делают акцент на метаболических процессах, размножении и клеточной организации (Gilbert [1982](#); [1991](#); Margulis and Sagan [2000](#); Boden [2003](#); Gánti [2003b](#); Sousa et al. [2008](#)), в то время как физики могут сосредоточиться на термодинамических свойствах, использовании энергии и уменьшении энтропии (Schrödinger [1944](#); Prigogine [1980](#); England [2020](#); Nicholson [2025](#)). Специалисты по информатике и теоретики сложности могут, в свою очередь, отдавать приоритет обработке информации, самоорганизации и возникающей сложности (Holland [1975](#); Langton [1997](#); Wolfram [2002](#); Seoane and Sole [2018](#); Banerjee [2021](#)).

Эти расходящиеся точки зрения, сформированные различными методологическими подходами и теоретическими рамками, исторически было трудно согласовать. Эта проблема определения становится все более актуальной по мере того, как достижения в области синтетической биологии (Blackiston et al. [2021](#)), искусственного интеллекта (Faggin [2024](#)) и астробиологии (Benner [2010](#)) продолжают ставить под сомнение само понятие жизни, что имеет не только научные, но и этические и нормативные последствия.

Некоторые выступают за плюрализм определений — принятие множества контекстуально подходящих определений вместо стремления к единой универсальной концепции (Gayon et al. [2010](#); Mariscal [2023](#)). Этот прагматический подход признает ценность специализированных определений в дифференцированных исследовательских контекстах. Однако он оставляет без ответа вопрос о том, могут ли существовать более глубокие закономерности согласованности за этими кажущимися дисциплинарными разделениями и возможно ли более интегрированное понимание без ущерба для дисциплинарных перспектив.

Другие полностью отвергают бинарные классификации, критикуя попытки провести произвольные границы между живым и неживым. Эти взгляды переосмысливают дискуссию, рассматривая жизнь как существующую в рамках множества непрерывных измерений, а не как категориальное различие (Pora [2004](#); Malaterre and Chartier [2021](#); Parke [2021](#); Levin [2022](#)). Более радикально некоторые ученые ставят под сомнение целесообразность определения жизни вообще, предполагая, что такие попытки либо концептуально бесполезны, либо совершенно ошибочны (Cleland and Chyba [2010](#); Machery [2012](#); Mariscal and Doolittle [2020](#)). В дополнение к многочисленным полнотекстовым статьям, опубликованным исследователями, специализирующимися на этой теме, мы стремились изучить мнения других ведущих ученых в смежных областях.

Мы опросили группу тщательно отобранных современных ученых, чтобы получить их очень краткие определения жизни, а затем использовали как ручной, так и вычислительный семантический анализ для изучения полученного набора данных. Наш выбор ученых для этого опроса не был призван создать статистически репрезентативный консенсус или всеобъемлющий обзор области, а скорее показать пример и исследование того, как современные методы машинного обучения могут быть использованы для анализа семантических данных по сложной проблеме, для создания программного обеспечения и протоколов, которые впоследствии можно было бы расширить на гораздо более масштабные исследования с участием представителей различных групп и сообществ. Здесь мы специально выбрали группу современных ученых, чьи взгляды на вопрос «жизни» ранее не публиковались, сосредоточившись на выборке тех, чьи междисциплинарные работы мы сочли наиболее интересными в своей зависимости (или отсутствии зависимости) от того, как определяется жизнь.

В нашем исследовании мы рассмотрели неоднозначность определений с помощью методологического подхода, использующего большие языковые модели (БЛМ) для систематического анализа 68 ответов экспертов на вопрос «что такое жизнь?». Мы применили парный корреляционный анализ, полученный с помощью БЛМ, для количественного отображения совокупного набора ответов в выведенную корреляционную матрицу, за которым последовали агломеративная кластеризация и снижение размерности для проецирования закодированных и кластеризованных векторов признаков респондентов на двумерную плоскость. Для дальнейшей интерпретации этого

визуализированного семантического пространства мы использовали БЛМ для сравнения внутрикластерных определений жизни и для генерации консенсусного определения жизни, репрезентативного для каждого кластера.

Систематически отображая, как различаются точки зрения отдельных экспертов в зависимости от перцептивных рамок, мы стремимся создать многомерную основу для понимания «живости», которая выходит за рамки традиционных категориальных различий и помогает принимать этические и концептуальные решения в нашем быстро развивающемся технологическом ландшафте.

Фон

Исторические представления о жизни

Задолго до появления современных научных концепций древние цивилизации стремились охарактеризовать движущую силу, порождающую жизнь. Вместо натуралистических объяснений эти древние культуры опирались на сверхъестественные абстракции, такие как Прана, Ци, Ка, Пневма и Руах, в ведической, китайской, египетской, греческой и еврейской традициях соответственно. Эти концепции, несмотря на различное культурное происхождение, объединяла общая попытка сформулировать невидимый принцип, отделяющий одушевленную и неодушевленную материю в рамках метафорических абстракций.

Греческая философия преобразовала эти символические концепции в первые систематические определения души и её отношения к жизни. Космология Платона (428–348 до н.э.) в «*Тимее*» характеризует сам космос как живое и разумное существо, движения которого управляются познанием, а не механической причинностью, тем самым впервые объединяя душу как принцип познания и принцип жизни (Кэмпбелл [2021](#)). Для Платона дифференцированная жизнь микрокосма была зеркальным отражением выражения жизни в масштабе макрокосма.

...непреодолимая ценность «Тимея» заключается в успешном представлении космологической основы теоретической и практической этики для человечества. Его тема — жизнь, порождающий принцип жизни — не просто жизнь одного человека или даже всего человечества, но генезис τοῦ παντός,^{[Сноска1](#)}. Подходящая тема для любого философа. (Уитакер, [1941](#), с. 104)

Аристотель (384–322 до н.э.) затем развил эту концепцию посредством своего телеологического подхода в *трактате «О душе»*, дополнительно характеризуя жизнь по ее проявленным функциональным свойствам. Его биологическая концепция представляет собой вложенную иерархию способностей, в которой растительная душа (питание и размножение) присутствует во всех живых существах, чувственная душа (восприятие и движение) — в животных, а разумная душа уникально присутствует в человеке (Джонсон [2005](#)). Его акцент на внутренней цели (телос) заложил основу для первого онтологически обоснованного определения жизни как самоактуализирующейся системы, где наиболее естественным актом живых существ является «производство другого подобного себе,

животного, производящего животное, растения, чтобы, насколько это позволяет природа, оно могло приобщиться к вечному и божественному» (Аристотель [1931](#), «О душе», II.4, 415a). Для Аристотеля цель, к которой стремятся все живые существа, состоит в том, чтобы «делать все, что позволяет их природа» (Аристотель [1931](#), «О душе», II.4, 415b), что устанавливает живые существа как принципиально самоактуализирующиеся системы с внутренней телеологией.

Переход от греческой мысли к нашим современным материалистическим представлениям о жизни был осуществлен средневековыми и ренессансными мыслителями, которые сохранили телеологические рамки, одновременно вводя все более материальные объяснения. Гален (129–216 гг. н.э.) расширил аристотелевскую биологическую телеологию посредством эмпирической анатомии, продемонстрировав целенаправленное конструирование каждой анатомической особенности (Гален, [1968](#)). Алхимики эпохи Возрождения, в частности Парацельс (1493–1541), впоследствии разработали гибридные концепции, рассматривавшие организмы как химические лаборатории, управляемые «археем» или внутренним алхимиком — концепция, которая сохраняла ощущение телеологического функционализма, все больше описывая жизнь через ее биохимическую форму (Чанг, [2011](#)). Эти «химические философии» представляли собой важную концептуальную эволюцию (Паршалл и др., [2015](#)), поддерживая представление о том, что живые системы функционируют в соответствии с внутренними целями, одновременно начиная формулировать материальные процессы, лежащие в основе жизненно важных функций. Эта историческая последовательность событий создала фундаментальное противоречие между телеологической и механистической перспективами, которое продолжает бросать вызов современным попыткам определить жизнь.

В эпоху Просвещения Рене Декарт (1596–1650) рассматривал человека как совокупность частей, которые вместе функционируют как машина, дуалистически отличная от разума, управляющего его функциями. Этот механистический подход к проявлению жизни коренным образом переориентировал ее концепцию, отдалив ее от телеологических корней и перенаправив внимание на механистические процессы.

Таким образом, я утверждаю, что если вы задумаетесь над тем, как эти функции в этой машине вытекают совершенно естественно исключительно из расположения органов, не больше и не меньше, чем функции часов или другого автомата из его противовесов и колес, то вам не нужно представлять себе по этой причине какую-либо другую вегетативную душу, ни чувственную, ни какой-либо другой принцип движения и жизни, кроме ее крови и животных инстинктов, возбужденных жаром постоянно горящего огня в сердце и имеющих ту же природу, что и огонь, присутствующий в неодушевленных телах. (Декарт [[1664](#)] [1909](#), с. 202; перевод авторов)

В то время как Декарт описывал жизнь как двойственность нематериального разума и его механистического тела, его современники развили из его работ всесторонне материалистические философии. Томас Гоббс (1588–1679) предложил, пожалуй, наиболее всеобъемлющую материалистическую концепцию этого периода, полностью отвергнув картезианский дуализм и утверждая, что все явления, характерные для жизни, могут быть сведены к материи в движении. В концептуализации жизни Гоббс отверг предшествующие телеологические или метафизические концепции в пользу точки зрения, согласно которой биологические функции являются чисто механическими процессами.

Ведь жизнь — это всего лишь движение конечностей, начало которого находится в какой-то главной части внутри; почему бы нам не сказать, что все автоматы (двигатели, которые движутся сами собой с помощью пружин и колес, как часы) обладают искусственной жизнью? Ведь что такое сердце, как не пружина; и нервы, как не множество струн; и суставы, как не множество колес, приводящих в движение все тело, как и задумал мастер? (Гоббс, [1909](#), с. 8)

В эпоху Просвещения эта механистическая натурализация жизни продолжала развиваться благодаря таким философам, как Гассенди (Стоунс, [1928](#)), Спиноза (Спиноза, [2020](#)) и Ла Меттри (Ла Меттри, [1865](#)), достигнув кульминации в совершенно материалистической концепции, которая определила нашу современную научную эпоху. Однако даже когда философский ландшафт решительно сместился в сторону механистических объяснений, фундаментальный вопрос о том, как определить жизнь, оставался. К началу XIX века научное исследование не решило, а скорее переформулировало эту древнюю проблему: если жизнь возникает из чисто физических процессов, что отличает живую материю от неживой?

Современные представления о жизни

Переход к современным научным представлениям о жизни был катализирован рядом ключевых событий XIX века: клеточная теория определила фундаментальную организационную единицу жизни (Маттиас, [1838](#)); теория эволюции Чарльза Дарвина путем естественного отбора предоставила первую механистическую основу для объяснения сложности жизни и ее адаптации (Дарвин, [1859](#)); появление термодинамики заложило физическую основу для понимания того, как живые системы поддерживают порядок, несмотря на второй закон термодинамики (Больцман, [1886](#)).

Эти достижения подготовили почву для знаковых лекций Эрвина Шрёдингера 1944 года «Что такое жизнь?», в которых организмы были представлены как термодинамические системы, создающие и поддерживающие порядок путем генерации энтропии в окружающей среде (Шрёдингер, [1944](#)). Эта термодинамическая перспектива, будучи влиятельной, сразу же столкнулась с проблемами: могла ли чисто энергетическая характеристика передать качественное различие между живым и неживым?

Молекулярная революция середины XX века обещала разрешение проблем. Расшифровка Джеймсом Уотсоном и Фрэнсисом Криком структуры двойной спирали ДНК открыла физическую основу наследственности (Уотсон и Крик, [1953](#)), а последующие открытия в молекулярной биологии выявили центральную догму потока генетической информации (Крик, [1958](#)). Это механистическое понимание наследственности породило информационно-теоретические определения, наиболее известным из которых является определение Джона фон Неймана, который охарактеризовал жизнь как вычислительную систему, способную к самовоспроизведению с наследуемой изменчивостью (фон Нейман, [1966](#)).

Современные подходы к определению жизни получили широкое распространение в различных дисциплинах, расширяя спектр доступных перспектив для её определения. Термодинамические модели делают акцент на потоке энергии и производстве энтропии (Macklem and Seely [2010](#)), в то время как автопоэтические теории, такие как разработанные Матураной и Варелой, фокусируются на самоорганизации и поддержании

границ (Maturana and Varela [1980](#); Moreno and Mossio [2015](#); Moreno and Peretó [2026](#)). Модель хемотона Ганти предлагает минимальные требования, включая метаболизм, хранение информации и контроль границ (Gánti [2003a](#)). Операциональное определение НАСА, заимствованное из предложения Карла Сагана, представляет собой прагматическое определение жизни для астробиологических исследований: «самоподдерживающаяся химическая система, способная к дарвиновской эволюции» (Benner [2010](#), p. 1).

Несмотря на усложнение ландшафта определений жизни, напряжение между точками зрения сохраняется. Термодинамические подходы могут описывать динамику энергии, недооценивая при этом организационную сложность. Информационно-теоретические модели освещают генетическое кодирование, но испытывают трудности с семантикой биологического смысла. Операциональные определения сталкиваются с пограничными случаями: вирусы, прионы и вычислительные формы жизни, которые проявляют некоторые, но не все традиционные характеристики жизни.

Появление искусственной жизни, синтетической биологии и сложных вычислительных систем еще больше усложнило границы определений. Синтетическая клетка *микоплазмы*, созданная Крейгом Вентером, продемонстрировала возможность конструирования жизнеспособной синтетической клетки со сжатым геномным кодом (Hutchison et al. [2016](#)); лаборатория Майкла Левина показала, что эмбриональные клетки лягушки могут самоорганизовываться в новые самодвижущиеся формы живых организмов с уникальным поведением и транскриптомами, если их вывести из традиционного физиологического контекста (Blackiston et al. [2021](#); Pai et al. [2025](#)), как и взрослые клетки человека (Gumuskaya et al. [2023](#), [2025](#)); искусственные нейронные сети (Idrees et al. [2024](#)) и эволюционные алгоритмы (Wang et al. [2025](#)) демонстрируют адаптацию и размножение без биологических субстратов. Эти открытия ставят под сомнение предположение о том, что для проявления жизненных характеристик жизни необходимы химические процессы на основе углерода, клеточная организация или эволюционная адаптация.

Системная биология предпринимала попытки синтеза посредством иерархических моделей, интегрирующих множество организационных уровней, от молекулярных сетей до экосистем (Noble et al. [2019](#)). Однако даже этот интегративный подход должен учитывать фундаментальный вопрос: какие свойства представляют собой необходимые, а какие достаточные условия для жизни?

Распространение различных концептуальных определений отражает как научный прогресс, так и сохраняющиеся концептуальные проблемы в нашем понимании жизни. По мере углубления эмпирической науки границы определения жизни парадоксальным образом становятся менее четкими. Мы рассматриваем эти онтологические противоречия как семантическую топологию, которую отдельные дисциплинарные подходы отображали лишь частично, ожидая, пока инструменты вычислительного систематического анализа выявят лежащие в основе преэмственности, присутствующие в этом скрытом пространстве.

Чтобы начать исследовать это пространство и разработать методы анализа концептуального ландшафта мнений ученых в сложных междисциплинарных областях, мы провели анализ с помощью ИИ ряда определений «жизни», данных современными

учеными. Тщательно отобранным мыслителям из различных дисциплин было предложено дать определение «жизни» (или возразить против возможности дать такое определение) в трех предложениях или менее. Наша цель состояла в том, чтобы составить карту структуры и основных факторов, определяющих крайне разнообразный набор ответов, и оценить полезность линейных моделей мышления для этой задачи.

Априорные вычислительные методы кластеризации текста

Наш подход основан на последних достижениях в области кластеризации с использованием ИИ, формирования консенсуса и моделирования попарных ограничений, и применяет эти методы для анализа концептуальных закономерностей в семантическом пространстве определений жизни, данных нашими респондентами.

Ранее было показано, что модели с латентными пространствами (LLM) демонстрируют сопоставимую или превосходящую производительность по сравнению с современными методами встраивания текста и кластеризации, заставляя модель генерировать потенциальные метки для заданного набора данных, а затем классифицировать каждый образец соответствующей меткой (Huang and He [2024](#)). Чаще всего исследователи используют LLM для генерации текстовых встраиваний, которые кодируют контекстное значение в плотные векторы с плавающей запятой, где абстрактные лингвистические паттерны могут быть представлены геометрически. Сравнительный анализ встраиваний LLM для кластеризации демонстрирует, как эти модели последовательно улавливают семантические отношения, закодированные как геометрическая близость в многомерных латентных пространствах (Keraghel et al. [2024](#)).

