

*Д.И. Дубровский*

### **Коммуникативный подход к анализу критериев оценки социальной значимости развития искусственного интеллекта\***

*Дубровский Давид Израилевич* – доктор философских наук, профессор, главный научный сотрудник. Институт философии РАН. Российская Федерация, 109240, г. Москва, ул. Гончарная, д. 12, стр. 1; e-mail: ddi29@mail.ru

Коммуникативный подход представляет собой анализ информационного процесса, включающего участие различных социальных субъектов в их взаимосвязи друг с другом (феномен полисубъектности), но в то же время и с системами искусственного интеллекта (ИИ). При этом выясняются позитивные и негативные аспекты влияния новаций ИИ на социальную деятельность. Акты коммуникации осуществляются в форме информационной причинности, которая отличается от физической причинности, но не противоречит ей. Проведен анализ специфики информационной причинности, обусловленной принципами изофункционализма систем и инвариантности информации по отношению к физическим свойствам ее носителя. Показано, что психическая (ментальная) причинность является видом информационной причинности и тем самым получает научное объяснение на основе информационного подхода. Это целиком относится к объяснению информационного управления и психического управления, что позволяет более основательно осмыслить их роль в познавательной, коммуникативной, практической деятельности социального субъекта и, как показано в статье, открывает более широкие возможности для решения вопросов о социогуманитарных основаниях критериев оценки инноваций, использующих достижения ИИ и методов цифровизации.

**Ключевые слова:** искусственный интеллект (ИИ), социальная значимость ИИ, критерии оценок, коммуникативный подход, социальный субъект, информационные процессы, кодирование, передача, декодирование информации, информационная причинность, психическая причинность, информационное управление, информационная безопасность

---

\* Работа выполнена при поддержке РНФ, проект № 21-18-00184 «Социогуманитарные основания критериев оценки инноваций, использующих цифровых технологии и искусственный интеллект».

## Структура, механизмы и функции коммуникативных процессов

Коммуникативный процесс осуществляется определенным субъектом, имеющим соответствующие цели, ценностные установки, практические и иные задачи взаимодействия с другими субъектами и различными объектами. Он представляет собой процесс выбора, передачи и использования определенной информации как в собственных интересах, так и в интересах других субъектов, преследуя большей частью задачи управления ими или управления избранным объектом (природным, социальным, техническим). Коммуникативный процесс, независимо от его конкретного содержания и ценностных установок, является **инвариантным** по своей структуре **информационным процессом**, в том смысле, что включает одни и те же основные компоненты и операции. Несмотря на то что они хорошо известны, их все же стоит четко выделить и проанализировать, взяв за систему отсчета активный субъект (инициатора коммуникации).

Здесь мы видим: 1) выбор паттерна «готовой» информации (или специально сформированной); 2) кодирование ее (часто в форме общедоступного кода, скажем, средствами русского языка), но иногда создание для этого оригинального кода, понятного лишь избранным коммуникантам и недоступного для других; 3) передача этой информации по существующему обычному каналу связи; но в случае необходимости по специальному, хорошо защищенному в целях безопасности каналу; 4) декодирование ее получателем в форме достаточной для «понимания» («принятия») и использования в его целях управления.

Таковы основные структурные компоненты коммуникативного процесса, каждый из которых имеет свои специфические ресурсы для его совершенствования в рамках данного социального субъекта, постоянно вынужденного решать задачи общения, успешной самоорганизации, сохранения своей функциональной целостности и дееспособности.

Здесь, однако, возникают крайне сложные вопросы, касающиеся типизации субъектов по их видам деятельности и выяснения их места в чрезвычайно разветвленной, противоречивой динамической структуре полисубъектных социальных сред, образующих в итоге мировой социум. Потребность обращения к этому всеобъемлющему субъекту вызвана тем, что любой выделяемый нами социальный субъект является элементом или подсистемой другого, более сложного социального субъекта или связан в определенной степени различными взаимодействиями и взаимозависимостями со многими другими субъектами.

