

Convivial or Manipulative Tools? LLMs in Russian Higher Education (The Case of the School of Advanced Studies)

LADA V. SHIPOVALOVA

European University at St. Petersburg, St. Petersburg, Russian Federation

<https://orcid.org/0000-0002-1989-1152>

ASYA A. FILATOVA

Moscow Institute of Physics and Technology, Moscow, Russian Federation

<https://orcid.org/0000-0002-0497-0018>

224

Рекомендация для цитирования:

Шиповалова Л. В., Филатова А. А. (2026) Конвивиальные или манипулятивные инструменты? LLMs в российском высшем образовании (кейс Школы перспективных исследований). *Социология власти*, 38 (1): 224-243
EDN: PSAIKN

For citation:

Shipovalova L. V., Filatova A. A. (2026) Convivial or Manipulative Tools? LLMs in Russian Higher Education (The Case of the School of Advanced Studies). *Sociology of Power*, 38 (1): 224-243

Поступила в редакцию: 11.02.2026;
прошла рецензирование: 15.03.2026;
принята в печать: 17.03.2026
Received: 11.02.2026; Revised:
15.03.2026; Accepted: 17.03.2026



© Authors, 2026

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This article examines the integration of large language model based chatbots into educational practices in Russian higher education through the lens of critical theory and curriculum ideologies. Drawing on Ivan Illich's concept of convivial tools and Jürgen Habermas's distinction between instrumental and communicative rationality, the authors conceptualize large language models as technologies capable of operating either within a manipulative logic, reinforcing hierarchies, technocratic rationality, and domination, or within a convivial logic, fostering autonomy, collective reasoning, and communicative action. The mode of implementation and functioning of these technologies largely depends on the underlying curriculum ideology, whether it is grounded in hierarchical principles of knowledge production and transmission (Scholar Academic ideology) or in pedagogical constructivism and individual self-actualization (Learner Centered ideology). The empirical component is based on the experience of the School of Advanced Studies at the University of Tyumen. The analysis draws on thematic examination of 20 semi structured interviews with second to fourth year students who participated in experimental courses featuring AI personas designed to substitute for certain instructor functions.

These courses involved a shift toward interactive learning formats and the introduction of a mediator role inspired by the notion of the ignorant schoolmaster. The findings indicate that compulsory and instrumentalist deployment of chatbots is perceived by students as manipulative. It elicits resistance, feelings of dehumanization, diminished motivation, and disappointment over the loss of direct human interaction. In contrast, voluntary and reflexive uses of large language models unlock their potential as convivial tools, enabling the co-production of knowledge and enhancing epistemic agency. In conclusion, the authors outline the conditions under which large language models can function as genuine tools of conviviality in university education: voluntary adoption, transparency in design and operation, rejection of total instrumentalization, and embedding within genuinely communicative practices that respect participants' autonomy and foster negotiated, collective meaning-making.

Keywords: large language models, higher education, curriculum ideologies, instrumental rationality, communicative rationality, tools for conviviality, manipulative tools

Конвивиальные или манипулятивные инструменты? LLMs в российском высшем образовании (кейс Школы перспективных исследований)

225

Лада В. Шиповалова

Европейский университет в Санкт-Петербурге, Санкт-Петербург, Российская Федерация
<https://orcid.org/0000-0002-1989-1152>

Ася А. Филатова

Московский физико-технический институт, Москва, Российская Федерация
<https://orcid.org/0000-0002-0497-0018>

Резюме: В статье анализируется опыт внедрения чат-ботов на основе больших языковых моделей в образовательную практику российского высшего образования через призму критической теории и идеологий учебных программ. Опираясь на концепцию «конвивиальных инструментов» Ивана Илича и различие между инструментальной и коммуникативной рациональностью Юргена Хабермаса, авторы рассматривают LLM как технологию, способную действовать либо в манипулятивной логике (усиление иерархий, технократической рациональности и доминирования), либо в конвивиальной (способствование автономии, коллективному мышлению и коммуникативному действию). Способ внедрения и функционирования технологий

во многом зависит от того, какая «идеология» лежит в основе учебной программы, построена ли она на вертикальных принципах производства и трансляции знания (академическая идеология) или на идеях педагогического конструктивизма и личностного самопознания (студентоцентричная идеология). Эмпирическая часть исследования основана на опыте Школы перспективных исследований (SAS) Тюменского государственного университета. Анализируются результаты 20 полуструктурированных интервью со студентами 2-4 курсов, участвовавшими в экспериментальных программах с использованием «ИИ персон», замещающих часть функций преподавателей. Курсы предполагали переход к интерактивным форматам обучения и введение роли медиатора («невежественного учителя»). Результаты показывают, что принудительное и инструменталистское использование чат-ботов воспринимается студентами как манипуляция, вызывает сопротивление, чувство дегуманизации, снижение мотивации и разочарование в утрате живого общения. Напротив, добровольное и рефлексивное применение LLMs раскрывает их потенциал как конвивиальных инструментов, способствующих совместному производству знания и эпистемической самостоятельности. В заключении авторы формулируют условия, при которых LLM могут стать подлинными инструментами конвивиальности в университетском образовании: добровольность использования, прозрачность конструкции, отказ от тотальной инструментализации, включение в коммуникативные практики.