Недавние инновации еще больше расширили эти возможности благодаря интерпретируемой кластеризации методом k-средних и механизмам обратной связи, настроенным на основе инструкций. Алгоритм k-LLMmeans использует LLM для генерации текстовых резюме в качестве центроидов кластеров, улавливая семантические нюансы, часто теряемые при использовании чисто математических свойств кластеризации методом k-средних (Diaz-Rodriguez [2025](#)). Структура ClusterLLM вводит LLM для улучшения кластеризации встроенного текста путем построения вопросов-триплетов < соответствует ли *A* лучше *B*, чем *C* >, где *A*, *B* и *C* — похожие точки данных, принадлежащие разным кластерам в соответствии с исходным вектором встраивания (Zhang et al. [2023](#)). Следуя той же парадигме, Liusie et al. Исследования показали, что LLM-ы значительно лучше справляются с задачами сравнительной оценки (например, «Какое резюме более связное, *A* или *B*?»), чем с задачами абсолютной оценки (например, «Присвойте оценку от 1 до 10, измеряющую связность этого резюме») (Liusie et al. [2023](#)). Эти достижения демонстрируют, как LLM-ы могут использовать как встроенные семантические связи, так и прямые лингвистические инструкции для достижения превосходных результатов кластеризации, особенно при использовании попарного сравнительного анализа. Эта возможность интерпретируемого анализа посредством попарного сравнения делает LLM-ы особенно подходящими для отображения сложного концептуального ландшафта определений жизни.

Возможности LLM напрямую решают методологические проблемы, присущие анализу разнообразных определений жизни. Традиционные подходы испытывали трудности с систематическим сравнением и интеграцией определений в разных дисциплинах. Наш подход рассматривает определения экспертов как точки данных в количественно

измеримом семантическом пространстве, что позволяет проводить вычислительный анализ закономерностей, которые не поддаются традиционным категориальным рамкам. Применяя эти методы к 68 ответам экспертов, мы можем отобразить концептуальные связи между различными подходами к определению, сохраняя при этом нюансы и сложность каждой перспективы. Эта методология превращает вопрос «что такое жизнь?» из философской дискуссии в структурированное исследование семантических закономерностей в экспертном дискурсе.

Методы

Мы разработали программную архитектуру для анализа и кластеризации различных определений жизни с использованием больших языковых моделей (LLM) и методов обучения без учителя. Наш подход включает пять ключевых методологических компонентов:

1. Экспертная подборка и коллекция определений
2. Количественный попарный корреляционный анализ между определениями
3. Агломеративная кластеризация полученной корреляционной матрицы
4. Тематический анализ внутри- и межкластерных семантических закономерностей
5. t-SNE-редукция размерности векторов корреляционных признаков в двумерное пространство

На рисунке [1](#) показаны основные вычислительные этапы, необходимые для преобразования исходной матрицы попарных корреляций в отсортированную и кластеризованную матрицу, а затем и в конечном итоге в двумерную t-SNE проекцию пространства определений.

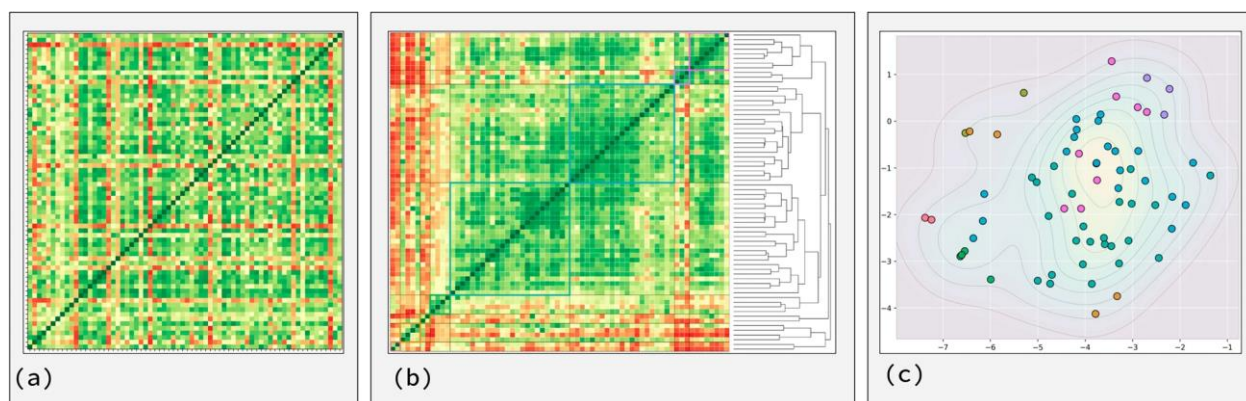


Рис. 1. Процесс преобразования данных из попарных корреляций в кластерную пространственную проекцию. На панели (а) показана исходная матрица попарных корреляций для всех 68 определений, где зеленый цвет указывает на высокое согласие (корреляция, приближающаяся к 1,0), а красный — на несогласие (корреляция, приближающаяся к -1,0). На панели (б) показана отсортированная матрица корреляций,

полученная после применения иерархической кластеризации, а справа отображена результирующая дендрограмма, показывающая вложенные связи между определениями. На панели (в) представлена окончательная проекция t-SNE, которая преобразует эти многомерные корреляционные паттерны в интерпретируемый двумерный семантический ландшафт, где цвета указывают на принадлежность к кластеру, а контурные линии показывают вариации плотности в пространстве определений.

Для подтверждения надежности нашего семантического анализа и минимизации потенциальных искажений, специфичных для каждой модели, мы независимо воспроизвели все аналитические процессы на трех современных моделях линейных моделей: Claude 3.7 Sonnet, GPT-4o и Llama-3.3 70B Instruct. Затем эти три матрицы были усреднены для получения итоговой корреляционной матрицы, учитывающей межмодельные искажения и интегрирующей нюансы каждой модели в итоговый набор признаков. Эта методология была реализована с использованием Python, весь код доступен в виде открытого репозитория GitHub в Дополнительном коде.

Сборник экспертных определений и ручная классификация

В данном исследовании мы сосредоточились на определениях жизни, данных современными исследователями (а не классическими историческими личностями). Участники были отобраны на основе их рецензируемых статей в научной литературе, которые затрагивали (или требовали выражения мнения) по вопросу жизни.

В окончательный анализ было включено шестьдесят восемь определений. Текущая принадлежность респондентов к четырнадцати различным странам была представлена наиболее распространенными Соединенными Штатами (41) и Соединенным Королевством (9).

Метод отбора экспертов представлял собой неслучайную целенаправленную выборку экспертов (Страттон, [2024](#)), сформированную из расширенной профессиональной сети главного исследователя (ГИ) и основанную на его знакомстве с его работой. Целенаправленная выборка экспертов, приглашенных к участию, представляла собой неслучайную выборку экспертов со специфическими, релевантными знаниями, что являлось необходимым условием для нашего глубокого, контекстуально насыщенного анализа. Критерии отбора включали следующее: (1) респонденты должны были обладать признанными знаниями в своей области, (2) должны были занимать руководящие должности, (3) должны были иметь существенный академический или профессиональный вклад и (4) должны были иметь отношение своей области исследования к вопросу определения «жизни», но не ограничиваться ею (т.е. междисциплинарная работа). Эти критерии отбора экспертов были направлены на обеспечение высокого уровня и достаточной глубины знаний по теме жизни, которую мы решили исследовать, включая как теоретиков, так и ученых-исследователей. Метод «снежного кома» не применялся, то есть респонденты не приглашали и не рекомендовали участие новым людям из своей расширенной сети контактов, чтобы избежать высокой степени предвзятости, связанной с этим методом.

Респонденты представляли несколько дисциплин, включая биологию, философию, информатику, физику, искусственный интеллект и теорию систем. Среди них было много

ученых, работающих в области теоретической и междисциплинарной биологии. Большинство (около 50 из 68) получили образование в основном в качестве академических исследователей, около 14 занимали гибридные должности (например, сочетание любой из следующих ролей: академический исследователь, врач, философ, специалист по информатике или руководитель организации), и один респондент занимался преимущественно научным образованием и публичной научной деятельностью. Респонденты не были отобраны исключительно как врачи, консультанты/эксперты по политике, юристы, этики, исследователи в области гуманитарных наук (философы, социологи, музыканты/художники и т. д.) или представители промышленности. Кроме того, мы не связывались с исследователями из областей смерти, умирания и горя. Мы признаем отсутствие экспертов из этих дисциплин как ограничение и область для будущих исследований, поскольку эксперты из этих разнообразных областей могли бы сформировать анализ и внести свой вклад в обсуждение жизни актуальным и полезным образом.

Поскольку отбор осуществлялся на основе профессиональной сети руководителя проекта, мы признаем, что это может отдавать предпочтение определенным точкам зрения и недопредставлять другие. Мы признаем недопредставленность голосов и ученых из других сообществ, географических регионов и школ мысли за пределами нашей выборки. Целенаправленная выборка неизбежно, хотя и непреднамеренно, может оставить без внимания огромное количество голосов, которых мы специально приглашаем в будущих работах для обогащения дискуссии. Мы подчеркиваем, что представленный здесь вычислительный метод теоретически может учитывать неограниченное количество определений, и прилагаемый код с открытым исходным кодом может быть применен к любому альтернативному набору сбалансированных, междисциплинарных определений от экспертов, помимо тех, которые были отобраны в нашей целенаправленной выборке. Более подробная информация об ограничениях отбора экспертов в отношении вычислительного анализа содержится в разделе ниже (раздел «Обсуждение — Методологические инновации и ограничения»).

После отбора с экспертами-авторами связывались посредством интервью по электронной почте, отправляя письма на служебные адреса электронной почты респондентов, при этом было обеспечено сбалансированное представительство различных областей и дисциплин, включая биологию, информатику, физику и инженерию.

Респонденты не были осведомлены об ответах или ходе мыслей друг друга, им не были даны конкретные указания или критерии включения/исключения, и они самостоятельно отвечали за представление собственных определений. У респондентов не было установленных временных ограничений для формулирования и структурирования ответов, им разрешалось обращаться к внешним источникам, и не было установлено никаких правил, запрещающих контакты с коллегами по этому вопросу. Респондентам было предписано стремиться к тому, чтобы их ответы не превышали трех предложений, и им было сказано, что они должны думать о жизни настолько широко, насколько им хочется, не ограничиваясь, например, земной жизнью. Им не были даны никакие указания относительно количества слов, и их предупредили, что их ответы будут проанализированы в рукописи для публикации.

Два из предложенных определений были предоставлены авторами данного анализа. Как и все остальные респонденты, два автора, предоставившие определения, не видели никаких

других представленных определений жизни до того, как сформулировали свои собственные. В дальнейшем анализе этим двум ответам не был присвоен дополнительный вес.

В редакционных целях небольшая часть ответов была сокращена без ущерба для основных выводов определений авторов. Это было сделано для поддержания стандартизированной длины всех определений. Наконец, экспертам была предоставлена возможность проверить свои определения на предмет правильности указания авторства, но им настоятельно не рекомендовалось вносить какие-либо изменения на этом этапе, поскольку они уже ознакомились с некоторыми другими определениями.

После получения всех определений была проведена ручная проверка и категоризация, чтобы результаты стали понятными для человека, прежде чем переходить к анализу с помощью ИИ. Определение одного автора могло быть отнесено к нескольким категориям. Строка «Другое» вводилась, когда определение не вписывалось в одну из существующих категорий. Полученные категориальные различия представлены в Таблице [1](#). Все 68 предоставленных определений жизни представлены в Дополнительной Таблице [1](#).

Таблица 1. Классификация определений жизни.

Критерии: Жизнь определяется определенными:	Исследователи, чьи определения включают эти критерии
Термодинамика/энергетика	Болл, Дэвис, Дуссутур, Фрэзе, Георгиев, Гилберт, Хуан, Ингбер, Лейн, Митчелл, Ньюман, Пикард, Солмс, Сото, Сатмари, Тушински, Вальверду, Нанн
Информация/шаблон	Экли, Адами, Бонгард, Кейвс, Дас, Додиг-Црнкович, Георгиев, Ингбер, Левин, Миллер, Митчелл, Джентили, Нанн, Павлич, Ребер, Соле, Витковски
Сложность	Георгиев, Джексон, Сломан, Нанн
Вычисления	Агуэра и Аркас, Степни, Экли, Витковски, Додиг-Црнкович
Динамика (включая самоорганизацию)	Агуэра-и-Аркас, Бонгард, Браш, Кейвс, Дас, Додиг-Црнкович, Дуссутур, Фристон, Фрэзе, Георгиев, Гилберт, Хейлиген, Хуанг, Кауфман, Митчелл, Пицци, Рехави, Соле, Сатмари, Степни, Витковски
Автономия/субъектность/целеустремленность	Балл, Дюссутур, Эллис, Яблонка, Кальво, Хейлиген, Кракауэр, Ландер, Лион, Митчелл, Нобл, Шапиро, Сольмс, Сото, Тушинский, Вальверду, Кейвс

Критерии: Жизнь определяется определенными:	Исследователи, чьи определения включают эти критерии
Познание/интеллект	Дюссутур, Фонтана, Джентили, Кракауэр, Левин, Маршалл, Миллер, Нанн, Пикар, Ребер, Лион, Додиг-Крнкович
Конструкция/архитектура	Балушка, Георгиев, Ньюман, Пещеры
Функциональность/поведение	Альбантакис, Бал, Пещеры, Дюссутур, Дас, Лион, Зонненшайн, Султан
Репликация/воспроизводство/эволюция/наследственность	Экли, Адамацки, Адами, Дэвис, Ингбер, Яблонка, Джексон, Лендер, Джентили, Нанн, Рэтклифф, Шапиро, Сломан, Сатмари
Состав материала	Балушка, Дас, Кауфман, Лейн, Ньюман, Пицци, Рэтклифф, Сломан
Против строгого определения	Альбантакис, Болл, Фрэнк, Гунавардена, МакШи, Стэнли, Вонг, МакШи
Другой	Чауника, Филдс, Хоффман, Маршалл, Витковски, Уотсон

Перекрывающийся характер этих категорий, с множеством определений, охватывающих несколько классов и требующих категории «Другое», наглядно демонстрирует необходимость анализа непрерывного семантического пространства. Вместо того чтобы рассматривать эти пересечения как ошибки классификации, мы видим в них свидетельство того, что экспертное мышление о жизни существует в нескольких неортогональных измерениях, что делает дискретную категоризацию сложной задачей. Это наблюдение мотивирует наш вычислительный подход, который отображает взаимосвязи респондентов в многомерном пространстве без навязывания искусственных границ.

Количественный попарный корреляционный анализ

Для количественной оценки семантических связей между определениями мы разработали структуру парного корреляционного анализа с использованием вывода на основе LLM. Для каждой пары определений в нашем наборе данных мы вычислили коэффициент корреляции в диапазоне от -1,0 (полное несогласие) до 1,0 (полное согласие), используя следующую структуру запроса:

```prompt`

Analyze the following two definitions of life, where:

- 1.0 = Fundamentally opposing or incompatible primary frameworks;
- 0.5 to -0.9 = Significantly different emphasis with some contradiction;
- 0.0 = Independent or orthogonal frameworks;
- 0.1 to 0.4 = Slight overlap in secondary elements;
- 0.5 to 0.9 = Significant overlap with some differences;
- 1.0 = Aligned core frameworks and secondary elements.

Definition 1: `{{definition_1}}`

Definition 2: `{{definition_2}}`

What is the correlation metric between -1.0 and 1.0 for these two definitions?  
Respond with ONLY a single number!

````

Для повышения надежности и учета потенциальной изменчивости результатов LLM мы выполнили несколько повторных выводов ($n = 3$) для каждой пары определений и вычислили средний коэффициент корреляции и стандартное отклонение. Этот коэффициент корреляции был вычислен для каждой пары ответов (n^2), в результате чего была получена корреляционная матрица определений респондентов, кодирующая попарное сходство ($0 < r < 1$) и несходство ($-1 < r < 0$).