Такие беспредельно сложные процессы коммуникации, конечно, не поддаются четкому анализу, но их надо иметь в виду, когда мы рассуждаем о критериях социальной значимости новаций в развитии ИИ и процессов цифровизации. Реалистический подход к этой проблематике вынуждает большей частью ограничиваться конкретными, жизненно значимыми для основательного исследования социальными субъектами. Прежде всего, это наша страна в целом, ее государственные, социальные, экономические структуры, системы образования, здравоохранения, обеспечения обороноспособности и информационной безопасности, все те социальные организации, от деятельности которых зависит благополучие нашего народа. При этом, однако, мировой социум, земная

социальная самоорганизация обладает рядом всем известных общих негативных свойств, которые, как это предельно ясно, должны стоять на первом плане при рассмотрении коммуникативных проблем в международном аспекте и, в связи с этим, в масштабах нашей страны. Предельное обострение глобального кризиса мировой цивилизации ставит перед нами чрезвычайно актуальные, судьбоносные вопросы, связанные в том числе с использованием новаций ИИ, и мы к ним еще вернемся.

Но надо подчеркнуть, что отмеченные выше трудности выделения определенного субъекта, выяснения его многоуровневых и многоплановых связей с другими субъектами относятся и к тем случаям, когда перед нами, казалось бы, такие знакомые, сравнительно «узкие», «понятные» субъекты, как, скажем, российская система образования или здравоохранения. Задача определения критериев социальной значимости новаций ИИ для этих систем требует глубокого теоретико-методологического анализа с позиций постнеклассической научной рациональности [Лекторский, 2001; Лекторский 2018]. Это вызвано тем, что здесь необходима интеграция, своего рода конвергенция в единой концептуальной структуре множества разнородных факторов и явлений: учет **полисубъектности** коммуникативных процессов и того обстоятельства, что в большинстве случаев мы имеем дело с **гибридной реальностью**, в которой объединяются **субъектные, цифровые и физические** составляющие.

Понятно, что искомые критерии оценки цифровых новаций, включая роль физических факторов, относятся к определенному субъекту. Но тут также возникают проблемные ситуации, касающиеся выделенного субъекта. Они часто обусловлены тем, что в тени оставляют различие **надличностного субъекта** (им является в данном случае система образования) и **личности как полноправного субъекта коммуникации** со своими специфическими свойствами и интересами. Одно дело, когда оценка прилагается к системе образования и ее разным подсистемам и элементам (школы, техникумы, институты), другое, когда мы переходим на уровень реальной личности, которая в конечном итоге должна служить главным субъектом оценки, влияния новаций ИИ, включающей как положительные, так и негативные значения.

Хорошо известно, что нередко нам официально сообщают положительную оценку, относящуюся к системе в целом или к отдельным крупным образовательным учреждениям, в то время как мы знаем о многочисленных случаях, когда учащиеся, их родители и сами педагоги резко отрицательно оценивают различные способы использования цифровых новаций. Конечно, часто такой разрыв в оценках – следствие чрезмерной бюрократизации системы образования, которая ведь тоже является своеобразным феноменом коммуникации, представляющим системный эффект, производимый структурами более высокого ранга.

Все это указывает на то принципиальное обстоятельство, что **параметр личности** как субъекта коммуникации с ее специфическими свойствами субъективной реальности, самополагания и самоорганизации явно недостаточно учитывается при осмыслении коммуникативных процессов и определения оценки значения в них новаций ИИ. Соотношение личностного и надличностного – методологически важный пункт в исследовании полисубъектности

коммуникативных процессов. Здесь остаются и многие другие актуальные вопросы, касающиеся создания и использования эффективных эпистемологических средств междисциплинарных исследований.

