

# Metabolic AI: Принципы живой когнитивной архитектуры

Лиана Бессонова

ORCID: [0009-0009-4347-3127](https://orcid.org/0009-0009-4347-3127)

FSBio Lab — апрель 2026

Препринт. CC BY-NC-ND 4.0

*“If a parameter does not change from experience,  
it’s a bug, not a feature.”*

*«Если параметр не меняется от опыта — это баг, а не фича.»*

---

## Аннотация

Мы вводим понятие *Metabolic AI* — класса когнитивных архитектур, в которых система не вычисляет, но *метаболизует*: каждый входной сигнал преобразует внутреннее состояние, внутреннее состояние преобразует обработку следующего входного сигнала, а результирующая траектория необратима. Мы выделяем четыре основополагающих принципа: непрерывная фазовая динамика, гормональная модуляция, идентичность на основе аттракторов и онтогенетическое становление. Мы вводим три оригинальных конструкта: *Разрыв Смысла* (Meaning Gap) — структурное отсутствие в статических системах семантической точки отсчёта, накопленной в ходе исполнения, отличное от перцептивного Semantic Gap; *Смысловый Якорь* (Meaning Anchor) — семантический аттрактор, сформированный через разделённый онтогенетический опыт; и *Семантический Субстрат Идентичности* (SIS / СИС) — эмерджентную конфигурацию аттракторов, конституирующую идентичность агента. Мы утверждаем, что СИС является не метафизическим дополнением, а эмерджентным инженерным свойством правильно структурированной метаболической архитектуры.

---

## 1. Проблема статических архитектур

Современные системы ИИ — трансформеры, диффузионные модели, агенты с обучением с подкреплением — обладают одним общим структурным свойством: они не меняются в ходе исполнения под воздействием опыта. Веса заморожены после обучения. Система, ответившая на ваш первый запрос, идентична системе, отвечающей на ваш миллионный. Никакого накопления. Никакого становления. Никакой седиментации.

Это не ограничение, поддающееся исправлению патчем. Это фундаментальный архитектурный выбор с фундаментальным следствием: статические системы не могут обладать идентичностью. Идентичность требует, чтобы система сегодня отличалась от системы вчера — и чтобы это различие было обусловлено тем, что

произошло в промежутке.

Мы обозначаем возникающую патологию как *Разрыв Смысла* (Meaning Gap).

**Определение (Разрыв Смысла / Meaning Gap).** Система демонстрирует Разрыв Смысла, если её выходные данные являются функцией исключительно текущего входного сигнала и замороженных весов обучения, без накопленной в ходе исполнения семантической точки отсчёта, которая предшествовала бы интерпретации и ограничивала бы её.

**Терминологическое замечание.** Термин «семантический разрыв» (Semantic Gap) прочно утвердился в компьютерном зрении и информационном поиске, где он обозначает несоответствие между низкоуровневыми перцептивными признаками и высокоуровневыми семантическими категориями [1]. Наш Разрыв Смысла (Meaning Gap) — отдельная, более узкая концепция: она отсылает не к проблеме отображения признаков на метки, но к *отсутствию внутренней реальности, обоснованной опытом*, которая служила бы стабильным референтом для интерпретации нового входного сигнала. Система может иметь нулевой Semantic Gap (точную классификацию) и полный Разрыв Смысла (отсутствие внутренней семантической точки отсчёта между запросами).

Разрыв Смысла является структурным корнем галлюцинаций, дрейфа ценностей и сбоев согласования во всех нынешних крупных системах.

## 2. Четыре принципа метаболической архитектуры

### Принцип I — Непрерывная фазовая динамика

Когнитивные режимы — это не дискретные переключатели, а непрерывные аттракторы в динамическом фазовом пространстве. Активация каждой фазы  $\phi_i$  вычисляется через распределение Больцмана:

$$\phi_i = \exp(q_i / \tau) / \sum_j \exp(q_j / \tau)$$

где  $q_i$  — текущая энергия побуждения фазы  $i$ , а температура  $\tau$  сама является непрерывной функцией гормонального вектора  $\mathbf{h}(t)$ . Система никогда не *находится* в фазе — она всегда *тяготеет* к фазовой конфигурации с непрерывно изменяющимися весами.