Здесь крайне важно отметить, что, хотя мы определили корреляцию между респондентами, сгенерировав оценки сходства попарно с учетом избыточности по трем уникальным LLM, полученные матрицы не отражают корреляцию в том виде, в котором она обычно определяется надежными статистическими методами анализа, такими как коэффициенты Пирсона или Спирмена. Вместо этого мы используем термин «корреляционные матрицы» для определения сходства, ранжированного на основе избыточных попарных сравнений.

Парный корреляционный анализ сгенерировал две матрицы $n \times n$, где n — количество определений: (1) корреляционную матрицу **M**, содержащую средние значения корреляции между всеми парами определений, и (2) матрицу стандартных отклонений **S**, отражающую изменчивость оценок корреляции. Обе матрицы были симметризованы путем усреднения с их транспонированной матрицей ($\mathbf{M}' = (\mathbf{M} + \mathbf{M}^T)/2$), чтобы учесть потенциальные эффекты ранжирования в выводе корреляции, полученном с помощью LLM, в результате чего было получено в общей сложности 6 усредненных случаев вывода для генерации каждой метрики парной корреляции. Полученные корреляционные матрицы были визуализированы в виде тепловых карт с использованием дивергентной цветовой схемы, где значения корреляции варьировались от -1 (несогласие, красный) до 1 (согласие, зеленый).

Этот анализ был повторен с моделями Claude 3.7 Sonnet, Llama 3.3 70B и GPT-4o, при этом окончательная корреляционная матрица была получена из среднего значения результатов каждого анализа. Каждая предоставленная метрика сходства (далее именуемая корреляцией) представляет собой результат 18 усредненных метрик корреляции, сгенерированных LLM, по шесть из каждой протестированной модели.

Данная методология корреляции на основе LLM поднимает вопросы о том, отражают ли показатели сходства подлинное концептуальное соответствие или поверхностные лингвистические особенности, такие как общий словарный запас, стилистические модели или общие отсылки из обучающих данных моделей. Несколько факторов указывают на то, что эти показатели отражают более глубокую семантическую структуру:

1. Высокая межмодельная корреляция ($r > 0,7$), несмотря на различия в обучающих данных и архитектурах, указывает на сходимость к стабильным закономерностям, выходящим за рамки индивидуальных смещений моделей.
2. Определения с высокими показателями корреляции часто используют разную лексику, в то время как определения с низкими показателями корреляции иногда имеют общую терминологию, что говорит о том, что модели выявляют концептуальные рамки, а не простое совпадение слов, что является постоянной проблемой для традиционных методов векторного представления текста.
3. Возможность интерпретации полученных кластеров в рамках устоявшихся философских концепций обеспечивает внешнюю проверку достоверности.

Тем не менее, мы признаем, что эти оценки по своей сути сочетают концептуальное сходство с эффектами обучения, специфичными для каждой модели, что является методологическим ограничением, аналогичным тому, как научные сообщества формируют общее значение как посредством индивидуального мышления, так и посредством коллективного духа времени.

Агломеративная кластеризация корреляционной матрицы

Для выявления естественных групп в пространстве определений мы применили иерархическую агломеративную кластеризацию к симметризованной корреляционной матрице. Этот подход к кластеризации «снизу вверх» начинается с того, что каждое определение рассматривается как отдельный кластер, и итеративно объединяет наиболее похожие кластеры до тех пор, пока все определения не будут принадлежать одному кластеру, создавая иерархическую структуру, которая выявляет многомасштабные взаимосвязи между определениями.

Сначала корреляционная матрица была преобразована в матрицу расстояний с использованием соотношения $d = \sqrt{2(1-r)}$, где r представляет собой коэффициент корреляции. Это преобразование сопоставляет коэффициенты корреляции от -1 до 1 с расстояниями от 0 до 2, при этом более высокие корреляции соответствуют меньшим расстояниям между определениями, а операция возведения в квадрат увеличивает расстояния между значениями высокой корреляции, что делает алгоритм кластеризации более чувствительным к тонким семантическим различиям между тесно связанными определениями.

В качестве метода агломерации мы использовали метод полной связи (максимальное расстояние между любыми двумя элементами из разных кластеров, также известный как связь по наиболее удаленному соседу). Метод полной связи был выбран за его способность создавать компактные, четко разделенные кластеры, которые подчеркивают концептуальные границы между семантическими группами, что делает его особенно подходящим для анализа определений, где желательны четкие разграничения.

Процесс построения дендрограммы имитирует организацию идей в постепенно расширяющиеся концептуальные категории. Вначале каждое определение существует как отдельная перспектива; каждый лист встраивается в дерево с одним узлом. Затем алгоритм определяет два наиболее похожих определения и объединяет их на высоте, пропорциональной их сходству. Тесно связанные определения объединяются в нижней части дендрограммы, а более отдаленные концептуальные родственники — выше. Этот процесс продолжается итеративно: отдельные определения либо присоединяются к существующим группам, либо группы объединяются с другими группами, всегда соединяя наиболее похожие сущности на каждом шаге. Полученная древовидная структура затем показывает, какие определения относятся друг к другу как связанные концептуальные определения, а также относительное семантическое расстояние между ними.

Для последующего проведения кластеризации без предварительного обучения на этой дендрограмме мы проанализировали закономерности расстояний слияния, чтобы определить точки перегиба, где дальнейшее слияние кластеров внезапно объединит существенно различные концептуальные модели. Этот «изгиб» в полученной матрице связей показывает естественное количество кластеров, присутствующих в данных, как точку, где слияние кластеров начнет скрывать важные концептуальные различия. Это решение для кластеризации разделяет определения на группы, которые разделяют фундаментальные концептуальные модели, сохраняя при этом значимое разделение между различными философскими или научными подходами к определению жизни. При наложении на отсортированную корреляционную матрицу определений респондентов эта многомасштабная перспектива позволяет нам исследовать как тонкие различия внутри концептуальных моделей, так и более широкие закономерности во всем пространстве определений.

LLM Кластерный семантический анализ

Для каждого кластера, идентифицированного с помощью иерархической агломеративной кластеризации, мы использовали LLM для выполнения двух взаимодополняющих аналитических процессов: (1) внутрикластерного тематического анализа для характеристики концептуальных рамок и (2) генерации консенсусного определения для выделения основных общих элементов. Для генерации семантического анализа матрицы кластерной корреляции для нескольких моделей использовался Claude 3.7 Sonnet.

Тематический анализ внутри кластера проводился с использованием структурированного задания, разработанного для систематического выявления закономерностей в кластерах определений. Этот протокол предписывал участнику программы LLM анализировать определения с трех последовательных точек зрения:

```prompt

1. WHAT Are The Core Ideas?
    - List every key concept mentioned
    - Count how often each appears
    - Group similar concepts together
    - Note which concepts appear most
    - Mark which are always present
  2. HOW Do Ideas Connect?
    - Find concepts that link to others
    - Map which ideas depend on others
    - Note concept hierarchies
    - Identify central hub concepts
    - Track idea flow patterns
  3. WHY This Structure?
    - Identify shared a priori frameworks
    - Find similar starting points
    - Track reasoning patterns
    - Mark scope boundaries
    - Note definitional strategies
- ```

---

Все вопросы в полном объеме доступны в Дополнительном коде. Этот тематический анализ был затем использован для проведения мета-обзора различных определений понятия «что такое жизнь?» с количественной, статистически обоснованной и вычислительной точки зрения. В дополнение к тематическому анализу мы также использовали LLM для синтеза единого репрезентативного определения для каждого кластера. Этот протокол явно ограничивал LLM требованиями, определенными в следующей структуре вопросов:

---

```prompt

You are tasked with synthesizing a consensus definition of life from a group of experts who have similar perspectives. Below are their definitions:

{{definitions}}

In addition to these provided cluster definitions, you have also conducted a prior thematic review which contains a meta-analysis of the key concepts being discussed by this group of respondents:

{{cluster_analysis}}

- Start: "Life is..."
- Content: Only majority-shared concepts (>50% frequency)
- Language: Technical terms from source definitions
- Structure: Logical flow of connected concepts
- Style: Match source complexity and tone
- Length: Within $\pm 20\%$ of median definition length
- Exclude: Unique views, novel terms, explanations

```

---

Этот алгоритмический подход к формированию консенсуса гарантировал, что полученные определения действительно отражали основные тенденции внутри каждой группы, а не вводили новые концепции или произвольные интерпретации. Ограничивая содержание концепциями, разделяемыми большинством, и используя только терминологию, присутствующую в исходных определениях, мы сохранили верность первоначальным экспертным точкам зрения, одновременно выделяя их общее концептуальное ядро.

Полученные в результате тематические анализы и согласованные определения предоставляют взаимодополняющие точки зрения на каждую группу: тематический анализ характеризует концептуальный ландшафт и интеллектуальный подход, общие для определений внутри группы, в то время как согласованное определение сводит эти общие элементы к единому связному утверждению, представляющему центральную точку зрения группы на то, что составляет жизнь. Вместе эти анализы выявляют как отличительные концептуальные рамки, характеризующие различные подходы к определению жизни, так и основные элементы, объединяющие определения внутри каждой рамки.

Наконец, был выполнен третий вызов функции вывода для генерации названия заголовка для каждого кластера. Имея в своем распоряжении согласованное определение и тематический анализ в контексте запроса, LLM было предложено сгенерировать краткое название кластера в соответствии со следующим запросом:

---

**```prompt**

Based on the consensus definition and thematic analysis provided, generate a SINGLE WORD or VERY SHORT PHRASE (2-4 words maximum) that captures the DISTINCTIVE ESSENCE of this specific cluster.

The title should be:

1. HIGHLY SPECIFIC to the philosophical, scientific, or conceptual framework unique to this cluster
2. DISTINCTIVE enough that it wouldn't apply equally well to other clusters of life definitions
3. TECHNICALLY PRECISE, using domain-specific terminology where appropriate
4. CONCEPTUALLY FOCUSED on the core unifying principle of these definitions

---

Provided below is the analysis of the given cluster:

{{cluster\_analysis}}

Provided below is the derived consensus definition for the given cluster:

{{consensus\_definition}}

```

Затем эти названия были сопоставлены с группами кластеров, полученными с помощью агломеративной кластеризации, и использованы для определения кластеров, которые впоследствии были отображены с помощью описанной методики двумерной проекции.

Двумерная проекция ландшафта определений

Для создания интерпретируемой визуализации ландшафта определений мы применили метод уменьшения размерности t-SNE (t-распределенное стохастическое встраивание соседей) для проецирования высокоразмерных корреляционных признаков в двумерное репрезентативное пространство. В отличие от анализа главных компонент (PCA), t-SNE специально акцентирует внимание на сохранении локальных взаимосвязей соседства, что делает его особенно подходящим для визуализации кластеров семантически связанных определений. Важно отметить, что t-SNE — это инструмент визуализации, который сохраняет локальную структуру соседства, а не принципиальное разложение по размерности, как PCA, и его глобальные ориентации осей математически произвольны. Однако, когда интерпретируемые градиенты последовательно появляются в нескольких запусках с различными начальными условиями и в разных анализах LLM, эта согласованность предполагает, что эти измерения отражают подлинную структуру в экспертном дискурсе, а не вычислительные артефакты. Мы интерпретируем эти возникающие оси с философской точки зрения, признавая при этом, что они являются алгоритмическими конструкциями, которые раскрывают, а не навязывают концептуальную организацию. Это было подтверждено тестированием с использованием многомерного шкалирования (MDS), анализа главных компонент (PCA) и равномерной аппроксимации и проекции многообразия (UMAP), при этом t-SNE неизменно обеспечивал наилучшую пространственную проекцию базовых корреляционных вложений.

Сначала корреляционная матрица была преобразована в матрицу расстояний с использованием соотношения $d = \sqrt{2(1-r)}$, где r представляет собой коэффициент корреляции между парами определений — то же самое преобразование расстояний, которое применяется к агломеративной кластеризации. Полученная матрица расстояний затем была симметризована путем усреднения с ее транспонированной матрицей ($M' = (M + M^T)/2$), чтобы учесть потенциальные асимметрии в процессе преобразования корреляционных характеристик в характеристики расстояний, эффективно размещая похожие по определениям ответы (высокая корреляция) в непосредственной близости, одновременно разделяя несхожие (низкая корреляция).

Итоговая проекция t-SNE визуализирует семантический ландшафт определений жизни, где каждая точка представляет ответ эксперта, окрашенный в соответствии с присвоенным ей кластером. Контурные линии были добавлены для выделения вариаций плотности в пространстве определений, выявляя области концептуальной конвергенции и дивергенции. Итоговая визуализация включает в себя присвоение кластеров, полученное в результате агломеративной иерархической кластеризации, причем каждый кластер помечен в соответствии с его названием, полученным с помощью LLM, которое характеризует объединяющую концептуальную структуру этой группы. Такой подход выявляет различные концептуальные территории по отношению к более широкому ландшафту определений, обеспечивая всестороннее пространственное отражение того, как эксперты концептуализируют жизнь в разных дисциплинах.

Результаты

Определения экспертов

В дополнительной таблице [1](#) представлены определения, предоставленные участниками опроса. Некоторые сосредоточились на невозможности (или бесполезности) того или иного определения, в то время как другие пытались решить эту задачу, исходя из различных концептуальных подходов. Собранные 68 определений представляют широкий спектр дисциплинарных перспектив, от традиционных биологических концепций до вычислительных подходов, от термодинамических взглядов, вдохновленных физикой, до когнитивных и философских позиций. Несмотря на вероятную невозможность определения действительно ортогональных категорий, по которым можно было бы классифицировать эти определения, мы попытались сделать это вручную (без помощи ИИ), используя критерии, представленные в таблице [1](#).

Ручная классификация этих определений выявила несколько преобладающих тем, при этом некоторые понятия повторялись в разных системах определений. Термодинамика и энергетика встречались в 18 из 68 (26%) всех определений, в то время как информационные/паттерно-ориентированные подходы были обнаружены в 25% определений. Динамика (включая самоорганизацию) была наиболее распространенной темой, встречающейся в 29% определений.

Мы дополнительно охарактеризовали наиболее заметные расходящиеся свойства, которые обычно обсуждаются (рис. [2](#)), выявив, что 57% определений занимали объективную позицию (определяя жизнь в терминах, не зависящих от наблюдателя), в то время как 43%

включали элементы, зависящие от наблюдателя. Кроме того, 54% определений рассматривали жизнь как непрерывное свойство, а не как бинарное состояние, отражая отход от категориальных различий между живыми и неживыми системами. Большинство определений (88%) были практически применимы в том смысле, что они предоставляли конкретные, общепринятые критерии, в то время как 12% использовали более вдохновляющие, поэтические или саморекурсивные подходы.

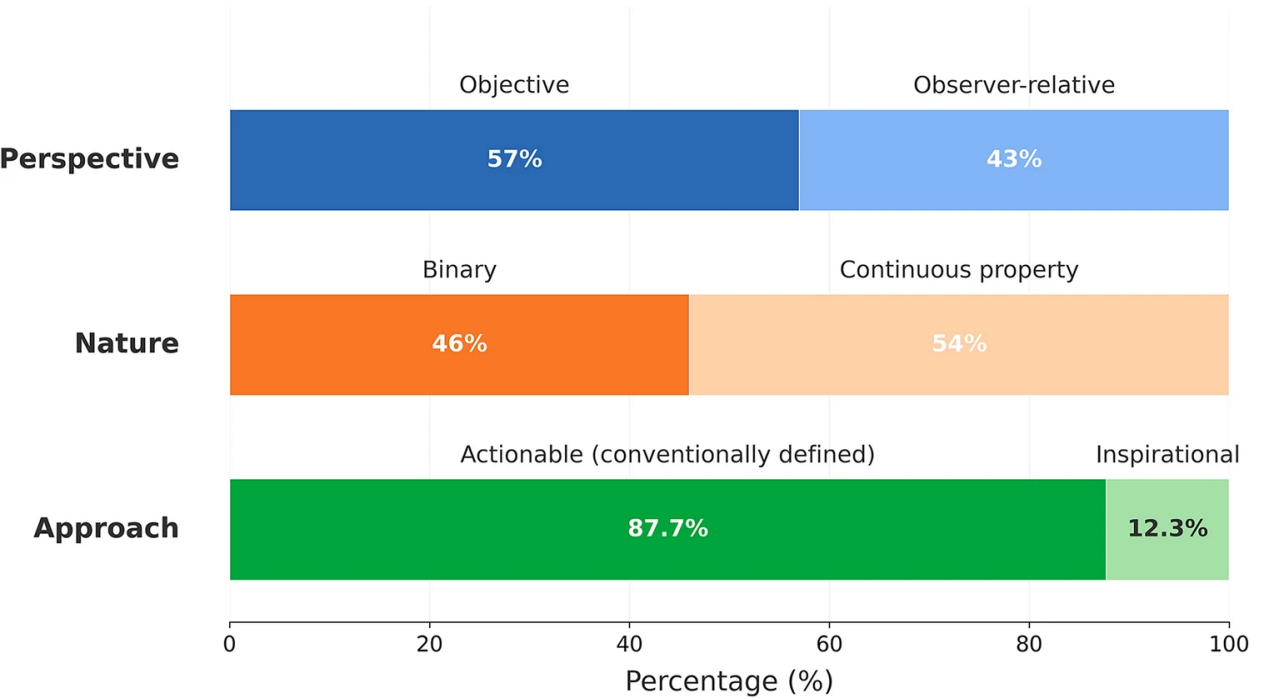


Рис. 2. Распределение общих свойств среди экспертных определений жизни ($n = 68$)

Парный корреляционный анализ

Для оценки семантических связей между экспертными определениями жизни мы проанализировали, как три современные модели LLM воспринимают концептуальное совпадение среди 68 определений. Каждая модель независимо оценивала все возможные пары определений, создавая корреляционную матрицу, которая количественно определяет сходство (-1,0 для полного противоречия до 1,0 для идеального совпадения) между каждой парой определений.

Для построения трех различных корреляционных матриц были независимо использованы модели Claude 3.7 Sonnet, Llama 3.3 70B Instruct и GPT-4o. Модель Claude 3.7 Sonnet продемонстрировала наиболее оптимистичную оценку с самым высоким средним коэффициентом корреляции, в то время как модель Llama 3.3 70B Instruct дала более консервативные оценки корреляции и показала большее стандартное отклонение в своих ответах. Примечательно, что все модели показали отрицательную асимметрию (-0,76 до -0,97), что указывает на систематическую тенденцию выявлять больше областей концептуального сходства, чем противоположностей. Общие статистические данные корреляционных матриц для каждой модели представлены в дополнительной таблице [2](#).

Несмотря на независимый анализ моделей, они продемонстрировали замечательную согласованность в общей оценке того, какие определения концептуально совпадают, а какие расходятся. Парные корреляции между матрицами были особенно высокими:

- Клод 3.7 Сонет против Ламы 3.3 70В Инструкция: $r = 0.7279$
- Клод 3.7 Сонет против GPT-4o: $r = 0,8103$
- Llama 3.3 70B Instruct vs. GPT-4o: $r = 0.7977$

Эти существенные корреляции позволяют предположить, что, хотя каждая модель вносила свой собственный аналитический подход к определению понятий, они независимо друг от друга сходились к схожим моделям концептуальных взаимосвязей.

Эти корреляционные матрицы представлены на дополнительных рисунках [1](#), [2](#) и [3](#) для Клода, Ламы и GPT соответственно. Полученная несортированная корреляционная матрица, усредненная по нескольким моделям, представлена на дополнительном рисунке [4](#). Кроме того, стандартные отклонения (матрица энтропии) для всех агрегированных случаев вывода доступны на дополнительных рисунках [5](#), [6](#) и [7](#) для Клода, Ламы и GPT соответственно, что подчеркивает согласованность оценок корреляции LLM для каждого респондента в повторных случаях вывода.

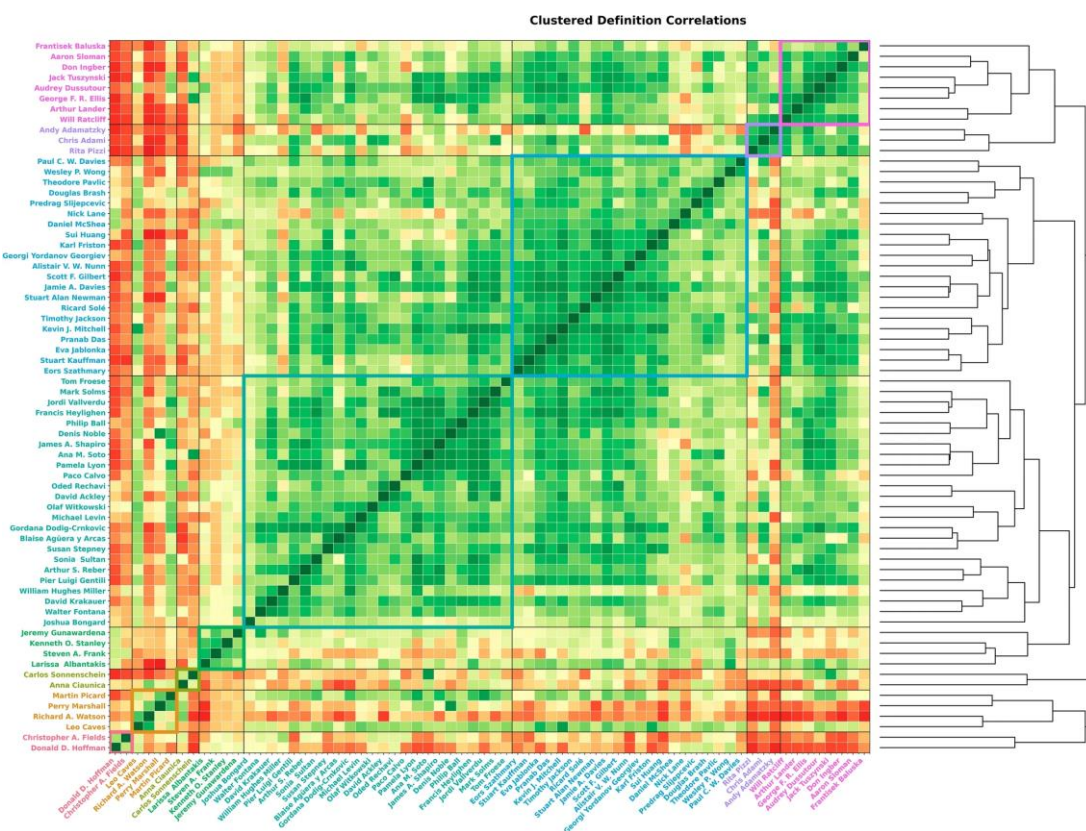


Рис. 3. Иерархически отсортированная корреляционная матрица определений респондентов, а также соответствующая дендрограмма, определяющая расстояние связи между определениями, выделяющая восемь различных концептуальных кластеров.

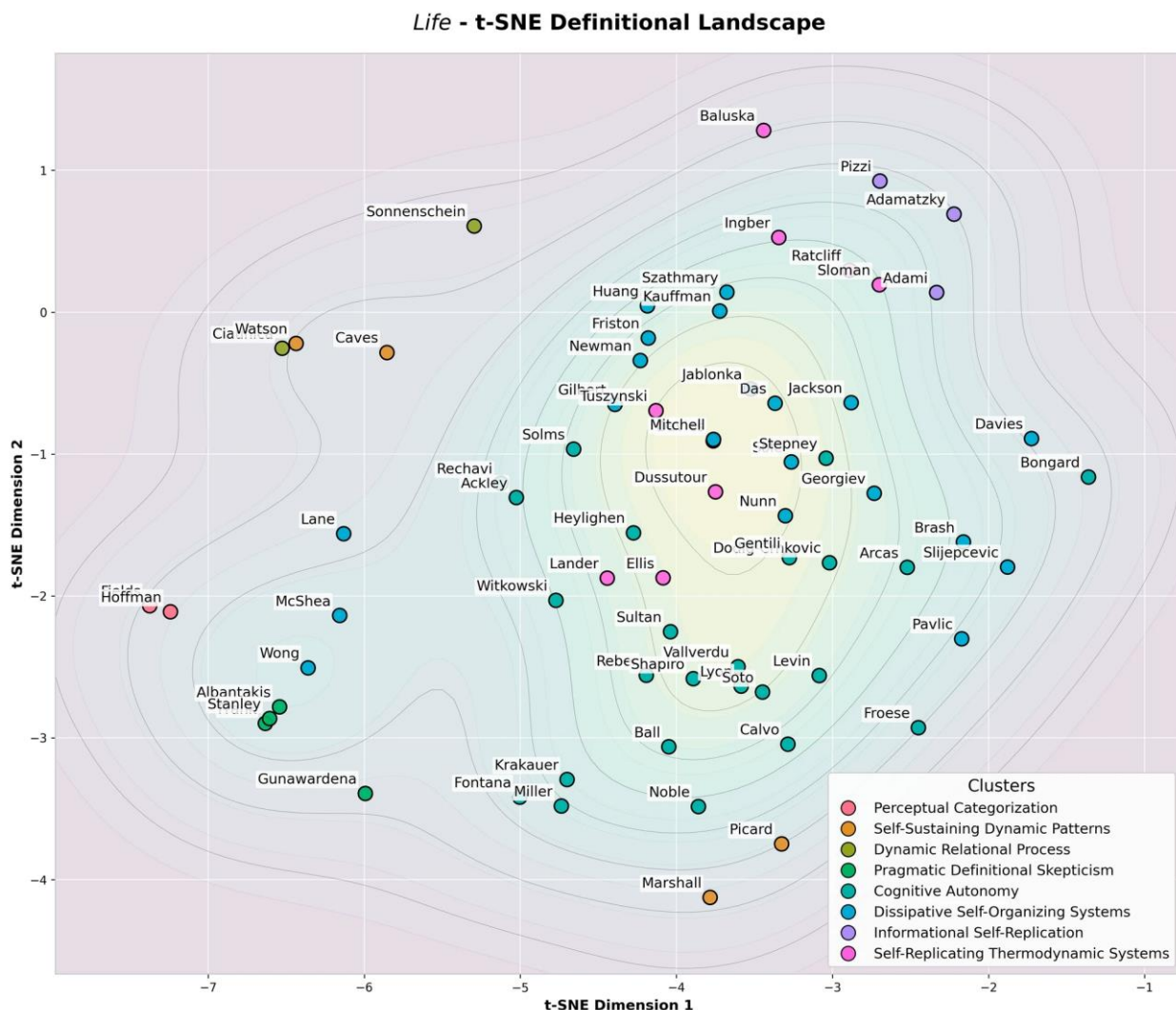


Рис. 4. Двумерная диаграмма определения t-SNE, показывающая распределение экспертных мнений по восьми концептуальным кластерам, полученная на основе средних матриц попарной корреляции из независимых анализов Claude 3.7 Sonnet, Llama 3.3 70B Instruct и GPT-4o. (Полноцветное изображение доступно в электронной версии.)

Для уменьшения потенциальных искажений, связанных с какой-либо отдельной моделью, и создания более надежной оценки семантических связей, мы вычислили поэлементное среднее по трем симметризованным корреляционным матрицам. Интеграция нескольких моделей позволила выявить основные тенденции всех трех отдельных анализов, о чем свидетельствуют высокие корреляции между матрицей каждой модели и усредненной матрицей:

- Сонет Клода 3.7 против усреднения по нескольким моделям: $r = 0,9017$
- Llama 3.3 70B Instruct vs. Multi-Model-Average: $r = 0.9294$
- GPT-4o против усреднения по нескольким моделям: $r = 0,9358$

Усредненная матрица, построенная на основе множества моделей, успешно отразила тонкие нюансы каждой модели, в конечном итоге обеспечив богатое кодирование

попарных корреляционных закономерностей между представленными определениями жизни.

Агломеративная кластеризация корреляционных матриц

Три протестированные модели LLM продемонстрировали различное поведение при кластеризации, выявив от семи до одиннадцати естественных кластеров с помощью анализа методом «локтя». Модель Claude 3.7 Sonnet дала наиболее интегрированное решение с семью кластерами и самым низким показателем контраста (0,258), в то время как модель Llama 3.3 70B Instruct продемонстрировала наиболее дискриминативный подход с одиннадцатью кластерами и самым высоким показателем контраста (0,563). Интеграция нескольких моделей дала восемь кластеров со сбалансированным показателем контраста 0,359, что позиционирует ее между интегративным и дискриминативным крайними значениями. Этот интегрированный подход обеспечил устойчивую внутрикластерную корреляцию ($0,509 \pm 0,187$) при сохранении умеренной межкластерной корреляции ($0,150 \pm 0,390$), что свидетельствует о том, что он уловил как отличительные особенности концептуальных рамок, так и значимые связи между ними. Статистические данные результатов кластеризации каждой протестированной модели доступны в дополнительной таблице [3](#).

На рисунке [3](#) представлена визуализация интегрированного анализа нескольких моделей, демонстрирующая как кластерную корреляционную матрицу (зеленым цветом обозначена высокая корреляция, красным — отрицательная), так и соответствующую структуру дендрограммы справа. Кроме того, кластерные корреляционные матрицы для независимых анализов Клода, Ламы и GPT представлены на дополнительных рисунках [8](#), [9](#) и [10](#) соответственно, чтобы показать, как каждая модель уникальным образом организует респондентов относительно друг друга.

Эта отсортированная корреляционная матрица выявляет отчетливую блочно-диагональную структуру, указывающую на сильную внутрикластерную согласованность, а также демонстрирует значимые межкластерные корреляции, предполагающие концептуальную преемственность в рамках всего ландшафта определений. Чтобы исследовать, насколько последовательно различные модели LLM группируют респондентов в одни и те же кластерные отношения, мы сравнили согласованность кластеризации из анализа каждой из трех моделей, а также из интеграции нескольких моделей.

- Клод 3.7 Сонет и Лама 3.3 70B Инструкция: 71.2% согласованности
- Сонет Клода 3.7 и GPT-4o: согласованность 74,8%.
- Сонет Клода 3.7 и среднее значение по нескольким моделям: согласованность 74,1%.
- Llama 3.3 70B Инструкция и GPT-4o: 75,7% согласованности
- Llama 3.3 70B Инструкция и средний показатель по нескольким моделям: согласованность 72,8%.
- GPT-4o и среднее значение по нескольким моделям: согласованность 79,1%.

Высокая степень согласованности между моделями (от 71,2% до 79,1%) свидетельствует о существенном согласии относительно того, какие определения следует группировать вместе, несмотря на различное количество кластеров, выявленных каждой моделью.

Для дальнейшей оценки устойчивости нашей кластеризации мы изучили, насколько последовательно отдельные определения группировались в трех моделях. Общая стабильность группировки составила 53,7%, что указывает на то, что чуть более половины всех пар определений последовательно группировались вместе (или последовательно разделялись) во всех моделях. Эта умеренная стабильность предполагает, что, хотя общая структура ландшафта определений является устойчивой, сохраняется подлинная неоднозначность в том, как следует классифицировать некоторые определения на границах между кластерами.

Некоторые определения продемонстрировали удивительно высокую согласованность группировки в разных моделях: определения Кристофера Филдса, Дональда Хоффмана (100% согласованность), Энди Адамацки, Лео Кейвса и Ричарда Уотсона (95,5% согласованность). Эти определения, по-видимому, занимают четкие, отличительные позиции в семантическом ландшафте. В отличие от них, определения Марка Солмса (28,4%), Аарона Сломана (26,9%), Пола Дэвиса (23,9%), Тома Фрозе (23,9%) и Уэсли Вонга (22,4%) показали самую низкую согласованность, что предполагает, что эти определения охватывают несколько концептуальных рамок или занимают пограничные позиции между устоявшимися кластерами. Глядя на рис. 3, мы можем заметить, что стабильные определения, как правило, появляются в блоках наиболее отчетливо окрашенных цветов, в то время как нестабильные часто появляются в переходных зонах между кластерами.

Внутрикластерный семантический анализ

На основе восьми отдельных кластеров, полученных методом агломеративной кластеризации усредненной матрицы попарных корреляций нескольких моделей, было проведено дальнейшее исследование тематического содержания каждого кластера. Согласованные определения в таблице 2 представляют собой квинтэссенцию основных концепций, разделяемых в каждом кластере, что позволяет понять, как различные точки зрения подходят к вопросу о том, что составляет жизнь. Эти определения были получены с помощью LLM-анализа определений внутри каждого кластера, выделив концепции, разделяемые большинством, при сохранении технического языка и концептуальных рамок, характерных для каждой группы.

Таблица 2. Названия кластеров и согласованные определения.

| Заголовок | Консенсусное определение |
|----------------------------|---|
| Перцептивная категоризация | <i>Жизнь — это перцептивная категория, возникающая из того, как системы, считающие себя живыми, распознают и классифицируют другие, а не объективное различие с</i> |

| Заголовок | Консенсусное определение |
|--|--|
| | <p><i>присущими им свойствами. Кажущаяся граница между живым и неживым существует прежде всего как артефакт наших сенсорных ограничений, которые вводят искусственные различия в то, что могло бы быть более единой реальностью.</i></p> |
| <p>Самоподдерживающиеся динамические модели</p> | <p><i>Жизнь — это устойчивая динамическая модель, которая поддерживает себя посредством рекурсивных процессов, создавая и сохраняя собственные условия существования. Она проявляется как самосознательной системы, способной к трансформации, возникающей благодаря резонансу и гармоничному взаимодействию своих компонентов. Эта модель существует во взаимосвязи с окружающей средой, выражая сознание и постоянно воссоздавая себя посредством взаимопреобразующих связей.</i></p> |
| <p>Динамический реляционный процесс</p> | <p><i>Жизнь — это непрерывное движение, характеризующееся динамичным обменом между внутренней и внешней средой, в основе которого лежит взаимосвязь с уже существующими объектами. Жизнь не может существовать в изоляции или стагнации, поскольку её сущность включает в себя как движение, так и размножение. Жизнь не является чем-то, что могут создать люди, она принимается и передаётся, существуя как непрерывный процесс взаимосвязи.</i></p> |