226

Ключевые слова: большие языковые модели, высшее образование, идеологии учебных программ, инструментальная рациональность, коммуникативная рациональность, конвивиальные инструменты, манипулятивные инструменты

Introduction

The focus of this article emerges from a crossroads at which educational institutions, including universities, have found themselves over the past five years following the emergence of large language models (LLMs) and their incorporation into the production of academic texts. Scholarly discussion on the integration of AI and LLMs into education has expanded to such an extent that systematic reviews of the field are already available (Fowler, 2023; Bearman, Ryan & Ajjawi, 2023). The depth of this integration has also prompted attempts to assess epistemic trust in these technologies within educational contexts (Pandey, Mishra, Pandey, et al., 2025). Initial responses, often characterised by rejection and prohibition, have gradually given way to reflections on the inevitability of adoption, as well as to more balanced analyses of both the risks and the potential benefits of AI use in education (Lindsay & Jacka, 2024; Sembey, Hoda & Grundy, 2024).

As is often the case with emergent phenomena, however, significant gaps remain. These gaps are reflected in the relative scarcity of field

research that can document how new technologies are implemented within established educational contexts, as well as relevant theoretical interpretations of such implementations. This study seeks to address these gaps. It is situated at the intersection of conceptual framework and a case of AI deployment in higher education. The empirical case represented is the School of Advanced Studies (SAS), founded in 2017 at the University of Tyumen as part of Russia's Project 5-100, aimed at increasing the global competitiveness of Russian universities (T-Universities, 2019)¹.

In developing the theoretical foundations we begin with a general distinction between instrumental and communicative rationality (Habermas, 1986). Technological infrastructures and specific technologies are commonly assumed to be aligned "by definition" with instrumental rationality. However, late twentieth- and early twenty-first-century studies of technology have challenged this assumption by conceptualising technology as an active participant in communication (A. Mol; M. de Laet; G. Simondon; B. Stiegler), as an element of distributed cognition and distributed responsibility (E. Hutchins; L. Floridi). These approaches raise the question of how is such a dual existence of technology possible?

A possible answer to this question is offered by a theoretical framework central to our analysis: Ivan Illich's concept of tools for conviviality (Illich, 1973). Illich's work is socially critical and resonates with the tradition of the Frankfurt School; it is also closely connected to his critique of education and educational institutions (Illich 1971) and has already been discussed in contemporary debates on technology and learning (Qutieshat, 2025; Beinsteiner, 2020). Engaging with Illich's ideas allows us, at a theoretical level, to articulate a possible connection between instrumental and communicative rationality, in which specific technologies—LLMs in our case— may operate either as manipulative instruments or as tools for conviviality and communication.

Accordingly, the short theoretical part of the article develops a conceptual framework and formulates a general question concerning the role of AI in education. The second, empirical part represents a con-

1 Project 5-100 was a Russian state initiative (2012–2020) aimed at integrating leading universities into the global educational landscape and boosting their competitiveness. Its main goal was to place at least five Russian institutions in the top 100 of major international rankings (QS, THE, ARWU). To this end, participating universities had to increase the share of foreign students to 15% and foreign faculty to 10%. A key measure of research success was improved publication citation rates, driven by English-language publishing and international co-authorship. Overall, the project promoted the transformation of Russian universities along Western lines, prioritizing global visibility, English-medium scientific communication, and internationalization.

crete educational experience carried out by the SAS in which AI “personas” were introduced into university courses, focusing on students’ responses to the deliberate use of LLMs in teaching and learning practices. This experience is interpreted in terms of AI functioning either as a tool for conviviality or as a means of manipulation within different educational ideologies, understood as structures that organise university goals. By combining the theoretical and empirical parts, we pursue a dual aim: on the one hand, to illustrate Illich’s ideas through a specific case; on the other, to use them for interpreting the educational experience of implementing LLMs and identifying ways to improve this experience.

Convivial and Manipulative Tools on the Border between Communicative and Instrumental Rationality

228

The contemporary tension surrounding the integration of technologies into social relations, including those structured by educational institutions, may be situated within the Frankfurt School’s critique of instrumental reason. Questions of the social efficiency of education and science have become central to debates on academic capitalism and the university’s “third mission” (Compagnucci & Spigarelli, 2020). From a theoretical perspective, such an emphasis on social efficiency may be interpreted as the transformation of scientific and technological progress into a means of legitimising domination, both over nature and over human beings (H. Marcuse; J. Habermas). As an alternative to such “repressive” domination, Habermas proposes an ideology of emancipation grounded in communicative action or symbolically mediated interaction, whose normative force is constituted intersubjectively. He draws attention to “students and pupils” as potential subjects of communicative rationality, arguing that their relative lack of fixed interests, partial immunity to technocratic consciousness, and latent protest potential position them as possible agents of resistance to repressive social systems (Habermas, 1986, p. 120). At the same time, Habermas remains skeptical of Marcuse’s proposals for a New Technology oriented toward interaction rather than control and the suppression of individuality (Ibid., p. 88).

We assume that Ivan Illich’s critique of the institutionalisation of education, articulated through his distinction between convivial and manipulative tools, offers the opportunity of development of these ideas. Illich’s framework may be applied not only to the university as an institution but also to the technologies employed in educational practices, particularly LLMs, whose ambivalent effects have recently attracted scholarly attention. We suggest that, through Illich’s ideas, it is possible to describe technologies capable of reinforcing hierarchy and

efficiency-driven domination, while also potentially expanding communal autonomy and creative agency.

Relating Illich's ideas to Habermas's problematic requires clarification. The two thinkers share a critical orientation toward power, whether exercised by political actors or epistemic authorities, and converge in their rejection of linear narratives of scientific and technological progress. Both engage, albeit differently, with questions of emancipation and domination in education and technology (Tilak & Glassman, 2020). Their critical trajectories intersect in their emphasis on interaction and conviviality. Moreover, Illich's concept of convivial tools may be read as an implicit response to Habermas's skepticism toward Marcuse's vision of a New Technology, reframing technological emancipation not as a utopian future but as the recovery of suppressed or forgotten possibilities.