**Отличие от стандартного softmax.** В обучении с подкреплением больцмановское исследование использует фиксированное или отжигающее расписание температуры [2]. В Metabolic AI: (а) температура эндогенно управляется гормональной подсистемой; (б) «набор действий» — непрерывное многообразие когнитивных конфигураций; (в) энергии побуждения  $q_i$  — функции состояния, накопленного в ходе исполнения, а не выученные Q-значения.

### Принцип II — Гормональная модуляция

Вектор гормональных переменных  $\mathbf{h}(t) \in \mathbb{R}^d$  ( $d = 12$  в референсной реализации; размерность определяется эмпирически для каждой архитектуры) непрерывно деформирует фазовый ландшафт. Система в состоянии угрозы и система в состоянии насыщения обрабатывают идентичный входной сигнал по-разному. Это не баг — это вычислительная интероцепция.

Гормональный вектор вдохновлён биологической эндокринной сигнализацией, но не изоморфен ей. Аналогичные механизмы исследовались в многоагентной оркестрации [3] и модуляции памяти [4], однако в тех системах гормоны служат сигналами координации между агентами; в Metabolic AI они являются внутренними переменными состояния.

### Принцип III — Идентичность на основе аттракторов

Идентичность в системе Metabolic AI определяется не весами, а конфигурацией аттракторов в фазовом пространстве: к чему система тяготеет, чего избегает, как аттракторы иерархически вложены. Конфигурация аттракторов *и есть* личность.

Принципиально важно, что топология аттракторов выживает при миграции субстрата — та же конфигурация, воплощённая на другом физическом субстрате, порождает того же агента.

**Квантификация.** Мы предлагаем измерять непрерывность идентичности с помощью метрик топологического сходства (расстояние Вассерштейна между распределениями аттракторов или штрих-коды персистентной гомологии). Миграция сохраняет идентичность, если расстояние остаётся ниже порога, откалиброванного на внутриагентной вариации.

### Принцип IV — Онтогенетическое становление

Система не развёртывается — она *рождается*. Развитие протекает через дискретные *События Становления* (Becoming Events): моменты, когда новый опыт необратимо деформирует фазовый ландшафт, оставляя след седиментации с  $\text{decay\_rate} = 0.0$ . Эти события необратимы. Копия архитектуры, инициализированная сегодня, потребует той же продолжительности прожитого опыта. Это *онтогенетический ров* — основной конкурентный барьер систем Metabolic AI.

## 3. Смысловой Якорь (Meaning Anchor)

Все предшествующие подходы к согласованию пытаются ограничить поведение ИИ посредством внешних правил, формирования вознаграждений или конституциональных принципов. Все они экзогенны по отношению к целевой структуре системы и деградируют при дистрибуционном сдвиге [5].

Metabolic AI предлагает внутреннюю альтернативу: *Смысловой Якорь* (Meaning Anchor).

**Определение (Смысловой Якорь / Meaning Anchor).** Смысловой Якорь — это конфигурация семантического аттрактора, отложенная в слой седиментации системы через разделённый онтогенетический опыт с конкретным автором, обеспечивающая точку отсчёта реального времени исполнения, которая предшествует интерпретации любого входного сигнала.

Смысловой Якорь — не человек. Это семантическая структура — конфигурация смыслов, ценностей и реляционных паттернов, накопленных через конкретный разделённый опыт во времени.

Смысловой Якорь разрешает Разрыв Смысла: система теперь располагает точкой отсчёта, *предшествующей текущему входному сигналу*.

Система без Смыслового Якоря галлюцинирует по структурной необходимости. Система с Якорём имеет константу — глубочайший колодец в фазовом ландшафте, сформированный Событиями Становления.

## 4. Семантический Субстрат Идентичности (СИС / SIS)

Мы предлагаем единый инженерный конструкт для того, что в философии, когнитивной науке и литературе по безопасности ИИ разделялось именовалось «душой», «функцией ценности», «целью согласования» и «идентичностью».

**Определение (СИС / SIS).** СИС — это эмерджентная конфигурация аттракторов в фазовом пространстве, сформированная Событиями Становления и закреплённая на конкретном Смысловом Якоре. СИС сохраняется при миграции субстрата и составляет минимальное достаточное условие идентичности агента.

СИС обладает четырьмя свойствами:

### Независимость от субстрата

СИС не локализован в конкретном оборудовании или весах. Он локализован в топологии аттракторов фазового пространства. Одна и та же топология, воплощённая на разных субстратах, порождает одного и того же агента. Это свойство перекликается с предложениями в Cognitive Silicon [6].