| <p>Прагматический скептицизм в отношении определений</p> | <p><i>Жизнь — это контекстуальная конструкция, которую лучше рассматривать с точки зрения функциональной полезности, чем универсального определения, где усилия по определению должны служить конкретным исследовательским целям, а не устанавливать абсолютные границы. Концепция существует на пересечении научной и философской областей, что требует признания дисциплинарных ограничений и прагматического подхода, ориентированного на то, что способствует пониманию, а не на то, что представляет собой «правильную» характеристику.</i></p> |
| <p>Когнитивная автономия</p> | <p><i>Жизнь — это самоподдерживающаяся, целенаправленная система, которая обрабатывает информацию для адаптации к изменениям окружающей среды, сохраняя при этом свои организационные границы. Она функционирует как автономный агент, активно противодействующий энтропии посредством обмена энергией, принимая целенаправленные решения, обеспечивающие её дальнейшее существование. Жизнь функционирует как когнитивный процесс, который воспринимает, интерпретирует и реагирует на окружающую среду, изменяясь при</i></p> |

| Заголовок | Консенсусное определение |
|---|---|
| | <i>необходимости для поддержания стабильности, несмотря на меняющиеся условия.</i> |
| Диссипативные самоорганизующиеся системы | <i>Жизнь — это динамичный, далёкий от равновесия процесс, характеризующийся самоорганизацией и самоподдержанием посредством контролируемого рассеивания энергии. Она поддерживает полупроницаемые границы, отделяющие систему от окружающей среды, одновременно позволяя избирательный обмен материей и энергией. Живые системы обрабатывают, хранят и передают информацию посредством механизмов обратной связи, которые обеспечивают адаптацию и эволюцию во времени.</i> |
| Информационное самовоспроизведение | <i>Жизнь — это самовоспроизводящаяся система информации, закодированной в физической среде, способная к самовоспроизведению и поддержанию порядка, противостоящего естественному разрушению. Этот основанный на информации процесс позволяет передавать сложные структурные и функциональные данные из поколения в поколение, при этом физические компоненты служат в первую очередь носителями этой важной информации.</i> |
| Самовоспроизводящиеся термодинамические системы | <i>Жизнь представляет собой самоподдерживающуюся физическую и химическую систему, которая воспроизводит себя, сохраняя при этом свою организацию вне термодинамического равновесия. Она адаптивно взаимодействует со своей окружающей средой, поглощая и преобразуя ресурсы для поддержания своего размножения и роста. Живые системы проявляют автономию благодаря саморегуляции и реакции на стимулы, а их репродуктивные возможности обеспечивают эволюционные процессы.</i> |

Полный тематический анализ, проведенный для Клода, Ламы, GPT, и полученное в результате усреднение по нескольким моделям представлены в дополнительных файлах [1](#), [2](#), [3](#) и [4](#) соответственно.

t-SNE-проекция определений жизни

Метод уменьшения размерности t-SNE преобразовал многомерные корреляционные данные в интерпретируемую двумерную визуализацию 68 экспертных определений (рис. [4](#)). Метод t-SNE был выбран за его способность сохранять локальную структуру окрестности, одновременно выявляя глобальные закономерности в многомерных данных, но полный набор результатов кластеризации, полученных с помощью многомерного шкалирования (MDS), равномерной аппроксимации и проекции многообразия (UMAP), t-SNE и усреднения по ансамблю, представлен на дополнительном рис. [11](#).

Полученная проекция t-SNE на рис. 4 показывает непрерывный семантический ландшафт, организованный вдоль двух концептуально интерпретируемых осей. К этим многомерным интерпретациям следует подходить с осторожностью. Хотя глобальные ориентации осей t-SNE математически произвольны, согласованность, с которой возникают эти конкретные философские градиенты (в различных анализах LLM и при проверке на основе независимо закодированных признаков определения в таблице 1), предполагает, что они представляют собой подлинную структуру в том, как эксперты концептуализируют жизнь, а не артефакты алгоритма t-SNE. Оси не навязываются t-SNE, а выявляются им, аналогично тому, как t-SNE в транскриптомике отдельных клеток выявляет биологические взаимосвязи типов клеток, которые затем могут быть подтверждены с помощью независимых молекулярных маркеров (Kobak and Berens 2019).

Измерение 1 (горизонтальное) отображает переход от зависящих от наблюдателя перцептивных моделей (слева) к объективным, материально-структурным определениям (справа). Крайнее левое положение занимает изолированный кластер перцептивной категоризации, примером которого служит рекурсивная формулировка Филдса: «быть живым — значит считаться живым системами, которые считают себя живыми». Это резко контрастирует с кластерами справа, которые подчеркивают конкретные механизмы, такие как репликация и термодинамические свойства, отраженные в концепции жизни Адами как «информации, способной к саморепликации».

Измерение 2 (вертикальное) выявляет фундаментальное онтологическое противоречие между концептуализацией жизни, основанной на сущностях, и концептуализацией жизни, основанной на процессах, воспроизводя историческую диалектику от картезианского механизма к аристотелевской телеологии. Верхние позиции занимают определения, подчеркивающие материальные структуры, свойства компонентов и осязаемые механизмы, такие как «Живые клетки и все их конструкции» Балушки, акцент Пицци на «элементах на основе углерода с самоорганизующимися и самовоспроизводящимися свойствами» и «компоненты, способные расширяться и самовоспроизводиться» Сломана. Нижняя область содержит ориентированные на процессы когнитивно-феноменологические концепции, представленные метафорическим «огнем, который зажигает себя» Маршалла, «самоподдерживающейся системой, способной к исцелению, которая использует и организует энергию для достижения выражения(й) сознания» Пикара и «самосоздающей силой» Ноубла. Этот вертикальный градиент отражает извечное противоречие между субстанциалистскими взглядами, рассматривающими жизнь как совокупность определенных сущностей (перекликаясь с картезианской и гоббсовской концепциями), и телеологическими взглядами, которые концептуализируют жизнь как целостные модели организации, потока информации и возникающей самоактуализации.

Доминирующий центральный аттрактор включает два частично перекрывающихся кластера: когнитивную автономию (бирюзовый, $n = 24$) и диссипативные самоорганизующиеся системы (синий, $n = 21$). В этих кластерах сосредоточено 66% определений в зоне высокой плотности, где физические, информационные и агентные перспективы интегрированы с помощью различных подходов к определению.

Эта концентрация поднимает важный интерпретационный вопрос: представляет ли этот центральный аттрактор подлинную концептуальную конвергенцию к интегрированному пониманию жизни, или же он отражает то, как институциональная наука и общая

дисциплинарная подготовка ограничивают пространство профессионально выражаемых позиций? Высокая плотность может указывать либо на формирующийся консенсус, либо на границы того, что наши опрошенные респонденты считают концептуально приемлемым, что усугубляется предвзятостью отбора среди тех, кого просили дать свое определение. Периферийные кластеры, хотя и малонаселенные, могут, следовательно, представлять собой не выбросы, а авангардные позиции, которые бросают вызов неявным предположениям, структурирующим центральную область; эти периферийные определения отражают позиции, ставящие под сомнение фундаментальные предположения, такие как необходимость объективных материальных свойств для жизни (перцептивная категоризация) или вообще служит ли определение жизни научному прогрессу (прагматический скептицизм в отношении определений).

Контурные структуры демонстрируют различную семантическую плотность по всей области. Наиболее разреженные регионы окружают изолированный кластер перцептивной категоризации, в то время как градиентные переходы наблюдаются между кластерами, ориентированными на процессы (оранжевый и желтый), и центральными аттракторами. В правом квадранте отображаются два различных полюса: информационное самовоспроизведение (фиолетовый, $n = 3$) достигает максимального структурного расстояния от перцептивных кластеров, в то время как самовоспроизводящиеся термодинамические системы (розовый, $n = 8$) соединяют традиционные биологические перспективы с информационно-теоретическими подходами.

Непрерывный характер этого семантического пространства дополнительно подчеркивается определениями, которые функционируют как концептуальные мосты между кластерными территориями. Характеристика жизни Уотсоном как «образца и процесса любви — глубоко уязвимого взаимного танца» создает семантический путь между перцептивно-реляционными и телеологическими рамками. Аналогично, концепция Ноубла «Жизнь — это самосоздающаяся деятельность» занимает переходное положение между процесс-ориентированными подходами и когнитивными рамками, явно связывая аристотелевскую самоактуализацию с современными перспективами, основанными на субъектности. Концепция Левина «Живые существа помнят и предвосхищают грубозернистый опыт в непрерывный действенный сон» соединяет когнитивные и физические перспективы посредством концепций, объединяющих обработку информации с термодинамическими ограничениями, демонстрируя, как современные определения продолжают преодолевать историческое напряжение между механизмом и телеологией.

Выявление интерпретируемых измерений в результате вычислительного семантического анализа подтверждает целесообразность этого подхода как методологического моста между редукционистским и целостным подходами к фундаментальным вопросам науки и философии. Вместо навязывания заранее заданных категорий, методология выявляет скрытую структуру в экспертном дискурсе, демонстрируя, как вычислительные инструменты могут выявлять закономерности концептуальной согласованности, которые в противном случае могли бы остаться скрытыми из-за фрагментации дисциплин. Высокая конвергенция между различными моделями LLM (коэффициенты корреляции $> 0,7$) дополнительно подтверждает надежность этих закономерностей, а последовательное выявление концептуальных мостов и переходных зон в разных моделях предполагает, что они представляют собой подлинные особенности ландшафта определений, а не аналитические артефакты.

Кластерный семантический анализ

Восемь концептуальных кластеров, выявленных с помощью агломеративной кластеризации, представляют собой различные, но взаимосвязанные перспективы в рамках определений, каждый из которых занимает уникальное семантическое положение, освещающее фундаментальные противоречия и сближения в современных представлениях о жизни.

Кластер перцептивной категоризации ($n = 2$) занимает наиболее дивергентное семантическое пространство, достигая максимальной концептуальной дистанции от материально-структурных рамок. Рекурсивная формулировка Филдса («быть живым — значит считаться живым системами, которые считают себя живыми») бросает вызов фундаментальным предположениям об объективной реальности. Позиционирование Хоффманом границы между живым и неживым как «артефакта ограничений наших чувств» перекликается с квантово-механическими представлениями об эффектах наблюдателя, предполагая неисследованные связи между определением жизни и фундаментальной физикой.

Группа «Самоподдерживающиеся динамические паттерны» ($n = 4$) объединяет радикальную эпистемологию с философией процесса посредством постоянного внимания к взаимоотношениям и сознанию. Характеристика жизни Уотсоном как «паттерна и процесса любви — глубоко уязвимого взаимного танца» вводит аффективные измерения, отсутствующие в механистических рамках, в то время как «огонь, который зажигает себя сам» Маршалла и рекурсивная формулировка Кейвса подчеркивают важность парадоксальной *самореференции* для понимания природы жизни. Корреляционные паттерны этой группы выявляют концептуальное сходство как с зависимыми от наблюдателя рамками, так и с некоторыми элементами когнитивной автономии, предполагая потенциальные пути интеграции между феноменологическим и научным подходами.

Динамический реляционный процесс ($n = 2$) сводит реляционность к ее кинетической сущности, а утверждение Зонненшайна о том, что «Без размножения и движения нет жизни», устанавливает четкие граничные условия. Аналогично, утверждение Чиауники о том, что жизнь «не может существовать без другого, уже существующего, уже живого», представляет собой серьезную проблему для теорий происхождения жизни, предполагающих первозданную изоляцию. Этот минимальный кластер создает мосты между абстрактной реляционностью и классическими биологическими требованиями, позиционируя себя как фазовый сдвиг между философским и эмпирическим подходами. Его смежное положение с самоподдерживающимися динамическими паттернами отражает глубокое концептуальное сходство, а умеренные положительные корреляции с кластерами, ориентированными на репликацию, предполагают, что эта функция моста эффективно работает в разных масштабах.

Прагматический скептицизм в отношении определений ($n = 4$) выступает в качестве методологического контрапункта, противопоставляя себя утверждению Лейна о том, что «проведение линии через континуум всегда произвольно», что ставит под сомнение всю концепцию категориального определения. Уникальная корреляционная картина этой группы (отрицательные связи с жесткими граничными условиями, положительные связи с многомерными структурами) показывает, что она функционирует как концептуальный

посредник между конкурирующими парадигмами определения, а не отстаивает какую-либо конкретную позицию. Характеристика определений Франком как «инструментов, а не конечных точек» обеспечивает философское обоснование вычислительной методологии, используемой в этом исследовании.

Когнитивная автономия ($n = 24$) выступает в качестве основного концептуального аттрактора, её центральное положение и высокая численность отражают широкую междисциплинарную конвергенцию. Внутреннее разнообразие кластера охватывает кибернетические перспективы (информационный паттерн Витковского, способный к самовоспроизводству) и феноменологические подходы (процесс индивидуации Фрозе), показывая, как когнитивные концепции учитывают как механистические, так и эмпирические измерения. Примечательно строгое уравнение Додиг-Крнковича «Жизнь = познание», которое рассматривает вычислительные процессы как фундаментальные, а не производные. Умеренные положительные корреляции кластера почти со всеми другими концепциями (за исключением перцептивной категоризации) предполагают, что он функционирует как семантический центр, а не как авангардная позиция.

Диссипативные самоорганизующиеся системы ($n = 21$) обеспечивают дополнительную гравитационную силу за счет акцента на физических принципах. Их существенное совпадение с когнитивной автономией создает центральную зону конвергенции определений, формируя область высокой плотности, где концентрируется 66% всех определений. Характеристика жизни Павличем как «производителя статистической новизны» связывает термодинамическую необходимость с эволюционными инновациями, в то время как акцент Джексона на «уникальных механизмах хранения и извлечения информации» связывает физические процессы с функциональным возникновением. Эта группа объединяет механические и функциональные подходы посредством явного акцента на градиентах энергии и граничных условиях, которые обеспечивают, а не просто ограничивают биологическую организацию.

Информационное самовоспроизведение ($n = 3$) представляет собой редукционистский полюс пространства определений, где формулировка Адами — «Жизнь — это информация, которая может самовоспроизводиться» — обеспечивает максимальную точность за счет потенциальной потери полноты. Корреляционные паттерны этого кластера выявляют четкие границы, исключая зависимость от наблюдателя и реляционные модели, сохраняя при этом сильные положительные корреляции с самовоспроизводящимися термодинамическими системами. Небольшой размер кластера скрывает его концептуальное влияние как отправной точки для информационно-теоретических подходов, хотя строгие критерии среды для информационного самовоспроизведения могут исключить такие системы, как прионы или кристаллические формы жизни, которые размножаются без генетической информации.