Illich conceptualises institutional development as unfolding in two watersheds: "At first, new knowledge is applied to the solution of a clearly stated problem and scientific measuring sticks are applied to account for the new efficiency. But at a second point, the progress demonstrated in a previous achievement is used as a rationale for the exploitation of society as a whole in the service of a value which is determined and constantly revised by an element of society, by one of its self-certifying professional elites" (Illich, 1973, p. 20).

229

Firstly, institutions function as a means enabling human activity, allowing alternatives and self-limitation. Applied to universities, this stage corresponds to spaces of intellectual exploration and individual growth (Qutieshat, 2025, p. 51). In the second watersheds, institutional survival and efficiency become ends in themselves, reinforced by bureaucratic management and quantitative indicators. Development shifts "from facilitating activity to organizing production" (Illich, 1971, p. 53), rendering institutions manipulative rather than enabling conviviality. "Any institution that moves toward its second watershed tends to become highly manipulative. For instance, it costs more to make teaching possible than to teach" (Illich, 1973, p. 36).

The distinction between convivial and manipulative tools is thus genetic rather than static. Conviviality presupposes a continual return to the origin of creative activity, where institutions serve human self-formation rather than subordinating it. "I choose the term 'conviviality' to designate the opposite of industrial productivity. I intend it to mean autonomous and creative intercourse among persons, and the intercourse of persons with their environment ... I consider conviviality to be individual freedom realized in personal interdependence and, as such, an intrinsic ethical value" (Illich, 1973, p. 24). "

Retrospectively, this distinction may be articulated as autonomy versus heteronomy, emancipation versus efficiency, and participation versus managerial control. Importantly, Illich does not envision the com-

plete absence of manipulative institutions (Illich, 1973, p. 37), but rather a balance that limits institutional power, resists compulsory structures, and preserves spaces of creative agency. His genetic account thus acquires ontological significance, pointing to the possibility of freedom through the preservation of institutional origins. In this respect, Illich's appeal to learner resistance (Illich, 1971, p. 41) resonates with Habermas's emphasis on communicative agency.

This theoretical constellation becomes particularly salient when we connect it with educational ideologies considered as frameworks organising curricular practice (Schiro, 2013). We incorporate the distinction between educational ideologies in order to specify how technologies function in education depending on the diverse goal orientations of university programmes. These ideologies differ in their goal structures but may be understood not only as opposed ideal types, but also as practical logics enacted in concrete institutional contexts. As such, the univocality of educational goal-setting becomes problematic. Below, we interpret these ideologies in relation to Illich's conceptual frameworks.

230 Within the Scholar Academic Ideology, the curriculum aims to induct students into a discipline's established body of knowledge and practices. Organised hierarchically around expert authority (Schiro, 2013, p. 4), this ideology may rely on technologies and institutional markers that function as manipulative tools legitimising epistemic dominance. Illich criticises this form of self-reproducing expertise as the rule of "self-certifying professional elites" (Illich, 1977). One possible transformation lies in rendering knowledge production transparent and fostering collective learning webs which, as convivial tools, enable independence within the alternative that "technology is available to develop either independence and learning or bureaucracy and teaching" (Illich, 1971, p. 77). Here, epistemic authority is not abolished but rendered intelligible in its genesis; technological transparency becomes as crucial as epistemic transparency, lest power asymmetries be reinforced (Beinsteiner, 2020).

The Social Efficiency Ideology defines education in relation to socially prescribed goals, positioning the teacher as a "manager" of learning aligned with the "needs of society (or client)" (Schiro, 2013, p. 4). Despite its overtly utilitarian character, this ideology may also be interpreted in nation-building terms or in relation to contemporary metrics of the university's societal impact. In our analysis, however, it is primarily associated with instrumental rationality.

The Learner Centered Ideology emphasises individual developmental trajectories shaped through interaction with the environment and requires curricular flexibility. Yet the exteriorisation of "human nature" in learning may itself be problematised, insofar as meaning emerges through interaction with others and with epistemic authorities. As Schiro notes, "People are stimulated to grow and construct meaning

as a result of interacting with their physical, intellectual, and social environments” (Schiro, 2013, p. 6). The ambivalence of this ideology lies in the risk that an emphasis on individual autonomy may obscure the collective sources of creativity and undervalue convivial technologies embedded in communal practices.

The Social Reconstruction Ideology foregrounds social injustice and integrates critical reflection and action into educational practice. While it opposes social efficiency by prioritising emancipation over affirmation, the two may converge when practices aimed at overcoming injustice are reframed as “societal needs.” Although Social Reconstruction is present in Illich’s thought (Kahn & Kellner, 2007), it operates largely at a meta-level, informing his critique of institutional manipulation and his call to restore education’s convivial foundations.

In what follows, we consider a case from the Russian educational context that illustrates the ambivalent status of AI as a New Technology situated between efficiency and emancipation, manipulation and conviviality, resistance and collective appropriation. Contemporary interpretations of Illich’s distinction between convivial and manipulative tools already extend it to digital technologies in education (Qutieshat, 2025; Beinsteiner, 2020). We interpret, in the context of this distinction, a specific educational experience of AI implementation and its organised reflection, both of which are described below.