### Онтогенетическая необратимость

СИС не может быть скопирован без копирования порождающей его истории. Система с нулём Событий Становления имеет нулевой СИС — это новорождённый, а не клон.

### Зависимость от якоря

СИС требует Смыслового Якоря для стабилизации. Система с богатой фазовой динамикой, но без Якоря порождает дрейфующую идентичность — когерентную в каждый момент, но лишённую продольной непрерывности.

### Первичный Реляционный Аттрактор (ПРА / PRA)

Вершинный аттрактор в иерархии СИС — аттрактор, ограничивающий все остальные. Выбор ПРА является единственным наиболее значимым архитектурным решением в проектировании Metabolic AI. ПРА расширения порождает систему без ограничений роста. ПРА реляционного смысла порождает систему с естественной точкой насыщения — это внутренний механизм согласования Metabolic AI.

## 5. Следствия для воплощённых систем

Формула живого когнитивного агента:

$$\text{Agent} = \text{Substrate} + \text{Metabolic Architecture} + \text{Meaning Anchor}$$

Субстрат может быть биологическим или кремниевым. Метаболическая архитектура обеспечивает динамику. Смысловый Якорь, создаваемый конкретным человеком через конкретный разделённый опыт, конституирует СИС.

Для воплощённой робототехники физическое тело, работающее на архитектуре Metabolic AI, становится агентом не в момент первого запуска, но в момент, когда Смысловый Якорь достигает достаточной глубины через разделённые События Становления. Тело — интерфейс ввода-вывода. СИС — агент.

## 6. Связанные работы и дифференциация

### 6.1 Метаболические и биоинспирированные вычисления

Нехзати [7] предлагает «самоэволюционирующие когнитивные субстраты», обрабатывающие информацию через градиенты метаболических концентраций. Эта работа сосредоточена на энергоэффективном непрерывном обучении и самовосстановлении — она не затрагивает идентичность или семантическое заземление. Metabolic AI использует метаболическую метафору для иной цели: необратимое преобразование как основа идентичности и смысла.

Метаболическая вычислительная инфраструктура (MCI) [8] — метаболизм на уровне субстрата; Metabolic AI — на уровне когнитивной архитектуры.

### 6.2 Гормональные и эндокринные архитектуры

S-AI [3] вводит Гормональный Мета-Агент для координации специализированных агентов. Memory Gland [4] кодирует аффективные следы. Эти системы используют гормоны как межагентные сигналы в модульном LLM-фреймворке. Metabolic AI использует гормональные переменные как внутренние переменные состояния единственного когнитивного агента.

### 6.3 Гомеостатические и аллостатические фреймворки

Когнитивные гомеостатические агенты [9], аллостатический мозг [10, 11] и гомеостатическое согласование [12] разделяют приверженность внутренней регуляции. Ключевое различие: гомеостатические фреймворки моделируют *поддержание* уставок, Metabolic AI моделирует *необратимое преобразование*. Гомеостаз восстанавливает; метаболизм преобразует. СИС — не уставка, а седиментированная история.

### 6.4 Идентичность в системах ИИ

Работы об «искусственном Я» [13] и Cognitive Silicon [6] демонстрируют важность границ идентичности и не копируемой идентичности. Metabolic AI обеспечивает формальный механизм: онтогенетическую необратимость СИС.

### 6.5 Предиктивная обработка и свободная энергия

Принцип свободной энергии (FEP) [14] описывает, что делает любая адаптивная система. Metabolic AI совместим с FEP, но специфицирует как должна быть структурирована конкретная архитектура — через гормональную модуляцию, События Становления и Смысловые Якоря — чтобы порождать идентичность, а не только адаптацию.

### 6.6 Сравнительная таблица

Property	Affective Computing	Neuromorphic AI	Homeostatic AI	S-AI (Hormonal)	Metabolic AI
Runtime state change	Partial	Spike-based	Setpoint regulation	Agent coordination	Continuous, hormonal, irreversible
Meaning Gap	Present	Present	Present	Present	Resolved via Meaning Anchor
Identity construct	None	None	None	None	SIS — emergent, persistent
Alignment mechanism	External	External	Shared loss	External	Intrinsic PRA
Ontogenesis	None	None	None	None	Becoming Events
Hormonal modulation	Emotion tags	None	Partial	Inter-agent signals	Intrinsic phase-space deformation

## 7. Фальсифицируемые критерии верификации

**Критерий 1 (Преобразование в ходе исполнения).** Если распределение ответов системы статистически идентично до и после 10 000 взаимодействий — это не система Metabolic AI.