В.Е. Лепским показано, что для успешной разработки этой крайне сложной, многомерной проблематики методы классической эпистемологии недостаточны. Здесь необходимо широкое использование идей и методов постнеклассической научной рациональности, принципов постнеклассической эпистемологии. Опираясь на эти средства, им была разработана и предложена концепция саморазвивающихся полисубъектных (рефлексивно-активных) сред (кибернетики третьего порядка), которая охватывает методологические вопросы формирования и функционирования гибридной реальности, разработки нелинейных моделей связей, характерных для таких сред. Эта концепция не ограничивается только сферой ИИ и цифровых технологий, но также включает институциональные, организационные и управленческие инновации, она формирует методологические основания для целостного представления и конвергенции **субъектов, средств и объектов** в познавательных процессах и разнообразных проявлениях деятельной активности полисубъектной среды [Лепский, 2010; Лепский 2018].

Важным достоинством концепции В.Е. Лепского является акцент на вопросах саморегуляции и самоорганизации, на оценке цифровых трансформаций в саморазвивающихся полисубъектных (рефлексивно-активных) средах; тем самым открываются новые перспективы для дальнейшей разработки этой столь актуальной проблематики.

В свете этой концепции мы можем более полно оценить социальное значение таких факторов коммуникативного процесса, как скорость, точность передачи информации, надежность способов ее кодирования и декодирования, сохранение и использование ее для целей управления.

### **Информационная причинность и акты управления в коммуникативных процессах**

Поскольку коммуникативные процессы представляют собой передачу и использование информации, возникает задача анализа тех факторов, которые определяют активность субъекта в этих процессах, его акты управления коммуникативными действиями как в отношении других субъектов, так и в его взаимодействиях с системами ИИ и физическими объектами. Это относится также к взаимодействию систем ИИ между собой без личного участия человека, ибо они совершаются на основе созданных им компьютерных и иных программ, и он так или иначе контролирует их реализацию. Во всех случаях эти действия осуществляются в форме **информационной причинности**, которая отличается от физической причинности. Поэтому важно рассмотреть ее особенности и функциональные возможности, которые играют существенную роль в использовании новаций ИИ в различных областях социальной жизнедеятельности.

Понятие информационной причинности – теоретический концепт **парадигмы функционализма**. Она в ряде существенных отношений альтернативна парадигме физикализма, которая до XIX века безраздельно господствовала

в науке, как и понятие физической причинности. Парадигма функционализма сформировалась к началу второй половины XX в. на основе достижений биологических, социальных и технических дисциплин, широкого использования системных и структурных подходов, особенно же исследований и результатов А. Тьюринга, Н. Винера и других выдающихся творцов компьютерных технологий и кибернетики. Парадигма функционализма выдвинула и обосновала **принцип изофункционализма систем: одна и та же функция может быть воспроизведена и реализована на разных по своим физическим свойствам носителях**; в эпистемологическом отношении это означает, что функциональные описания в принципе не могут быть выражены на языке физических описаний. Отсюда «Машина Тьюринга» и ее модели, независимые от физических описаний, само понятие информации, а применительно к ее производству и передаче **принцип инвариантности информации по отношению к физическим свойствам своего носителя** (возможность кодирования и передачи одной и той же информации самыми разными средствами и способами).

Соответственно определяется и понятие информационной причинности как не сводимое к физической причинности. Хотя информация необходимо воплощена в своем физическом носителе, она инвариантна по отношению к определенным физическим свойствам. А постольку в самоорганизующихся системах, таких как биологические и социальные, а соответственно, и в коммуникативных процессах многие явления не могут быть объяснены на основе физической причинности.

Когда действует физическая причина, то вызываемое ею следствие определяется исключительно величиной массы, энергии, пространственно-временными показателями. В случае информационной причины эти факторы недостаточны для определения следствия, а именно оно чаще всего имеет для нас главное значение. Когда я говорю вам: «Дайте мне, пожалуйста, посмотреть эту книгу» и вы передаете ее мне, то ожидаемое мной следствие вызывается, очевидно, не просто физическими свойствами звукового сигнала, а именно переданной и воспринятой вами информацией потому, что вы знаете русский язык, эту кодовую систему и в вашем мозгу мгновенно происходит расшифровка соответствующих кодов, т.е. достигается **понимание** информации. Если бы вы не знали русского языка, акт коммуникации не состоялся бы.