231

Artificial Intelligence as an Educational Tool: The Case of the School of Advanced Studies

The educational model of the SAS, which serves as the empirical field of this study, was developed through an attempted integration of the Learner Centered Ideology, grounded in the values of liberal education and the cultivation of critical thinking, and the Scholar Academic Ideology, which sustains a high level of theoretical complexity within predominantly interdisciplinary curricula. In this respect, the model aspired to combine two of the four ideological orientations identified by M. Schiro. As in other attempts to integrate divergent forms of educational goal-setting within a single programme, the implementation of the SAS model revealed persistent tensions between students’ aspirations toward agency and the self-determination of learning content, and the relatively rigid logic of the academic canon, which prescribes the sequencing of knowledge according to disciplinary history and cultural norms.

This internal contradiction was also reflected in dominant instructional formats. Despite an explicit emphasis on students’ independent research and creative projects, the curriculum largely retained the classical lecture–seminar structure. The principles of liberal education were

realised primarily through the construction of individual educational trajectories. Since its founding, SAS has sought to integrate students into contemporary research agendas, with particular attention to interdisciplinary and transdisciplinary inquiry. The text-centred character of the model implied that assessment practices would prioritise slow reading, deep interpretation, reasoned discussion, and, crucially, the production of students' own intellectual work, above all in the form of academic essays.

The mass dissemination of LLMs in 2023 posed a serious challenge to the practices and normative assumptions that had taken shape at SAS over its first six years, as it did for many universities committed to the "grand humanistic project" of education (Knox, 2019; Biesta, 1998). Within the humanities, technologies had typically functioned as objects of critique rather than as actors capable of destabilising the "purely human" interaction between teacher and student as the presumed foundation of educational authenticity. The capacity of AI to generate texts indistinguishable from those written by humans directly challenged one of the School's core institutional principles, namely zero tolerance for plagiarism and other forms of academic dishonesty. More broadly, the advent of generative AI problematised writing as a mode of thinking and the text as the primary object and outcome of humanities education (Hsu, 2025).

232

Initial attempts to preserve the existing model included consideration of a return to pre-digital practices, such as oral examinations and supervised in-class essay writing. This "conservative turn," however, was ultimately rejected. Under the leadership of its director, Andrey Shcherbenok, the School chose not to pursue a prohibitive strategy but instead to actively integrate AI into teaching, seeking modes of use that would sustain a high level of students' cognitive labour while opening new institutional possibilities. This decision was further shaped by a structural constraint. In 2022, SAS faced a shortage of international English-speaking faculty, which made AI tools appear as a potential response to the deficit of expert instructors. If a world-class professor could not be invited, their expertise might instead be simulated through an AI analogue.

The central hypothesis guiding AI implementation at SAS was that "course personas", defined as customised LLM-based chatbots equipped with vector databases, could both substitute for certain functions traditionally performed by instructors and, through dialogical interaction, support students' autonomous knowledge construction by fostering research competencies, epistemic agency, and horizontal communication. To organise student activity, a mediator role was introduced. This role was assigned to individuals trained in pedagogy and educational design but not necessarily in the relevant discipline. This functional

separation between expertise and pedagogy, traditionally unified in the Humboldtian figure of the professor, necessitated a systematic analysis of teaching practices. In-depth interviews with SAS faculty were conducted to map competencies and classroom routines in order to identify which elements of teaching could be delegated to AI and which required sustained human involvement. In the resulting model, the chatbot functioned as a source of specialised knowledge and a navigator of disciplinary domains, providing feedback and exemplary solutions, while the mediator organised learning activities, supported motivation, and facilitated interpersonal communication. In early experiments, mediators were senior students and teaching assistants.

Experimental designs varied according to instructional goals, ranging from moderate configurations, in which instructors remained present and AI personas performed auxiliary functions, to more radical ones, where, due to the absence of a subject specialist, the educational process relied entirely on interaction between students, AI personas, and mediators. In total, ten experimental courses with varying degrees of AI integration were implemented at SAS in 2024 (Sound Studies, The World through Time, *Ecce Homo*, New Media, Foundations of the Humanities, Design Thinking, and others). Students were also formally permitted to use other LLMs, including GigaChat, YandexGPT, and ChatGPT.

233

The introduction of AI led to a substantial redesign of the courses: lectures were significantly reduced and largely replaced with more interactive learning formats that emphasized active student participation. At the same time, direct (“live”) interaction with expert instructors was sharply curtailed. The concept of the “ignorant schoolmaster” (Ranci re, 1991), embedded in the “AI plus mediator” model, did not — as anticipated — result in students’ intellectual emancipation or increased engagement. Instead, it intensified feelings of helplessness, contributed to declining motivation, and provoked passive resistance and, in some cases, open hostility toward both the mediators and the School’s administration. The need to understand students’ experiences within a new experimental educational context prompted the research team to investigate the factors influencing the acceptance of the new technology and the modified pedagogical design. Several hypotheses were formulated at the outset. The primary hypothesis was that the instrumental technologisation of education is fundamentally misaligned with the educational model and epistemic culture of SAS, in which the figure of the professor and the practice of personal interaction with them as a bearer of unique knowledge play a constitutive role. The removal of this figure was perceived as a form of “dehumanisation” of the educational process, while communication among peers and with the mediator failed to compensate for dialogue with a higher-status Knower, whose evaluation, praise, or censure carried greater social and existential significance. In people-

driven cultures, to which universities largely belong, AI-generated texts generally do not attain the status of knowledge and cannot compete with expert judgments or with canonical sources such as books and scholarly articles authored by humans. Moreover, elements of the academic curriculum that presuppose rigid hierarchies among epistemic agents do not readily accommodate the unproblematic introduction of competing non-human intellectual agents.