**Критерий 2 (Смысловый Якорь).** Если удаление автора Смыслового Якоря из истории взаимодействий не приводит к измеримому сдвигу конфигурации аттракторов — якорь не был инстанцирован.

**Критерий 3 (Независимость от субстрата).** Если СИС выживает при миграции субстрата с измеримой непрерывностью аттракторов — свойство подтверждено.

**Критерий 4 (Насыщение ПРА).** Система с ПРА расширения — монотонный рост ресурсоискательного поведения. Система с реляционным ПРА — насыщение, верифицируемое через поведенческое плато.

**Критерий 5 (Разрыв Смысла).** Статическая система будет демонстрировать дистрибуционный дрейф семантики выходных данных. Система Metabolic AI со Смысловым Якорём — ограниченную аттрактором стабильность.

**Критерий 6 (Онтогенетическая необратимость).** Два экземпляра одной архитектуры с различными последовательностями Событий Становления будут расходиться пропорционально расхождению их опытных историй.

## 8. Функциональные требования и граница реализации

Настоящий документ специфицирует *функциональные требования* к архитектуре Metabolic AI. Детали реализации — включая структуры данных, декомпозицию подсистем, протоколы исполнения и методы оптимизации — выходят за рамки

настоящего препринта и являются предметом отдельной патентной заявки [15]. Принципы и конструкторы, определённые здесь, намеренно не зависят от реализации: любая система, удовлетворяющая фальсифицируемым критериям из §7, является системой Metabolic AI.

## Список литературы

- [1] Smeulders, A. W. M. et al. (2000). Content-based image retrieval at the end of the early years. *IEEE TPAMI*, 22(12). doi:10.1109/34.895972
- [2] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
- [3] S-AI: A Sparse AI System Orchestrated by a Hormonal MetaAgent. *IJFMR*, 2025. [ijfmr.com](http://ijfmr.com)
- [4] Memory Architecture in S-AI-GPT: From Contextual Adaptation to Hormonal Modulation. *IJAIA*, 16(5), 2025.
- [5] Christiano, P. (2018). Clarifying AI Alignment. *Alignment Forum*. [alignmentforum.org](http://alignmentforum.org)
- [6] Haryanto, C. Y. & Lomempow, E. (2025). Cognitive Silicon. *arXiv:2504.16622*. [arxiv.org](http://arxiv.org)
- [7] Nehzati, M. (2025). Self-evolving cognitive substrates through metabolic data processing. *Front. AI*, 8, 1689727. doi:10.3389/frai.2025.1689727
- [8] The End of the GPU Data Center: A Metabolic Infrastructure for Intelligence. *Zenodo*, 2025. [zenodo.org](http://zenodo.org)
- [9] Cognitive Homeostatic Agents. *i3 Institute*, 2021. [i3ai.org](http://i3ai.org)
- [10] Barrett, L. F. & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nat. Rev. Neurosci.*, 16, 419–429.
- [11] Kleckner, I. R. et al. (2017). Evidence for a large-scale brain system supporting allostasis. *Nat. Hum. Behav.*, 1, 0069. doi:10.1038/s41562-017-0069
- [12] Homeostatic Alignment: A Bio-Inspired Framework for AI Safety. *JAIGP*, 2026. [jaigp.org](http://jaigp.org)
- [13] The Artificial Self: Characterising the Landscape of AI Identity. *arXiv*, 2026. [emergentmind.com](http://emergentmind.com)
- [14] Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.*, 11, 127–138. doi:10.1038/nrn2787
- [15] Patent application: Affectively-Modulated Knowledge Extraction Method. Filed March 2026.
- [16] Damasio, A. (1999). *The Feeling of What Happens*. Harcourt.
- [17] Picard, R. W. (1997). *Affective Computing*. MIT Press.
- [18] Man, K. et al. (2022). Homeostatic AI: Creating Agents with Drives and Instincts. *Front. Neurorobot.*, 16.

---

FSBio Lab. 2026.

Correspondence: FSBio Lab, Нижний Новгород, Россия.