Впрочем, эту информацию я мог бы передать и многими другими способами. Но важен сравнительно простой для понимания, удобный для коммуниканта код. От этого во многом зависит результативность коммуникативного процесса и тем самым роль тех новаций ИИ, которые способствуют повышению качества коммуникации, взаимопонимания субъектов, организации их коллективной деятельности, но в то же время успешному управлению ими системами ИИ, а во многих отношениях и оптимизации взаимодействия последних между собой. Все это относится к проблемам программирования и развития систем ИИ, достижения последними способности осуществлять все более сложные функции, близкие к естественному интеллекту.

Таким образом, информационная причинность определяет специфический тип управления в коммуникативных процессах – **информационное управление** в отличие от физического управления (см. подробнее: [Дубровский, 2015;

Дубровский 2021]). Информационная причина действует на основе сложившейся у субъекта **кодовой зависимости** между определенными физическими явлениями и их функциональным значением для субъекта. Некоторые из них кратковременны (например, условный рефлекс), другие действительны для всей жизни индивида и, более того, всего существования вида и даже всей жизни на Земле (генетический код).

Информационная причинность не противоречит физической, так как не исключает, а предполагает участие физического фактора. Но она качественно расширяет диапазон научного объяснения и практического преобразования тех сложных процессов объективной реальности, которые характерны для биологических и социальных самоорганизующихся систем, а точнее – для биосоцио-технических систем, необходимо включающих личностный параметр (роль сознания, творческой активности, самоорганизации, ответственности социального индивида).

Столь большое место, отведенное мной информационной причинности, вызвано тем, что при рассмотрении критериев социальной значимости новаций ИИ не всегда четко различаются два аспекта: возникновение новаций в самих системах ИИ, с одной стороны, и новаций, вызванных этими новациями в системах социальной деятельности, с другой. Разумеется, оба аспекта тесно связаны, но в ряде случаев важную роль приобретает отдельный специальный анализ каждого аспекта. В этом отношении, когда речь идет о коммуникативной проблематике, возникает необходимость определения того круга технологических задач, решение которых сейчас особенно актуально. Имеется в виду то, что называют **«социальным заказом»** на создание новых технологий ИИ, обладающих первостепенной социальной значимостью. Это в первую очередь технологии бесперебойной связи и надежной системы управления, обеспечения информационной безопасности. Сюда относятся задачи эффективных способов передачи информации, ее кодирования, декодирования, в том числе для создания **интерфейсов «мозг-сознание-язык-компьютер-машина»**, разработка средств для более эффективного взаимодействия с интеллектуальными роботами, скажем, создание способов управления ими с помощью речи. Все это имеет прямое отношение к созданию Общего ИИ, который по своим возможностям приближается к функциям естественного интеллекта [Бурцев и др., 2021]. При рассмотрении этого аспекта особенно важен акцент на таких технологических достижениях ИИ, которые способны вызвать глобальные изменения социума в целом как положительного характера, так и несущие угрозу его гибели. Этот круг вопросов широко обсуждается в философской литературе, особенно в западной, и он должен занять большее место в наших исследованиях процессов антропотехнологической эволюции, темпы которой ускоряются.

Надо особо подчеркнуть, что для разработки всего обозначенного выше круга вопросов исключительно важное значение имеет проблема **психической (ментальной) причинности**. Это сразу видно, если даже ограничиться лишь коммуникативными процессами. Вряд ли надо доказывать фундаментальную роль психической причинности и, соответственно, психического управления во всех сферах общественной жизни, всюду, где действует личность. **Психическая причинность является видом информационной причинности,**

она в конечном счете является основой всех видов деятельности социального субъекта, наиболее ясно выражает значение того личностного параметра, о котором говорилось выше.