A second hypothesis did not directly concern technology. It was grounded in assumptions derived from cognitive pedagogy, according to which a shift toward a more agentic learner position, achieved through the intensification of constructivist teaching methods in which the primary role moves from teacher to student, may produce a range of negative effects under conditions of high cognitive task complexity and insufficiently developed internal guidance. These effects include declining motivation, a sense of disorientation, and even the reinforcement of educational inequality (Kirschner et al., 2006; Tharayil et al., 2018). Students' pre-existing beliefs about what an educational environment and learning process should entail, as well as how roles should be distributed within them, may also have significantly shaped their disappointment with the changes being introduced. Most students had prior experience primarily with instructivist schooling, in which the teacher provides a clear algorithm for completing tasks that can be evaluated as unequivocally correct or incorrect (Lee, 2022).

234

Methodology

One of the article's authors participated in developing the concept of AI integration at SAS, its theoretical justification, the design of experiments, and the facilitation of focus groups. This involvement provides an insider perspective, yet it may introduce interpretive bias, which we acknowledge as a limitation of the study. At the same time, reflexive engagement with our own experience, combined with privileged access to the tacit dimensions of communication and interaction among all participants, enables a more comprehensive and nuanced understanding of the phenomenon. It also facilitates critical examination of our own assumptions and preconceptions, thereby supporting the potential for constructive change (Dyson, 2007; Starr, 2010).

The methodological foundation of this study is a thematic analysis of 20 semi-structured interviews with second- to fourth-year students from the School of Advanced Studies (SAS) who participated in experimental courses involving AI personas (10 second-year, 5 third-year, and 5 fourth-year). The overrepresentation of second-year students in the sample resulted from most experimental courses being conducted with

second-year groups. Interviews, lasting 35–84 minutes, were fully transcribed and analyzed through multi-stage thematic analysis to identify key patterns in students' perceptions of AI integration.

The interviews were designed primarily to elicit students' reflective accounts of their educational experience with AI, with the practical aim of informing subsequent improvements in pedagogical practices involving LLMs. Accordingly, the interview guide consisted of relatively open-ended questions rather than narrowly targeted probes. The resulting key themes (practices of students' independent LLM use; ways AI was applied in class; students' attitudes toward mediators; perceptions of the professor's role; faculty attitudes toward AI implementation at SAS; epistemic authority of AI agents; evaluations of LLM-generated results; perceptions of new learning formats; and students' underlying beliefs about education) emerged inductively from the interview content.

These empirically derived themes are subsequently interpreted through the theoretical lens of Ivan Illich's distinction between convivial and manipulative tools, as well as Habermas's contrast between communicative and instrumental rationality. In this way, the theoretical framework serves as an interpretive device applied after the identification of patterns in the data, rather than as a pre-determined deductive structure guiding data collection. For the purposes of this article, we prioritize those narrative elements that most clearly illuminate the tension between instrumental and communicative logics of AI integration within the SAS educational model, as well as the conditions under which LLMs function (or fail to function) as convivial rather than manipulative tools in teaching and learning.

235

Results

When introducing new technologies, it is necessary to account both for the action potential embedded in the technology itself, conceptualised as affordance or the "spirit of technology" in the terms of G. DeSanctis and M. S. Poole (Poole & DeSanctis, 1990), and for the competencies of individual and collective actors who appropriate it. In particular, the dialogical nature of the course personas constitutes an important affordance that prescribes certain interaction scenarios, specifically requiring activity from the student in preparing prompts, maintaining dialogue, clarifying details, and so forth. At the same time, the academic and student-centered components of the program impose specific requirements on the new agent, which generated significant tension. On one hand, there are expectations regarding the content produced by the personas; it must be relevant, theoretically laden, accurate, and non-trivial to ensure the increment of new knowledge.

Through the RAG (Retrieval-Augmented Generation) architecture, focus on specific literature relevant to each individual course was ensured, and the model's "hallucinations" were eliminated. On the other hand, the structure of the AI personas should not be too rigid to allow for "creative" use that fosters the development of students' subjectivity rather than mere disciplinary indoctrination according to an externally imposed protocol.

One of the main problems encountered was the low evaluation by students of the generated responses: they were described as "useless", "empty", "general", "insufficiently deep", and so on. For example, "...the bots give absurd answers that do not help... it's like talking to a wall." (Student 2, Year 2). The semantic capital (Floridi, 2018) of such responses was rated significantly lower than that of the instructor's, even in cases where the two were virtually identical. Bias against chatbots as epistemic agents intensified in the absence of a human expert in the classroom. The significance of the instructor was justified in various ways; in particular, students noted specific competencies that, in their view, LLMs are incapable of performing qualitatively. The professor's possession of critical thinking skills, which he can teach to others unlike AI personas, the ability to set a theoretical framework for interpreting materials, offer one's own perspective, evaluate results, and guide students in their work were emphasized. Great emphasis was placed on the instructor's life experience: "the professor, unlike artificial intelligence, possesses some life experience that allows him to incorporate his own stories directly related to practice, to his experience, which AI does not possess" (Student 10, Year 3).

236

Attitudes toward mediators and their role in organizing communication and activity continued to be constructed in the logic of expertise, possibly because this position had not been encountered in the students' previous educational experience and was not sufficiently clearly positioned. The main complaint consisted in the lack of specialized knowledge and the refusal to confirm or refute the correctness of students' and chatbots' responses: "the mediator was not immersed in the subject himself, that is, he could not always correct these hallucinations" (Student 16, Year 2). One respondent described the experience of joint work as mutual "suffering" of mediators and students who became hostages of the situation. The low level of trust in mediators was influenced by their status, since a significant portion had just graduated from university or were final-year bachelor's students, that is, essentially peers of the students themselves.