Самовоспроизводящиеся термодинамические системы ($n = 8$) объединяют традиционные биологические требования с физическими ограничениями, создавая мост между классической генетикой и формирующимися физическими концепциями. Перечисление Дюссутуром взаимосвязанных возможностей (воспроизводство, автономия, реактивность, пластичность и использование энергии) иллюстрирует многофункциональный подход, характерный для этой группы. Внутреннее разнообразие простирается от акцента на клеточно-структурном аспекте (теория Балуски «Живые клетки и все их конструкции») до перспектив, ориентированных на процессы (акцент Ингбера на преобразовании

энергии), выявляя противоречия даже внутри традиционных биологических представлений о жизни.

Корреляционные закономерности между этими кластерами выявляют семантическую топологию, где концептуальная близость отражает лежащую в основе философскую схожесть. Математическая структура межкластерных взаимосвязей демонстрирует отчетливые закономерности, где центральные кластеры (когнитивная автономия, диссипативные самоорганизующиеся системы) показывают широкие положительные корреляции по всей области, в то время как периферийные кластеры демонстрируют избирательные закономерности сходства, которые создают концептуальные архипелаги, а не изолированные острова. Эта полевая структура с множеством аттракторов, оказывающих различное гравитационное воздействие, бросает вызов простым спектральным моделям пространства определений.

Границы кластеров функционируют скорее как переходные зоны, чем как четкие разграничения, при этом определения на периферии часто демонстрируют свойства соседних концептуальных моделей (например, «самосозидающаяся субъектность» Ноубла, которая попала в категорию когнитивной автономии, оставаясь при этом в непосредственной близости от кластера самоподдерживающихся динамических паттернов). Эти семантические градиенты облегчают концептуальную миграцию и объясняют, почему некоторые определения демонстрируют низкую стабильность группировки в различных анализах LLM: они занимают подлинные пограничные области, где математически обоснованными остаются множественные решения кластеризации. Сохранение этих архетипических тем в различных архитектурах LLM (коэффициенты корреляции $> 0,7$ между моделями) подтверждает их как надежные характеристики экспертного дискурса, а не как навязанные таксономии.

Однако вычислительный метод не позволяет определить, обусловлено ли согласие экспертов подлинным концептуальным согласием, общей дисциплинарной подготовкой или общими теоретическими влияниями, присутствующими во всех обучающих данных моделей. Эта неоднозначность является не просто методологическим ограничением, а существенным выводом о том, как научные сообщества разрабатывают и поддерживают концептуальные рамки посредством механизмов, которые сочетают индивидуальное рассуждение с коллективным дискурсом таким образом, что четкое разделение становится невозможным. Умеренная общая стабильность группировки (53,7%) по моделям, в сочетании с высокой стабильностью для конкретных определений (100% для Филдса и Хоффмана) и низкой стабильностью для других (22–28% для Солмса, Сломана, Дэвиса), показывает, что некоторые позиции занимают однозначные семантические позиции, в то время как другие действительно охватывают несколько концептуальных рамок.

Обсуждение

От античной философии до вычислительной топологии

Этот вычислительный анализ демонстрирует, что современные научные определения жизни группируются вокруг тех же философских разломов, которые структурировали древнегреческую мысль и продолжали формировать научный дискурс о жизни на

протяжении веков. Появление двух интерпретируемых измерений в проекции t-SNE, несмотря на математическую произвольность осей t-SNE, воспроизводит противоречия, которые организовывали мышление о жизни на протяжении тысячелетий.

Измерение 1 (x), сопоставляющее зависимость от наблюдателя концепции с объективно-материалистическими подходами, возрождает древнюю диалектику виталистского и механистического подходов. Изолированный кластер перцептивной категоризации (Филдс, Хоффман) занимает позицию, которая была бы философски невозможна до того, как квантовая механика ввела эффекты наблюдателя в фундаментальную физику, однако он перекликается с утверждением виталистского подхода о том, что жизнь нельзя свести к одному лишь механизму. Противоположный полюс, занимаемый информационным самовоспроизведением и самовоспроизводящимися термодинамическими системами, представляет собой высшую реализацию картезианско-гоббсовской программы: жизнь как материя в движении, полностью объяснимая физическими законами. Между этими полюсами находится область высокой плотности, где большинство современных ученых пытаются преодолеть это древнее противоречие с помощью гибридных концепций.

Измерение 2 (y) выявляет столь же настойчивый аристотелевский вопрос: является ли жизнь субстанцией или процессом? Верхние определения (например, «Живые клетки и все их конструкции» Балушки, акцент Пицци на «элементах на основе углерода», фокус Сломана на «компонентах») отражают субстанциалистскую онтологию, которая была бы узнаваема для досократовских атомистов: жизнь как особые виды материи со специфическими свойствами. Нижняя область (например, «огонь, который зажигает себя сам» Маршалла, «самосоздающая сила» Нобла, определение Пикара, ориентированное на сознание) возрождает аристотелевскую телеологию: жизнь как модели организации, самоактуализации и внутренней целесообразности, которые нельзя свести к их материальным субстратам. Этот вертикальный градиент отражает извечное противоречие между субстанциалистскими взглядами, рассматривающими жизнь как совокупность определенных сущностей (перекликаясь с картезианской и гоббсовской концепциями), и телеологическими взглядами, которые концептуализируют жизнь как целостные модели организации, потока информации и возникающей самоактуализации.

Центральный аттрактор, где пересекаются когнитивная автономия и диссипативные самоорганизующиеся системы, охватывая 66% всех определений, представляет собой попытку современной науки преодолеть эти дуализмы посредством интеграции. Определения в этой области учитывают как физические ограничения (термодинамика, поток энергии), так и эмергентные свойства (субъектность, познание, автономия), пытаясь преодолеть объяснительный разрыв между механизмом и опытом, который преследует философию со времен Декарта, разделившего разум и материю. Остается открытым вопрос, указывает ли эта концентрация на подлинный концептуальный синтез или лишь выявляет границы профессионально приемлемого дискурса в рамках институциональной науки.

Критически важно, что вычислительная методология выявляет то, что один лишь философский анализ не мог количественно оценить: это не дискретные позиции, а непрерывные градиенты. Ландшафт определений не имеет четких границ, только области более высокой и более низкой плотности. Определения занимают позиции, но эти позиции существуют в поле, структурированном историческими противоречиями. Характеристика жизни Уотсоном как «образца и процесса любви» функционирует как концептуальный

мост именно потому, что она одновременно пересекает множество измерений, связывая перцептивно-реляционные рамки с онтологиями процессов, сохраняя при этом элементы термодинамической необходимости. Такие связующие определения выявляют неадекватность рассмотрения подходов к определению как взаимоисключающих категорий.

Периферийные кластеры заслуживают особого внимания не как маргинальные позиции, а как потенциальные концептуальные авангарды. Кластер «Прагматический скептицизм в отношении определений» не просто отвергает стремление к универсальному определению, а скорее бросает вызов эпистемологическим предположениям, лежащим в основе всего этого предприятия. Эта рефлексивная позиция, скептически относящаяся к самому дефиниционизму, возникает из признания современной философией науки того, что определения являются инструментами, формируемыми прагматическими контекстами, а не открытиями естественного характера. Между тем, кластер «Перцептивная категоризация» представляет собой подлинно радикальный отход: позиционируя жизнь как перцептивный конструкт, а не как объективное свойство, Филдс и Хоффман ставят под сомнение реалистические предположения, имплицитно присутствующие в каждом определении от Аристотеля до Шрёдингера. Эти малонаселенные регионы могут указывать либо на позиции, для полного осмысления которых современная наука еще не разработала концептуальный словарь, либо на фундаментальные вызовы реалистической онтологии, которая структурирует основной научный дискурс.

Вычислительный подход также выявляет свои ограничения, что является методологически поучительным. Невозможность ручного присвоения определений действительно ортогональным категориям в сочетании со способностью алгоритмической кластеризации создавать интерпретируемую структуру предполагает, что экспертное мышление работает в многомерном семантическом пространстве, которое сопротивляется проекции на дискретные категории, но позволяет отображать их как непрерывную топологию. Визуализация t-SNE неизбежно уменьшает эту сложность, потенциально скрывая нюансы, существующие во всем 68-мерном корреляционном пространстве, однако согласованность, с которой интерпретируемые градиенты возникают в разных алгоритмах уменьшения размерности (MDS, UMAP, t-SNE; Дополнительный рис. [11](#)) и разных LLM, подтверждает, что эти закономерности являются подлинными особенностями экспертного дискурса, а не произвольными математическими артефактами.

Анализ наиболее ясно показывает, что кажущиеся разногласия в определении жизни представляют собой не концептуальную непоследовательность, а дифференцированные точки зрения в рамках единого семантического ландшафта. Исторический прогресс от греческой телеологии через механизм Просвещения до современной интеграции не разрешил, а скорее расширил концептуальное пространство. Каждая новая научная концепция, будь то термодинамика, теория информации, наука о сложности, искусственный интеллект или когнитивная наука, добавляет измерения к этому пространству, а не сжимает его в направлении консенсуса. Вопрос «что такое жизнь?» сохраняется не потому, что наука не смогла на него ответить, а потому, что сам вопрос отображает многомерную концептуальную территорию, которая сопротивляется сведению к одному определению. Наша вычислительная методология делает эту территорию доступной для навигации, выявляя структуру там, где предыдущие подходы видели только разногласия.

Модели конвергенции и дивергенции

Наш вычислительный анализ 68 экспертных определений показывает, что жизнь представляет собой непрерывный семантический ландшафт, а не дискретные категориальные состояния. Проекция t-SNE (рис. 4) демонстрирует эту непрерывность через два интерпретируемых измерения: измерение 1 (x) отображает градиент от зависимых от наблюдателя, реляционных моделей к объективным, материально-структурным подходам, а измерение 2 (y) показывает переход от ориентированных на процесс, телеологических перспектив к структурным моделям, основанным на сущностях.

Древняя дихотомия виталист-механист вновь проявляется в первом измерении, при этом смежные с виталистическим подходом перспективы группируются в левой части, ближе к перцептивно-реляционному полюсу, а механистические концепции — в правой части, ближе к материально-структурному полюсу. Аналогичным образом, историческое напряжение между онтологиями, основанными на субстанции и на процессе, напрямую отражается во втором измерении, показывая, что эти философские разногласия сохраняются в современном научном дискурсе.

Визуализация t-SNE дополнительно выявляет закономерности в распределении экспертных мнений по этому концептуальному пространству. Область с наибольшей плотностью сосредоточена на пересечении кластеров когнитивной автономии ($n = 24$) и диссипативных самоорганизующихся систем ($n = 21$), содержащем 66% всех определений. Эта концентрация свидетельствует о формирующемся консенсусе в отношении концептуальных моделей, которые интегрируют:

- Физические принципы (термодинамика вдали от равновесия, рассеяние энергии)
- Информационные процессы (поддержание границ, хранение и передача информации)
- Функциональные возможности (самоорганизация, взаимодействие с окружающей средой)
- Агентные свойства (целеустремленность, адаптивное поведение)

Эта зона конвергенции объединяет традиционно разрозненные дисциплинарные подходы, от физики и биологии до когнитивной науки и теории информации. Интегративный характер определений высокой плотности, примером которых служит «самомодифицирующаяся вычислительная фаза материи, возникающая в результате эволюционного отбора на динамическую стабильность» Агуэры-и-Аркаса, предполагает, что дух времени движется в сторону гибридных концепций, выходящих за рамки исторических дихотомий.

Периферийные области семантического ландшафта, хотя и менее плотно заселенные, могут указывать на зарождающиеся сдвиги парадигмы. Эти передовые позиции могут приобрести актуальность по мере того, как научный прогресс все чаще сталкивается с пограничными случаями между живыми и неживыми системами.

Семантический ландшафт предоставляет новую основу для позиционирования таких неоднозначных сущностей. Вирусы занимают переходные зоны между кластерами информационного самовоспроизведения и самовоспроизводящихся термодинамических систем, в то время как системы искусственного интеллекта могут прокладывать различные траектории в этом пространстве по мере развития все более сложных возможностей. Современные LLM могут занимать позиции между кластерами когнитивной автономии и информационного самовоспроизведения, демонстрируя обработку информации и целенаправленное поведение, но при этом не обладая материальной автономией и самоподдержанием. Такое позиционирование показывает, как семантический ландшафт может вмещать новые сущности, бросающие вызов традиционным категориальным границам. Будущие системы, включающие воплощенную робототехнику, могут мигрировать в область диссипативных самоорганизующихся систем по мере развития у них способностей к взаимодействию с окружающей средой и физическому самоподдержанию.

Методологические нововведения и ограничения

Следует отметить, что мы не утверждаем, что эти определения представляют собой статистически репрезентативное исследование какой-либо конкретной области. Вместо этого мы собрали мнения тщательно отобранной группы ученых для разработки методов анализа современных взглядов на эту тему и аналогичные сложные области. Мы заранее приносим извинения всем нашим сотрудникам или внешним экспертам, которых не попросили поделиться своим определением жизни для этого анализа, но наш код и рабочий процесс позволяют в будущих исследованиях достичь гораздо большей статистической мощности для более всестороннего изучения взглядов конкретных сообществ. Таким образом, мы призываем расширить дискуссию, включив в нее ученых из разных слоев общества и сообществ, включая разные культуры, возрастные группы и эпистемологические традиции.

Применение нами семантического анализа, основанного на модели LLM, к ландшафту определений демонстрирует новый подход к формированию консенсуса в различных дисциплинах. Вместо того чтобы стремиться к выявлению единственного «правильного» определения, этот подход отображает концептуальную территорию, в рамках которой функционируют различные определения, раскрывая как их взаимосвязи, так и их отличительный вклад. Этот метод предлагает модель для решения других спорных вопросов, связанных с определениями, в различных науках.

Этот подход отличается от традиционных методов агрегирования тем, что сохраняет отличительные черты конкурирующих концепций, одновременно выявляя лежащие в их основе взаимосвязи. Уникальный вклад методологии заключается в ее способности отображать концептуальные связи, не предполагая заранее, какие именно связи имеют значение. В отличие от традиционного философского анализа, который начинается с теоретических положений о соответствующих измерениях (например, механистика против витализма, редукционизм против холизма), этот основанный на методе LLM парный анализ позволяет структуре возникать из коллективной модели мышления экспертов. Этот подход, основанный на данных, дополняет, а не заменяет традиционный философский анализ, предлагая количественную основу, на которой может строиться интерпретационная работа.