Однако теоретическое обоснование психической причинности и психического управления всегда служило камнем преткновения для естественнонаучного объяснения, поскольку требовало решения так называемой «Трудной проблемы сознания». Суть ее в том, что осознаваемым психическим явлениям с их качеством субъективной реальности нельзя приписывать физические свойства (массу, энергию, пространственные характеристики). Как в таком случае можно теоретически корректно объяснить связь явлений субъективной реальности с деятельностью мозга и их способность управлять телесными процессами? Большой частью для ответа на этот вопрос предлагались редукционистские решения физикалистского типа. Однако редукция психических явлений к физическим, их отождествление вместо объяснения специфических свойств психических явлений фактически просто устраняет их как нереальные. В этом несостоятельность редукционистских подходов.

На протяжении нескольких десятилетий я занимался разработкой «Трудной проблемы сознания» под углом **научного** объяснения качества субъективной реальности в рамках проблемы «Сознание и мозг», оставляя в стороне метафизические дискуссии. В предложенной мной концепции [Дубровский, 1971; Дубровский 2015; Дубровский 2019] на основе информационного подхода проведено научное теоретическое обоснование психической причинности и психического управления, включая объяснение прежде всего качества субъективной реальности и задач методологии расшифровки мозговых кодов психических явлений. При этом мною широко использовались результаты современной нейронауки и ряда смежных с ней дисциплин ([Анохин, 2021] и др.), а также собственные феноменологические исследования субъективной реальности [Дубровский, 2002]. Разумеется, предложенная концепция должна пройти основательные критические испытания. В любом случае, однако, реальность информационной причинности и информационного управления не подлежит сомнению. Их роль в коммуникативных процессах должна быть предметом дальнейшего основательного исследования.

### Список литературы

Анохин, 2021 – Анохин К.В. Когнитом: В поисках фундаментальной нейронаучной теории сознания // Журнал высшей нервной деятельности. 2021. Т. 71. № 1. С. 39–71.

Бурцев и др., 2021 – Бурцев М.С., Бухвалов О.Л., Ведяхин А.А. и др. Сильный искусственный интеллект. На подступах к сверхразуму. М.: Интеллектуальная литература, 2021. 237 с.

Дубровский, 1971 – Дубровский Д.И. Психические явления и мозг. Философский анализ проблемы в связи с некоторыми актуальными задачами нейрофизиологии, психологии и кибернетики. М.: Наука, 1971. 386 с.

Дубровский, 1980 – Дубровский Д.И. Информация, сознание, мозг. М.: Высшая школа, 1980. 286 с. 2-е изд., доп. М.: ЛЕНАНД, 2021. 304 с.

Дубровский, 2002 – Дубровский Д.И. Проблема идеального. Субъективная реальность. М.: Канон+, 2002. 368 с.

Дубровский, 2015 – Дубровский Д.И. Проблема «Сознание и мозг»: Теоретическое решение. М.: Канон+, 2015. 208 с.

Лекторский, 2001 – Лекторский В.А. Эпистемология классическая и неклассическая. М.: Эдиториал УРСС, 2001. 256 с.

Лекторский, 2018 – Лекторский В.А. Человек и культура. Избранные статьи. СПб.: СПбГУП, 2018. 640 с.

Лепский, 2010 – Лепский В.Е. Рефлексивно-активные среды инновационного развития. М.: Когито-Центр, 2010. 255 с.

Лепский, 2021 – Лепский В.Е. Искусственный интеллект в субъектных парадигмах управления // Философские науки. 2021. Т. 64. № 1. С. 88–101.

Dubrovsky, 2019 – Dubrovsky D.I. “The Hard Problem of Consciousness”. Theoretical solution of its main questions // AIMS Neuroscience. 2019. Vol. 6. Issue 2. P. 85–103.