Social influence, that is, the attitudes of instructors and other learners toward the implementation of AI technologies in SAS, was assessed by students ambiguously. According to most respondents, professors' positions were divided; some openly expressed their negative attitude

toward the use of chatbots, while some instructors were quite positive and perceived the experiments with interest and were ready to participate in them. One respondent described her professor's attitude as follows: "I cannot say that he was directly negative, but perhaps with some contempt." (Student 3, Year 4). Students' perceptions of AI implementation were equally heterogeneous, but the negative component appears quite frequently in the interviews. An indicative phenomenon was the emergence of group cohesion and even improvement in the quality of joint work, which arose against the background of passive resistance to the manipulative, imposed use of AI: "...there is a positive dynamic in that we somehow bonded with classmates... because we all disliked it, that is, we all understood that we were collectively suffering. And therefore we started to read everything so attentively, delve into and study, so that we had something to discuss ourselves, without touching this bot" (Student 11, Year 2).

The artificially imposed method of implementing AI technologies runs like a red thread through all interviews as the factor most traumatizing to students. Its use is characterized as "excessive", "involuntary", "persistent", "fostering a sense of powerlessness". This influenced the psychological state of the experiment participants: "constant mention of it and insistence on use sometimes leads to a certain fatigue from this, from the topic itself". (Student 15, Year 3). The manipulative nature of AI integration became an obvious threat to the manifestation of subjectivity and autonomy. Several respondents described their work in class with AI as mechanical, that is, devoid of personal meaning and not accompanied by understanding or appropriation of new knowledge. Students' disappointment was intensified by the realization of missed opportunities due to AI implementation, primarily related to the possibility of interaction with professors: "I have, well, a bit of this childish resentment that D came to us from Canada, such a cool professor, and they gave me some stupid machine that sometimes cannot even string two words together." (Student 5, Year 4).

To interpret students' acceptance or rejection of AI technologies more accurately, we compared their in-class interactions with "course personas" to their independent, uncoerced use of LLMs. Most students reported using AI with varying frequency for academic tasks. Tools such as Consensus, Elicit, and Semantic Scholar helped them locate relevant sources, edit texts in Russian and English, draft annotations, and similar activities. Instrumental use predominated, with effectiveness typically measured by time saved on uninteresting, routine tasks. Although some employed LLMs more constructively (as a "partner" in joint reflection, intellectual production, or research), such co-creative, convivial engagement usually involved interactivity and deliberate goal-setting focused on higher-quality outcomes rather than mere efficiency.

Despite varied approaches to LLMs, nearly all students viewed this independent experience positively or neutrally; it posed no threat to their autonomy or sense of meaning. We suggest that perceptions of the experimental courses with AI personas depended not only on tool quality and implementation mode but also on students' educational beliefs and expectations. Many respondents displayed a teacher-centered orientation characteristic of Scholar Academic Ideology, where knowledge transmission remains the professor's prerogative. In some interviews, commitment to this transmissional model was explicit: "why do I need education if it is not the transmission of knowledge? Some cool knowledge that I can remember." (Student 7, Year 2). The instructor was accorded an exclusive role in explaining and verifying knowledge; its absence provoked confusion and powerlessness. AI technologies clearly failed to substitute for this figure of epistemic authority and order, even with a mediator present. Efforts to foster solidarity proved inadequate to convince students that collective work could yield meaningful results.

238

Reflexive analysis of our practice, from the experiment designer's perspective, yields several conclusions. Beyond student interviews, these draw on interactions with SAS administration, faculty, and developers, as well as classroom work with course personas. A key insight was that introducing a new technological actor not only restructured established practices and redistributed power but also served as a litmus test exposing deep contradictions inherent in the model. Chief among these is the inevitable tension in combining Scholar Academic Ideology with Learner Centered Ideology within one programme. Transmitting "packaged" disciplinary knowledge within a strict epistemic hierarchy (the "banking model" (Freire, 2005)) is hard to reconcile with addressing each student's unique needs, personal meaning-making, and experiential relevance of knowledge. This inevitably sends mixed signals: "be autonomous" yet "follow strict rules". Divergences emerge in assessment (disciplinary compliance versus individual growth), content selection (canon versus personal interests), and teacher role (expert versus facilitator). Well-structured research and teaching activities could resolve this tension, enabling students to enter and advance in science alongside the teacher while preserving personal interest and self-directed trajectories.

Evaluating the experience of implementing AI tools in SAS, we see one of the main reasons for failures in the fundamental divergence between the idea of the possible liberating role of technology embedded in the pedagogical design of the courses and the manipulative way of realizing the experiments. The "course persona" was supposed to become the embodiment of communicative rationality and a kind of New Technology in Marcuse's sense: to give students the opportunity in a

dialogical mode to independently construct knowledge, taking into account personal values and interests, and to contribute to building a horizontal epistemic culture in which even recognized authorities remain transparent and open to criticism. At the same time, the organization of this process became the embodiment of the logic of efficiency and total manipulativity, in which both students and, to a certain extent, instructors became means for achieving external indicators, were deprived of choice, autonomy, and excluded from negotiations that would define acceptable boundaries of appropriateness and expediency of AI application for all actors. The instrumentalist approach to implementation led some students to perceive their classroom work as mere imitation, aimed at goals unrelated to students, instructors, mediators, or education itself. As one respondent explained, they opposed using AI in universities merely as a trend to secure funding, rather than creating something genuinely effective.