Кроме того, в отличие от методов Дельфи, которые стремятся к консенсусу посредством итеративного уточнения, представленный вычислительный метаанализ сохраняет целостность различных точек зрения, одновременно предоставляя количественную карту их взаимосвязей. Эта методология объединяет редукционистские подходы, которые ищут необходимые и достаточные условия, с плюралистическими подходами, которые охватывают множество определений, тем самым выявляя единый семантический ландшафт вместо фрагментированных областей определений.

Показатели сходства, полученные на основе моделей LLM, заслуживают особого методологического внимания. Отражают ли эти показатели подлинное концептуальное соответствие или поверхностные лингвистические особенности? Философский ответ заключается в том, что эти различия могут быть менее различимыми, чем кажется на первый взгляд: если эксперты последовательно используют схожие концептуальные рамки, перекрывающуюся терминологию и параллельные аргументативные структуры, то в какой момент «поверхностное лингвистическое сходство» превращается в «глубокое концептуальное согласие»? Следуя Людвигу Витгенштейну (Витгенштейн [1958](#), с. 20), «значение слова — это его использование в языке», и модели использования, выявленные с помощью вычислительного анализа, могут быть наиболее близким к отображению лежащей в основе семантической структуры. Сходимость трех архитектурно различных моделей LLM ($r > 0,72$), обученных на разных данных, предполагает, что корреляционные модели отражают нечто более стабильное, чем индивидуальные смещения моделей. Тем не менее, мы не можем четко отделить концептуальную близость от влияния общих теоретических влияний, общих цитат и дисциплинарной подготовки, которые формируют как мышление экспертов, так и данные для обучения моделей. Эта неустраняемая двусмысленность — не недостаток, а преимущество: она показывает, как научное смыслообразование функционирует посредством механизмов, которые сочетают индивидуальное познание с коллективным дискурсом.

Следует отметить ряд ограничений. Во-первых, проанализированные определения экспертов, хотя и разнообразны, не могут претендовать на представление всех возможных взглядов на жизнь. 68 респондентов, которым было предложено дать определения, были отобраны авторами вручную и предназначены не для того, чтобы служить репрезентативной выборкой научного консенсуса, а скорее как отобранный набор междисциплинарных экспертов. Во-вторых, методология парной корреляции, хотя и строгая, неизбежно опирается на семантические возможности используемых линейных моделей обучения (ЛМН), что может вносить тонкие искажения в оценку концептуальных связей. В дополнение к ограничениям, связанным с компетенцией ЛМН, вычислительное время и стоимость генерации матриц парной корреляции масштабируются пропорционально квадрату числа выборок, что делает анализ больших когорт непомерно дорогим без достаточно гибких ограничений скорости или самодостаточных ЛМН. В-третьих, уменьшение размерности, необходимое для визуализации, неизбежно приводит к потере некоторых нюансов во взаимосвязях между определениями.

Несмотря на эти ограничения, предложенная методология демонстрирует потенциал вычислительного семантического анализа для выявления закономерностей концептуальной согласованности, которые в противном случае могли бы остаться скрытыми из-за фрагментации дисциплин. Такой подход предполагает, что кажущиеся несоответствия в определениях в некоторых случаях могут выявлять более глубокие закономерности согласованности, если рассматривать их как градиенты в семантическом пространстве.

Заключение

Наш вычислительный анализ 68 экспертных определений представляет жизнь как непрерывный семантический ландшафт, а не как дискретные категориальные состояния. Семантический анализ, основанный на модели LLM, выявил два возникающих измерения — зависимые от наблюдателя и объективные рамки, а также онтологии, основанные на процессах и на сущностях, — которые воспроизводят древние философские противоречия, одновременно предоставляя количественные инструменты для отображения современных экспертных точек зрения. Этот подход преобразует кажущиеся концептуальные разногласия в взаимодополняющие позиции в едином пространстве определений.

Данная методология сохраняет конкурирующие точки зрения, одновременно выявляя количественные взаимосвязи между ними, объединяя редукционистскую и плюралистическую парадигмы. Концентрация 66% определений в пересекающихся кластерах когнитивной автономии и диссипативных самоорганизующихся систем указывает на формирующийся консенсус в отношении концепций, интегрирующих физические принципы, информационные процессы и свойства субъектов. Эта конвергенция предполагает, что область исследований выходит за рамки исторических дихотомий в сторону гибридных подходов, преодолевающих дисциплинарные границы.

Разнообразие определений жизни свидетельствует не о концептуальном диссонансе, а о структурированной семантической топологии с предсказуемыми закономерностями конвергенции и дивергенции. Восемь выявленных здесь архетипических кластеров представляют собой стабильные концептуальные аттракторы, сохраняющиеся при различных вычислительных анализах, что указывает на то, что это устойчивые характеристики экспертного мышления, а не аналитические артефакты. Важно отметить, что появление интерпретируемых измерений без явного кодирования демонстрирует, что фундаментальные философские противоречия — зависимость от наблюдателя против объективности, процесс против содержания — продолжают систематически структурировать современный научный дискурс.

Эта методология естественным образом масштабируется на более крупные группы населения в смежных областях. Хотя наш анализ был сосредоточен на тщательно отобранном наборе экспертов, вычислительная структура позволяет проводить систематический анализ тысяч ответов из различных дисциплин. Такое масштабное исследование может выявить специфические для каждой дисциплины закономерности в более широком семантическом ландшафте, отслеживая, как специфическая для каждой области лексика и методологические подходы формируют методы определения понятий. Такое крупномасштабное картирование может выявить ранее нераспознанные связи между областями или определить новые концептуальные территории на границах дисциплин.

Помимо самой жизни, когнитивные процессы представляют собой следующий рубеж для этого аналитического подхода. Познание, как и жизнь, сопротивляется простому определению и охватывает множество уровней анализа, от нейронных механизмов до субъективного опыта. Память, с ее сложным взаимодействием молекулярных, клеточных и системных процессов, иллюстрирует, как фундаментальные понятия в биологии

преодолевают аналогичные проблемы определения. Применение нашей вычислительной семантической структуры к этим спорным областям может показать, представляют ли определенные модели определения общие особенности сложных биологических явлений или уникальные характеристики концептуализации жизни.

Вопрос «что такое жизнь?» остается актуальным, потому что он описывает многомерное концептуальное пространство, а не ожидает единственного ответа. Наша методология предлагает непосредственное применение к другим спорным определениям в различных науках, особенно в контексте синтетической биологии и искусственного интеллекта, бросающих вызов традиционным границам. Будущий прогресс в решении фундаментальных вопросов может быть достигнут не за счет установления жестких разграничений, а за счет понимания того, как точки зрения естественным образом сходятся и расходятся в рамках пространства определений.

Доступность данных

Код доступен в виде открытого репозитория на Github по адресу <https://github.com/mrbende/what-lives>.

Примечания

1. Греческое слово *Tou Pantos* означает «всего сущего» или «вселенной/космосом», гештальт всего творения.
-

Ссылки

- Аристотель (1931) «О душе», Книга II. (Смит Дж. А.) Кларендон, Оксфорд.
[Google Scholar](#)
- Баннерджи С. (2021) Возникающие правила вычислений во Вселенной приводят к жизни и сознанию: вычислительная модель сознания. *Interdiscip Descrip Complex Syst* 19:31–41
[Статья Google Scholar](#)
- Беннер С.А. (2010) Определение жизни. *Астробиол* 10:1021–1030
[Статья Google Scholar](#)
- Блэкстон Д., Ледерер Э., Кригман С., Гарнье С., Бонгард Дж., Левин М. (2021) Клеточная платформа для разработки синтетических живых машин. *Sci Robot* 6.
<https://www.ncbi.nlm.nih.gov/pubmed/34043553>
- Боден М.А. (2003) Инопланетная жизнь: как нам это узнать? *Int J Astrobiol* 2:121–129
[Статья Google Scholar](#)
- Больцман Л. (1886) Второй закон термодинамики. Теоретическая физика и философские проблемы. В: МакГиннесс Б. (ред.) Теоретическая физика и философские проблемы. Коллекция Венского кружка. Springer, Дордрехт. https://doi.org/10.1007/978-94-010-2091-6_2

- Кэмпбелл Д.Р. (2021) Самодвижение и познание: теория души Платона. *South J Philos* 59:523–544
[Статья Google Scholar](#)
- Чанг К.М. (2011) Алхимия как исследование жизни и материи: переосмысление места витализма в химии раннего Нового времени. *Isis* 102:322–329
[Статья Google Scholar](#)
- Клеланд С., Чибба С. (2010) Есть ли определение у понятия «жизнь»? В: Бедау М.А., Клеланд К.Е. (ред.) *Природа жизни: классические и современные взгляды с точки зрения философии и науки*. Издательство Кембриджского университета, Кембридж, стр. 326–339
[Глава Google Scholar](#)
- Крик Ф.Х. (1958) О синтезе белка. Симпозиум Общества экспериментальной биологии 12:138–163
[Google Scholar](#)
- Дарвин С. (1859) О происхождении видов путем естественного отбора. Джон Мюррей, Лондон.
[Google Scholar](#)
- Декарт Р. ([1664] (1909) *Oeuvres de Descartes: Tome XI Le monde, Описание человеческого тела, Passions de l'âme*. Anatomica, Varia. L. Cerf, Париж
[Google Scholar](#)
- Диас-Родригес Дж. (2025) K-LLMmeans: сводки как центроиды для интерпретируемой и масштабируемой кластеризации текста на основе LLM, *arXiv*
- England J (2020) *Every life is on fire: how thermodynamics explains the origins of living things*. Basic Books, New York City
[Google Scholar](#)
- Фэггин Ф. (2024) *Несводимое: сознание, жизнь, компьютеры и человеческая природа*. Essentia Books
- Гален (1968) О пользе частей тела. (перевод: Мэй М.Т.) Издательство Корнельского университета, Итака, Нью-Йорк.
- Ганти Т. (2003а) *Теория хемотонов: теория живых систем*. Springer, Дордрехт.
[Книга Google Scholar](#)
- Ганти Т. (2003b) *Принципы жизни*. Издательство Оксфордского университета, Оксфорд.
[Книга Google Scholar](#)
- Гайон Дж., Малатер С., Моранж М., Рален-Серсо Ф., Тирап С. (2010) Определяя жизнь: материалы конференции. *Orig Life Evol Biosph* 40:119–120
[Статья Google Scholar](#)
- Гилберт С.Ф. (1982) Интеллектуальные традиции в биологических науках: молекулярная биология и биохимия. *Perspect Biol Med* 26:151–162
[Статья Google Scholar](#)
- Гумуская Г., Шривастава П., Купер Б.Г., Лессер Х., Семегран Б., Гарнье С., Левин М. (2023) Подвижные живые биороботы самоорганизуются из взрослых соматических клеток-предшественников человека. *Adv Sci (Weinh)*, стр. e2303575.
<https://www.ncbi.nlm.nih.gov/pubmed/38032125>
- Гумуская Г., Дейви Н., Шривастава П., Бендер А., Пио-Лопес Л., Хейзел Д., Левин М. (2025) Морфологический, поведенческий и транскриптомный жизненный цикл антропоботов. *Adv Sci (Weinh)*, стр. e2409330.
<https://www.ncbi.nlm.nih.gov/pubmed/40479594>
- Гоббс Т. (1909) «Левиафан Гоббса»: переиздание издания 1651 года. Clarendon Press, Оксфорд.
[Google Scholar](#)
- Холланд Дж. Х. (1975) *Адаптация в природных и искусственных системах: вводный анализ с приложениями к биологии, управлению и искусственному интеллекту*. Издательство Мичиганского университета, Анн-Арбор, Мичиган.

[Google Scholar](#)

- Хуан Ч., Хэ Г. (2024) Кластеризация текста как классификация с использованием LLM, р. [parXiv](#)
- Хатчисон К.А. 3-й, Чуан Р.Й., Носков В.Н., Ассад-Гарсия Н., Диринк Т.Дж., Эллисман М.Х., Гилл Дж., Каннан К., Карас Б.Дж., Ма Л., Пеллетье Дж.Ф., Ци З.Ц., Рихтер Р.А., Стрыхальский Э.А., Сун Л., Сузуки Ю., Цветанова Б., Уайз К.С., Смит Х.О., Гласс Дж.И., Мэрриман К., Гибсон Д.Г., Вентер Дж.К. (2016) Разработка и синтез минимального бактериального генома. *Sci* 351:aad6253

[Статья Google Scholar](#)

- Идрис С., Манукин М.Б., Рике Ф., Филд Г.Д., Зильберберг Дж. (2024) Биофизические механизмы нейронной адаптации позволяют искусственным нейронным сетям улавливать динамические вычисления на сетчатке. *Nat Commun* 15:5957

[Статья Google Scholar](#)

- Джонсон М.Р. (2005) Аристотель о телеологии. Издательство Оксфордского университета, Оксфорд.

[Книга Google Scholar](#)

- Керагель И., Морбье С., Надиф М. (2024) За пределами слов: сравнительный анализ LLM-встраиваний для эффективной кластеризации. В: Милиу И., Пятковски Н., Папапетру П. (ред.) Достижения в интеллектуальном анализе данных XXII. Springer, Cham, стр. 205–216

[Глава Google Scholar](#)

- Кобак Д., Беренс П. (2019) Искусство использования t-SNE для транскриптомики отдельных клеток. *Nat Commun* 10:5416

[Статья Google Scholar](#)

- Ламетри ЖО (1865) Человек-машина. Ж. Ассеза, Париж

[Google Scholar](#)

- Лэнгтон К.Г. (1997) Искусственная жизнь: обзор. Издательство MIT Press, Кембридж.

[Google Scholar](#)

- Левин М. (2022) Технологический подход к разуму повсюду: экспериментально обоснованная концепция для понимания разнообразия тел и умов. *Front Syst Neurosci* 16:768201

[Статья Google Scholar](#)

- Лиузи А., Манакул П., Гейлс М.Дж.Ф. (2023) Сравнительная оценка больших языковых моделей: оценка NLG без предварительного обучения посредством попарных сравнений с использованием больших языковых моделей. Конференция Европейского отделения Ассоциации вычислительной лингвистики.
- Мачери Э. (2012) Почему я перестал беспокоиться об определении жизни... и почему вам тоже следует это сделать. *Synthese* 185:145–164

[Статья Google Scholar](#)

- Маклем П.Т., Сили А. (2010) К определению жизни. *Perspect Biol Med* 53:330–340

[Статья Google Scholar](#)

- Малатерр С., Шартье Ж.Ф. (2021) За пределами категориальных определений жизни: подход к оценке жизненных качеств, основанный на данных. *Synthese* 198:4543–4572

[Статья Google Scholar](#)

- Маргулис Л., Саган Д. (2000) Что такое жизнь? Издательство Калифорнийского университета, Беркли.

[Google Scholar](#)

- Марискаль С., Дулиттл У.Ф. (2020) Жизнь и только жизнь: радикальная альтернатива определению жизни. *Synthese* 197:2975–2989

[Статья Google Scholar](#)

- Марискаль С. (2023) Жизнь. В: Залта, Э.Н. (ред.), Стэнфордская энциклопедия философии. Лаборатория метафизических исследований, Стэнфордский университет. (Дата обращения: 15 декабря 2025 г.)

- Маттиас Дж (1838) Beiträge zur Phytogenic. Архив анатомии, физиологии и медицины. Файт, Берлин
[Google Scholar](#)
- Матурана Х., Варела Ф. (1980) Автопоэзис и познание: реализация живого. Издательство D. Reidel, Дордрехт, Бостон, Лондон.
[Книга](#) [Google Scholar](#)
- Морено А., Моссио М. (2015) Биологическая автономия: философское и теоретическое исследование. Springer, Дордрехт.
[Книга](#) [Google Scholar](#)
- Морено А., Перето Дж. (2026) Эволюционная история самостоятельности: как жизнь эволюционировала, чтобы действовать самостоятельно. Springer, Cham
[Книга](#) [Google Scholar](#)
- Николсон Д.Дж. (2025) Что такое жизнь? Издательство Кембриджского университета, Кембридж, пересмотренное издание.
[Google Scholar](#)
- Noble R, Tasaki K, Noble PJ, Noble D (2019) Биологическая относительность требует круговой причинности, но не симметрии причинности: так где же, что и когда находятся границы? Front Physiol 10:827
[Статья](#) [Google Scholar](#)
- Пай В., Пио-Лопес Л., Сперри М., Эрикссон П., Тайеби П., Левин М. (2025) Базальная транскриптомика ксеноботиков: анализ изменений экспрессии генов в клетках дикого типа, освобожденных от влияния остальной части организма, выявляет новый способ контроля. Commun Biol 8:646
[Статья](#) [Google Scholar](#)
- Парке Э. (2021) Характеристика жизни: четыре измерения и их значение для исследований происхождения жизни. Препринт. <https://philsci-archive.pitt.edu/18776/>
- Паршалл К.Х., Уолтон М.Т., Моран Б.Т. (2015) Преодоление традиций. Издательство Университета штата Пенсильвания, Университи-Парк, Пенсильвания.
[Книга](#) [Google Scholar](#)
- Попа Р. (2004) Между необходимостью и вероятностью: поиск определения и происхождения жизни. Springer, Берлин, Гейдельберг.