## **Communicative approach to the evaluation criteria analysis of the social significance of the artificial intelligence development**

*David I. Dubrovsky*

Institute of Philosophy, Russian Academy of Sciences. 12/1 Gonchamnaya Str., Moscow 109240, Russian Federation; e-mail: ddi29@mail.ru

The communicative approach is an analysis of the information process involving the participation of various social actors in their relationship with each other (the phenomenon of polysubjectivity), but at the same time with artificial intelligence (AI) systems. At the same time, the positive and negative aspects of the impact of AI innovations on social activity are being clarified. Acts of communication are carried out in the form of informational causality, which differs from physical causality, but does not contradict it. The analysis of the specifics of informational causality due to the principles of isofunctionalism of systems and the invariance of information in relation to the physical properties of its carrier is carried out. It is shown that psychic (mental) causality is a type of informational causality and thereby receives a scientific explanation based on the informational approach. This entirely relates to the explanation of information management and mental management, which allows us to more thoroughly comprehend their role in the cognitive, communicative, practical activities of a social subject and, as shown in the article, opens up wider opportunities for solving questions about the socio-humanitarian foundations of the criteria for evaluating innovations using AI achievements and digitalization methods.

**Keywords:** artificial intelligence (AI), social significance of the AI, evaluation criteria, communicative approach, social subject, information processes, encoding, transfer, information decoding, information causality, mental causality, information management, information security

**Acknowledgments:** The work was supported by the Russian Science Foundation, project No. 21-18-00184 “Social and humanitarian foundations for evaluation criteria for innovations using digital technologies and artificial intelligence”.

## **References**

Anokhin, K.V. Kognitom: “V poiskakh fundamental’noi neironauchnoi teorii soznaniya” [Cognitom: In search of a fundamental neuroscientific theory of consciousness], *Zhurnal vysshei*

*nervnoi deyatelnosti* [Journal of Higher Nervous Activity], 2021, vol. 71, no. 1, pp. 39–71. (In Russian)

Burtsev, M.S., Bukhvalov, O.L., Vedyakhin, A.A. et al. *Sil'nyi iskusstvennyi intellekt. Na podstupakh k sverkhrazumu* [Strong artificial intelligence. On the approaches to the supermind]. Moscow: Intellektual'naya literature Publ., 2021. 237 pp. (In Russian)

Dubrovsky, D.I. *Psichicheskie yavleniya i mozg. Filosofskiy analiz problemi v svyazi s nekotorymi aktualnymi voprosami neyrophiziologii, psichologii i kibernetiki* [Psychic phenomena and the brain: the Philosophical analysis of the problem due to some urgent problems of neurophysiology, psychology and cybernetics]. Moscow: Nauka Publ., 1971. 386 pp. (In Russian)

Dubrovsky, D.I. *Problema ideal'nogo. Sub"ektivnaya real'nost'* [The problem of the ideal. Subjective reality]. Moscow: Kanon+ Publ., 2002. 368 pp. (In Russian)

Dubrovsky, D.I. *Informatciya, Soznanie, Mozg* [Information. Consciousness. Brain]. Moscow: Higher School Publ., 1980. 286 pp. (In Russian)

Dubrovsky, D.I. *Problema "Soznanie I mozg": Teoreticheskoe reshenie* [The Problem "Consciousness and brain": The theoretical solution]. Moscow: Canon+ Publ., 2015. 208 pp. (In Russian)

Dubrovsky, D.I. "The Hard Problem of Consciousness'. Theoretical solution of its main questions", *AIMS Neuroscience*, 2019, vol. 6, issue 2, pp. 85–103.

Lectorsky, V.A. *Epistemologiya klassicheskaya i neklassicheskaya* [Classical and non-classical epistemology]. Moscow: Editorial URSS Publ., 2001. 256 pp. (In Russian)

Lectorsky, V.A. *Chelovek i kul'tura. Izbrannye stat'i* [Man and Culture. Selected articles]. Saint Petersburg: SPbUHSS Publ., 2018. 640 pp. (In Russian)

Lepskiy, V.E. *Refleksivno-aktivnye sredy innovatsionnogo razvitiya* [Reflexive-active environments of innovative development]. Moscow: Kogito-Center Publ., 2010. 255 pp. (In Russian)

Lepskiy, V.E. "Iskusstvennyi intellekt v sub"ektnykh paradigmakh upravleniya" [Artificial Intelligence in Subject-Oriented Control Paradigms], *Filosofskie nauki*, 2021, vol. 64, no. 1, pp. 88–101. (In Russian)