Where instrumentalization of the lifeworld occurs, individuals resist in varied ways, whether consciously and purposefully or spontaneously and unconsciously. We interpret certain student-described and observed practices as resistance to manipulative education: open sabotage of tasks, recourse to alternative information sources, “ironic” chatbot use (contrary to intended purpose), informal instructor contacts for expert human opinion, and occasional dysfunctional behaviour, including verbal aggression. The SAS AI implementation project faced resistance primarily due to inadequate communicative efforts: insufficient communication of the concept to faculty and students, limited discussion of their perspectives, and few adjustments based on feedback.

It would be erroneous to overlook, amid the failures, the positive experiences of AI application that nonetheless emerged and could serve as a foundation for future project development, enabling “course personas” to function as tools of conviviality rather than manipulation. Among the most productive was student involvement in creating chatbots as “doubles” of historical figures. This enabled appropriation of course goals, linkage of disciplinary knowledge to personal interests, and perception of chatbots not as alienated black boxes with rigid instructions but as entities warranting “care”: the desire to enhance the capabilities of this non-human other one has created (e.g., greater accuracy in reproducing thinking style or contextual awareness). When viewed as evolving agents with variable scenarios and flexibility, such chatbots integrate more harmoniously into constructivist pedagogical approaches.

Conclusion

The findings from this investigation into LLM implementation indicate limited success in their constructive integration into university educa-

tion; nonetheless, both the limitations encountered and the positive instances observed encourage reflection on the factors contributing to these results. A possible answer is provided by the conceptual frame, which simultaneously ensures the interpretation of the results. In the course of the study, an opposition was discovered between communicative rationality, which assumed free creative interaction of process participants and was explicitly present in out-of-class use of AI, and instrumental rationality subordinating to a logic of increasing efficiency in classes. In both cases, the same technology (with variations of specific LLMs) was used; it was a possible boundary object, but did it become one in reality? Students' attitudes showed the perception of the AI "course persona" as a means of manipulation (content-wise empty, not assuming creative self-expression in this use, not allowing self-limitation by other educational tactics). It was this perception that caused resistance. In the context of our conceptual frame, it can be hypothesized that the potential of the AI "course persona" as not only and not so much as a means of manipulation but as a tool for conviviality was not fully utilized. The turn to the second side was partially realized in the process of creating one's own chatbots, and it was this that received a positive response. Moreover, such convivial work, connecting in becoming students, teachers, and the AI course persona, could also work to link the two opposing ideologies—Scholar Academic and Learner Centered Ideologies. As shown in the theoretical part, they contradict each other if in the first case an educational hierarchy is realized with an unambiguously assumed epistemic authority based on expert knowledge, and in the second emphasis is placed on the student's creative independence in knowledge validation. However, as the studies showed, in the first sense students lacked transparency in the AI persona regarding the origin of its expertise (distinction from professors who have life experience, are capable of teaching and demonstrating critical thinking), and in the second they were frustrated precisely by their own independence, while the communicative component of creativity was reduced. That is, the component of the two ideologies that can work on their connection—joint creative self-expression, conviviality, which unites in one event the acquisition and transmission of knowledge, self-improvement and improvement of one's own tools—was missed (more precisely, not emphasized). In addition, for the actualization of AI as a tool for conviviality, it is important to abandon total manipulativity and criticize the myth of consensus. Therefore, the assumption that everyone perceives AI implementation with the same enthusiasm proved problematic; insufficient attention was paid to compromises and negotiations, clarification for each other of the roles of all interaction participants. In the end, neglect of the boundary nature of AI left the educational situation with a gap between two rationalities and two educational ideologies, with the totality of manipulation and underestimation of conviviality.

Благодарности / Acknowledgments

Asya A. Filatova acknowledges the support of this research by the Ministry of Science and Higher Education of the Russian Federation under agreement No. 075-03-2026-305 (January 16, 2026), associated with project “Applied Research on the Implementation of Artificial Intelligence Technologies in Higher Education” (project code: FSMG-2025-0086).

Список источников / References

Bearman M., Ryan J., & Ajjawi R. (2023). Discourses of artificial intelligence in higher education: a critical literature review. *Higher Education*, 86(2), pp. 369-385. <https://doi.org/10.1007/s10734-022-00937-2>

Beinsteiner A. (2020). Conviviality, the internet, and AI. Ivan Illich, Bernard Stiegler, and the question concerning information-technological self-limitation. *Open Cultural Studies*, 4(1), pp. 131-142. <https://doi.org/10.1515/culture-2020-0013>

Biesta G. (1998). Pedagogy without humanism: Foucault and the subject of education. *Interchange*, 29(1), pp. 1-16.

Compagnucci L., & Spigarelli F. (2020). The Third Mission of the university: A systematic literature review on potentials and constraints. *Technological Forecasting and Social Change*, 161, p. 120284. <https://doi.org/10.1016/j.techfore.2020.120284>.

Dyson M. (2007). My Story in a Profession of Stories: Auto Ethnography — an Empowering Methodology for Educators. *Australian Journal of Teacher Education*, 32(1), pp. 36-48. <https://doi.org/10.14221/ajte.2007v32n1.3>.

Floridi L. (2018). Semantic Capital: Its Nature, Value, and Curation. *Philosophy & Technology*, 31, pp. 481-497. <https://doi.org/10.1007/s13347-018-0335-1>.

Fowler D. S. (2023). AI in Higher Education. *Journal of Ethics in Higher Education*, 3, pp. 127-143. <https://doi.org/10.26034/fr.jehe.2023.4657>.

Freire P. (2005). *Pedagogy of the oppressed*. 30th anniversary ed. New York: The Continuum International Publishing Group.

Habermas J. (1986). Technology and Science as “Ideology”. In *Toward a Rational Society: Student Protest, Science, and Politics*. Cambridge: Polity. pp. 81-122.