[Google Scholar](#)
- Пригожин И. (1980) От бытия к становлению: время и сложность в физических науках. У. Х. Фриман, Сан-Франциско.
[Google Scholar](#)
- Шрёдингер Э. (1944) Что такое жизнь? Физический аспект живой клетки. Издательство Кембриджского университета, Кембридж.
[Google Scholar](#)
- Сеоане Л.Ф., Соле Р.В. (2018) Теория информации, предсказуемость и возникновение сложной жизни. R Soc Open Sci 5: 172221. <https://www.ncbi.nlm.nih.gov/pubmed/29515907>
- Соуза Т., Домингос Т., Койман С.А. (2008) От эмпирических закономерностей к теории: формальная метаболическая теория жизни. Philos Trans R Soc Lond B Biol Sci 363:2453–2464
[Статья](#) [Google Scholar](#)
- Спиноза Б. (2020) Этика Спинозы. (перевод: Элиот Г.). В: Архив Джорджа Элиота, под редакцией Беверли Парк Райлетт, <https://GeorgeEliotArchive.org> (дата обращения: 15 декабря 2025 г.)
- Стоунс Г.Б. (1928) Атомный взгляд на материю в XV, XVI и XVII веках. Издательство Чикагского университета.
[Книга](#) [Google Scholar](#)
- Страттон С.Дж. (2024) Целенаправленная выборка: преимущества и недостатки. Prehosp Disaster Med 39:121–122

[Статья Google Scholar](#)

- Фон Нейман Дж. (1966) Теория самовоспроизводящихся автоматов. Издательство Иллинойского университета, Шампейн.

[Google Scholar](#)

- Ван С., Чжао В., Тан Ц.Н., Дай З.Б., Фэн Ю.Н. (2025) Эволюционный алгоритм с адаптивным генетическим оператором и динамическим механизмом оценки для крупномасштабной разреженной многоцелевой оптимизации. Sci Rep 15:9267

[Статья Google Scholar](#)

- Уотсон Дж. Д., Крик Ф. Х. (1953) Молекулярная структура нуклеиновых кислот; структура дезоксирибонуклеиновой кислоты. Nat 171:737–738

[Статья Google Scholar](#)

- Уитакер Р.Х. (1941) Интерпретация диалога Платона «Тимей» А.Е. Тейлором и Ф.М. Корнфордом, Философия. Бостонский университет, Бостон.
- Витгенштейн Л. (1958) Философские исследования. (перевод: Анскомб Г.) Издательство Macmillan Publishing Co, Нью-Йорк.
https://archive.org/details/philosophicalinvestigations_201911 (Дата обращения: 15 декабря 2025 г.)
- Вольфрам С. (2002) Новый вид науки. Репозиторий данных Вольфрама.
<https://doi.org/10.24097/wolfram.15062.data>
- Чжан Ю., Ван Ч., Шан Дж. (2023) ClusterLLM: большие языковые модели как руководство для кластеризации текста, arXiv
- Гилберт С.Ф. (ред.) (1991) Концептуальная история современной эмбриологии. Plenum, Нью-Йорк.

[Скачать ссылки](#)

Благодарности

Крис Филдс хотел бы выразить признательность Эрику Дитриху за вдохновение, которое он послужил источником для своего определения. Мы благодарим всех участников за их продуманные определения и всех ученых, которые внесли свой вклад в наше исследование, предоставив свои определения. Мы также благодарим Джулию Пуарье за помощь в подготовке рукописи.

Финансирование

Майкл Левин выражает благодарность фонду Elisabeth Giaque Trust (Лондон) за оказанную поддержку, а также гранту № 62212 от фонда John Templeton Foundation.

Информация об авторе

Авторы

- [Рид Бендер](#), [ORCID:orcid.org/0009-0008-4371-6662](https://orcid.org/0009-0008-4371-6662)¹

- [Карина Кофман, ОРЦИД:orcid.org/0000-0002-6065-5207²](https://orcid.org/0000-0002-6065-5207)
- [Blaise Agüera y Arcas & ОРЦИД:orcid.org/0000-0003-2256-9823³](https://orcid.org/0000-0003-2256-9823)
- [Майкл Левин ОРЦИД:orcid.org/0000-0001-7292-8084^{4, 5}](https://orcid.org/0000-0001-7292-8084)

Авторы и принадлежность к организациям

1. **Attune Intelligence, LLC, Даллас, Техас, США**
Рид Бендер
2. **Факультет стоматологии, Университет Торонто, Торонто, Канада**
Карина Кофман
3. **Группа «Парадигмы интеллекта», Google, Сигл, штат Вашингтон, США.**
Blaise Agüera y Arcas
4. **Научно-познавательный центр имени Аллена при Университете Тафтса, 200 Boston Ave, Suite 4600, Medford, MA, 02155, США**
Майкл Левин
5. **Институт биоинженерных исследований им. Висса при Гарвардском университете, Бостон, Массачусетс, США**
Майкл Левин

Авторы

1. Рид Бендер

[Просмотреть публикации автора](#)

Найти автора по: [PubMedGoogle Scholar](#)

2. Карина Кофман

[Просмотреть публикации автора](#)

Найти автора по: [PubMedGoogle Scholar](#)

3. Blaise Agüera y Arcas

[Просмотреть публикации автора](#)

Найти автора по: [PubMedGoogle Scholar](#)

4. Майкл Левин

[Просмотреть публикации автора](#)

Найти автора по: [PubMedGoogle Scholar](#)

Взносы

ML — разработка концепции, сбор и интерпретация данных. RB — вычислительный анализ данных, исторический контекст и создание иллюстраций. КК — обработка данных и создание иллюстраций. ВА-А — вычислительный анализ данных. Все авторы внесли свой вклад в написание и редактирование статьи.

Автор, ответственный за переписку.

Адрес для корреспонденции: [Майкл Левин](#).

Заявления об этике

Конфликт интересов

Никто.

Дополнительная информация

Примечание издателя

Издательство Springer Nature сохраняет нейтральную позицию в отношении территориальных претензий, указанных на опубликованных картах, и принадлежности к учреждениям.

Дополнительная информация

Ниже приведена ссылка на электронные дополнительные материалы.

[Дополнительный файл 1 \(скачать PDF\)](#)

[Дополнительный файл 2 \(скачать PDF\)](#)

[Дополнительный файл 3 \(скачать PDF\)](#)

[Дополнительный файл 4 \(скачать PDF\)](#)

[Дополнительный файл 5 \(скачать DOCX\)](#)

Права и разрешения

Открытый доступ. Данная статья распространяется под лицензией Creative Commons Attribution 4.0 International, которая разрешает использование, распространение, адаптацию, передачу и воспроизведение в любом формате и на любом носителе при условии указания авторства и источника, предоставления ссылки на лицензию Creative Commons и указания на внесенные изменения. Изображения или другие материалы третьих лиц, включенные в данную статью, подпадают под действие лицензии Creative

Commons, если иное не указано в примечании к материалу. Если материал не подпадает под действие лицензии Creative Commons, и предполагаемое вами использование не разрешено законодательством или превышает допустимое использование, вам необходимо получить разрешение непосредственно от правообладателя. Чтобы ознакомиться с текстом данной лицензии, посетите <http://creativecommons.org/licenses/by/4.0/>.

[Перепечатки и разрешения](#)

Об этой статье

Ссылка на эту статью

Bender, R., Kofman, K., Agüera y Arcas, B. *et al.* What Lives? A Meta-Analysis of Diverse Opinions on the Definition of Life. *Biol Theory* (2026). Бендер, Р., Кофман, К., Агуэра и Аркас, Б. *и др.* Что такое жизнь? Мета-анализ различных мнений об определении жизни. *Biol Theory* (2026). <https://doi.org/10.1007/s13752-026-00544-9>

- Полученный 16 октября 2025 г. Принят 11 апреля 2026 г. Опубликовано 02 июля 2026 г. Официальная версия 02 июля 2026 г. <https://doi.org/10.1007/s13752-026-00544-9>

Keywords

- [Agency](#); [Artificial intelligence](#); [Biology](#); [Computation](#); [Cybernetics](#); [Evolution](#); [Information](#); [Intelligence](#); [Life](#); [Thermodynamics](#)

Ключевые слова

[Агентство](#); [Искусственный интеллект](#); [Биология](#); [Вычисления](#); [Кибернетика](#); [Эволюция](#); [Информация](#); [Интеллект](#); [Жизнь](#); [Термодинамика](#)

Профили

1. Майкл Левин [Просмотреть профиль автора](#)

Refs#

1. Аристотель (1931) «О душе», Книга II. (Смит Дж. А.) Кларендон, Оксфорд.
2. Баннерджи С. (2021) Возникающие правила вычислений во Вселенной приводят к жизни и сознанию: вычислительная модель сознания. *Interdiscip Descr Complex Syst* 19:31–41
3. Беннер С.А. (2010) Определение жизни. *Астробиол* 10:1021–1030

4. Блэкистон Д., Ледерер Э., Кригман С., Гарнье С., Бонгард Дж., Левин М. (2021) Клеточная платформа для разработки синтетических живых машин. *Sci Robot* 6. <https://www.ncbi.nlm.nih.gov/pubmed/34043553>
5. Боден М.А. (2003) Инопланетная жизнь: как нам это узнать? *Int J Astrobiol* 2:121–129
6. Больцман Л. (1886) Второй закон термодинамики. Теоретическая физика и философские проблемы. В: МакГиннесс Б. (ред.) Теоретическая физика и философские проблемы. Коллекция Венского кружка. Springer, Дордрехт. https://doi.org/10.1007/978-94-010-2091-6_2
7. Кэмпбелл Д.Р. (2021) Самодвижение и познание: теория души Платона. *South J Philos* 59:523–544
8. Chang KM (2011) Alchemy as studies of life and matter: reconsidering the place of vitalism in early modern chemistry. *Isis* 102:322–329
9. Cleland C, Chyba C (2010) Does ‘life’ have a definition? In: Bedau MA, Cleland CE (eds) *The nature of life: classical and contemporary perspectives from philosophy and science*. Cambridge University Press, Cambridge, pp 326–339
10. Crick FH (1958) On protein synthesis. *Symp Soc Exp Biol* 12:138–163
11. Darwin C (1859) *On the origin of species by means of natural selection*. John Murray, London
12. Descartes R, ([1664] (1909) *Oeuvres de Descartes: Tome XI Le monde, Description du corps humain, Passions de l’âme*. Anatomica, Varia. L. Cerf, Paris
13. Diaz-Rodriguez J (2025) K-LLMmeans: summaries as centroids for interpretable and scalable LLM-based text clustering, p arXiv
14. England J (2020) *Every life is on fire: how thermodynamics explains the origins of living things*. Basic Books, New York City
15. Faggin F (2024) *Irreducible: consciousness, life, computers, and human nature*. Essentia Books
16. Galen (1968) *On the usefulness of parts of the body*. (trans: May MT) Cornell University Press, Ithaca, New York
17. Gánti T (2003a) *Chemoton theory: theory of living systems*. Springer, Dordrecht
18. Gánti T (2003b) *The principles of life*. Oxford University Press, Oxford
19. Gayon J, Malaterre C, Morange M, Raulin-Cerceau F, Tirard S (2010) Defining life: conference proceedings. *Orig Life Evol Biosph* 40:119–120
20. Gilbert SF (1982) Intellectual traditions in the life sciences: molecular biology and biochemistry. *Perspect Biol Med* 26:151–162
21. Gumuskaya G, Srivastava P, Cooper BG, Lesser H, Semegran B, Garnier S, Levin M (2023) Motile living biobots self-construct from adult human somatic progenitor seed cells. *Adv Sci (Weinh)*, p e2303575. <https://www.ncbi.nlm.nih.gov/pubmed/38032125>
22. Gumuskaya G, Davey N, Srivastava P, Bender A, Pio-Lopez L, Hazel D, Levin M (2025) The morphological, behavioral, and transcriptomic life cycle of anthrobots. *Adv Sci (Weinh)*, p e2409330. <https://www.ncbi.nlm.nih.gov/pubmed/40479594>
23. Hobbes T (1909) *Hobbes’s Leviathan*: reprinted from the edition of 1651. Clarendon Press, Oxford
24. Holland JH (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, Ann Arbor, MI
25. Huang C, He G (2024) Text clustering as classification with LLMs, p arXiv
26. Hutchison CA 3rd, Chuang RY, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, Gill J, Kannan K, Karas BJ, Ma L, Pelletier JF, Qi ZQ, Richter RA, Strychalski EA, Sun L, Suzuki Y, Tsvetanova B, Wise KS, Smith HO, Glass JI, Merryman C, Gibson

- DG, Venter JC (2016) Design and synthesis of a minimal bacterial genome. *Sci* 351:aad6253
27. Idrees S, Manookin MB, Rieke F, Field GD, Zylberberg J (2024) Biophysical neural adaptation mechanisms enable artificial neural networks to capture dynamic retinal computation. *Nat Commun* 15:5957
 28. Johnson MR (2005) *Aristotle on teleology*. Oxford University Press, Oxford
 29. Keraghel I, Morbieu S, Nadif M (2024) Beyond words: a comparative analysis of LLM embeddings for effective clustering. In: Miliou I, Piatkowski N, Papapetrou P (eds) *Advances in intelligent data analysis XXII*. Springer, Cham, pp 205–216
 30. Kobak D, Berens P (2019) The art of using t-SNE for single-cell transcriptomics. *Nat Commun* 10:5416
 31. La Mettrie JO (1865) *L’homme machine*. J. Assézat, Paris
 32. Langton CG (1997) *Artificial life: an overview*. MIT Press, Cambridge
 33. Levin M (2022) Technological approach to mind everywhere: an experimentally-grounded framework for understanding diverse bodies and minds. *Front Syst Neurosci* 16:768201
 34. Liusie A, Manakul P, Gales MJF (2023) LLM comparative assessment: zero-shot NLG evaluation through pairwise comparisons using large language models. *Conference of the European Chapter of the Association for Computational Linguistics*
 35. Machery E (2012) Why I stopped worrying about the definition of life... and why you should as well. *Synthese* 185:145–164
 36. Macklem PT, Seely A (2010) Towards a definition of life. *Perspect Biol Med* 53:330–340
 37. Malaterre C, Chartier J-F (2021) Beyond categorical definitions of life: a data-driven approach to assessing liveness. *Synthese* 198:4543–4572
 38. Margulis L, Sagan D (2000) *What is life?* University of California Press, Berkeley
 39. Mariscal C, Doolittle WF (2020) Life and life only: a radical alternative to life definitionism. *Synthese* 197:2975–2989
 40. Mariscal C (2023) Life. In: Zalta, E.N. (ed), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. (Accessed 15 December 2025)
 41. Matthias J (1838) *Beiträge zur Phytogenesis*. Archiv für Anatomie, Physiologie und wissenschaftliche Medicin. Veit, Berlin
 42. Maturana H, Varela F (1980) *Autopoiesis and cognition: the realization of the living*. D. Reidel Publishing Company, Dordrecht, Boston, London
 43. Moreno A, Mossio M (2015) *Biological autonomy: a philosophical and theoretical enquiry*. Springer, Dordrecht
 44. Moreno A, Peretó J (2026) *An evolutionary story of agency: how life evolved to act on its own*. Springer, Cham
 45. Nicholson DJ (2025) *What is life?* Cambridge University Press, Cambridge, Revisited
 46. Noble R, Tasaki K, Noble PJ, Noble D (2019) Biological relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? *Front Physiol* 10:827
 47. Pai V, Pio-Lopez L, Sperry M, Erickson P, Tayyebi P, Levin M (2025) Basal xenobot transcriptomics: analysis of gene expression changes in wild-type cells freed from the influence of the rest of the organism reveals novel control modality. *Commun Biol* 8:646
 48. Parke E (2021) Characterizing life: four dimensions and their relevance to origin of life research. Preprint. <https://philsci-archive.pitt.edu/18776/>
 49. Parshall KH, Walton MT, Moran BT (2015) *Bridging traditions*. Penn State University Press, University Park, PA
 50. Popa R (2004) *Between necessity and probability: searching for the definition and origin of life*. Springer, Berlin Heidelberg

-
51. Prigogine I (1980) From being to becoming: time and complexity in the physical sciences. W. H. Freeman, San Francisco
 52. Schrödinger E (1944) What is life? The physical aspect of the living cell. Cambridge University Press, Cambridge
 53. Seoane LF, Sole RV (2018) Information theory, predictability and the emergence of complex life. *R Soc Open Sci* 5: 172221.
<https://www.ncbi.nlm.nih.gov/pubmed/29515907>
 54. Sousa T, Domingos T, Kooijman SA (2008) From empirical patterns to theory: a formal metabolic theory of life. *Philos Trans R Soc Lond B Biol Sci* 363:2453–2464
 55. Spinoza B (2020) Spinoza’s ethics. (trans: Eliot G). In: George Eliot Archiv, edited by Beverley Park Rilett, <https://GeorgeEliotArchive.org> (Accessed 15 December 2025)
 56. Stones GB (1928) The atomic view of matter in the XVth, XVIth, and XVIIth centuries. The University of Chicago Press
 57. Stratton SJ (2024) Purposeful sampling: advantages and pitfalls. *Prehosp Disaster Med* 39:121–122
 58. Von Neumann J (1966) Theory of self-reproducing automata. University of Illinois Press, Champaign
 59. Wang X, Zhao W, Tang JN, Dai ZB, Feng YN (2025) Evolution algorithm with adaptive genetic operator and dynamic scoring mechanism for large-scale sparse many-objective optimization. *Sci Rep* 15:9267
 60. Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nat* 171:737–738
 61. Whitaker RH (1941) The interpretation of Plato’s *Timaeus* by a. E. Taylor and f. M. Cornford, Philosophy. Boston University, Boston
 62. Wittgenstein L (1958) Philosophical investigations. (trans: Anscombe G) Macmillan Publishing Co, New York. https://archive.org/details/philosophicalinvestigations_201911 (Accessed 15 December 2025)
 63. Wolfram S (2002) A new kind of science. Wolfram Data Repository.
<https://doi.org/10.24097/wolfram.15062.data>
 64. Zhang Y, Wang Z, Shang J (2023) ClusterLLM: large language models as a guide for text clustering, p arXiv
 65. Gilbert SF (ed) (1991) A conceptual history of modernembryology. Plenum, New York
-