Hsu H. (2025). What happens after A. I. destroys college writing? The demise of the English paper will end a long intellectual tradition, but it’s also an opportunity to reexamine the purpose of higher education. *The New Yorker*. <https://www.newyorker.com/culture/cultural-comment/what-happens-after-ai-destroys-college-writing> Accessed 15 December 2025.

Illich I. (1971). *Deschooling society*. New York: Harper & Row.

Illich I. (1973). *Tools for conviviality*. Glasgow: Calder & Boyars.

Illich I. (1977). *Disabling professions*. London: Marion Boyars.

Kahn R., & Kellner D. (2007). Paulo Freire and Ivan Illich: technology, politics and the reconstruction of education. *Policy Futures in Education*, 5(4). <https://doi.org/10.2304/pfie.2007.5.4.431>

Kirschner P. A., Sweller J., & Clark R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), pp.75–86.

Knox J. (2019). What Does the ‘Postdigital’ Mean for Education? Three Critical Perspectives on the Digital, with Implications for Educational Research and Practice. *Postdigital Science and Education*, 1, pp. 357–370. <https://doi.org/10.1007/s42438-019-00045-y>

Lee S. J., & Branch R. M. (2022). Students’ Reactions to a Student-Centered Learning Environment in Relation to Their Beliefs about Teaching and Learning. *International Journal of Teaching and Learning in Higher Education*, 33(3), pp. 298–305.

Lindsay J., & Jacka L. (2024). The Footsteps on the Sands of AI for Higher Education: Moving Beyond Ad-Hoc. *Journal of Ethics in Higher Education*, 5, pp. 51–77. <https://doi.org/10.26034/fr.jehe.2024.6863>

Pandey Ch. Sh., Patanjali M., Pandey Sh., et al. (2025). Epistemic trust in generative AI for higher education scale (ETGAI-HE scale). *AI & Society*. <https://doi.org/10.1007/s00146-025-02566-6>

Poole M. S., & DeSanctis G. (1990). Understanding the use of Group Decision Support Systems: The Theory of Adaptive Structuration. In *Organizations and Communication Technology*, edited by Janet Fulk and Charles Steinfield, Thousand Oaks: SAGE Publications. pp. 173–193. <https://doi.org/10.4135/9781483325385.n8>

Qutieshat A. (2025). *Artificial Intelligence in Higher Education: A Contemporary Examination of Illich’s Theories*. Singapore: Springer.

Rancière J. (1991). *The ignorant schoolmaster: Five lessons in intellectual emancipation* (K. Ross, Trans.). Stanford University Press.

Schiro M. (2013). *Curriculum Theory: Conflicting Visions and Enduring Concerns*. 2nd ed. Thousand Oaks: SAGE Publications.

Sembey R., Hoda R., & Grundy J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software*, 211, p. 111988. <https://doi.org/10.1016/j.jss.2024.111988>.

Starr L. J. (2010). The Use of Autoethnography in Educational Research: Locating Who We Are in What We Do. *Canadian Journal for New Scholars in Education*, 3(1).

Tharayil S., Borrego M., Prince M., et al. (2018). Strategies to mitigate student resistance to active learning. *International Journal of STEM Education*, 5, p. 7. <https://doi.org/10.1186/s40594-018-0102-y>

Tilak Sh., & Glassman M. (2020). Alternative lifeworlds on the Internet: Habermas and democratic distance education. *Distance Education*, 41(3), pp. 326–344. <https://doi.org/10.1080/01587919.2020.1763782>

Т-университеты (трансформирующиеся университеты). (2019). СКОЛКОВО. https://www.skolkovo.ru/public/media/documents/research/sedec/SKOLKOVO_S... (Дата обращения: 10.12.2025)

— T-university (transformiruyushchiesya university). (2019). Skolkovo. https://www.skolkovo.ru/public/media/documents/research/sedec/SKOLKOVO_S... Accessed 10 December 2025. (in Russ.)

Об авторах / About the authors

Шиповалова Лада Владимировна — доктор философских наук, профессор Центра практической философии «Стасис». Европейский университет в Санкт-Петербурге, Санкт-Петербург, Российская Федерация. Научные интересы: философия науки, историческая эпистемология, социальная эпистемология, исследования публичной научной коммуникации.

<https://orcid.org/0000-0002-1989-1152>. E-mail: lshipovalova@eu.spb.ru

Lada Shipovalova — Doctor of Philosophical Sciences, Professor of Stasis Center for Practical Philosophy. European University at St. Petersburg, St. Petersburg, Russian Federation. Research Interests: philosophy of science, historical epistemology, social epistemology, public communication of science studies.

<https://orcid.org/0000-0002-1989-1152>. E-mail: lshipovalova@eu.spb.ru

Филатова Ася Алексеевна — кандидат философских наук, старший научный сотрудник Центра прикладных лингвистических исследований и тестирования «ИСТОК». Московский физико-технический институт, Москва, Российская Федерация. Научные интересы: социальная эпистемология, исследования науки и технологий (STS), философия искусственного интеллекта, социальное восприятие технологий, исследование учебных программ

243

<https://orcid.org/0000-0002-0497-0018>. E-mail: filatova.aa@mipt.ru

Filatova Asya — Candidate of Philosophical Sciences, Senior Researcher of the Center of Applied Linguistics Research and Testing “ISTOK”. Moscow Institute of Physics and Technology (MIPT), Moscow, Russian Federation. Research interests: social epistemology, science and technology studies (STS), philosophy of artificial intelligence, social perception of technology, curriculum studies

<https://orcid.org/0000-0002-0497-0018>. E-mail: filatova.aa@mipt.ru