

## GENERATIV INTELLEKT MATNLARIDAGI GALLYUTSINATSIYA HODISASINING KOMPYUTER-LINGVISTIK VA SEMANTIK TAHLILI

**To'ychiyev Abror Xabibulla o'g'li**

Renessans ta'lim universiteti

Assistent o'qituvchisi

[Tuychiyev13@gmail.com](mailto:Tuychiyev13@gmail.com)

+998932963378

**Annotatsiya.** Ushbu maqolada generativ sun'iy intellekt (LLM) modellari tomonidan yaratiladigan matnlarda uchraydigan gallyutsinatsiya hodisasi kompyuter-lingvistik va semantik nuqtai nazardan tahlil qilinadi. Dasturiy modellar yaratgan matnlar ko'p hollarda grammatik jihatdan to'g'ri, ravon va sintaktik tugallangan bo'lsa-da, ularda faktologik va semantik buzilishlar yuzaga keladi. Tadqiqotda gallyutsinatsiyalarning leksik-semantik, propozitsional, referensial va konseptual darajalardagi ko'rinishlari atroflicha o'rganilgan. Kompyuter lingvistikasidagi zamonaviy entropiya va semantik vektorli modellarga tayanib, matndagi mazmuniy uzilishlar va semantik gallyutsinatsiyalarni aniqlash hamda ularni baholash mezonlari ishlab chiqilgan.

**Kalit so'zlar:** generativ sun'iy intellekt, katta til modellari (LLM), gallyutsinatsiya, kompyuter lingvistikasi, semantik tahlil, propozitsiya, referentsiya, kogerentlik, semantik entropiya.

## КОМПЬЮТЕРНО-ЛИНГВИСТИЧЕСКИЙ И СЕМАНТИЧЕСКИЙ АНАЛИЗ ФЕНОМЕНА ГАЛЛЮЦИНАЦИЙ В ТЕКСТАХ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

**Аннотация.** В данной статье рассматривается феномен галлюцинаций в текстах, генерируемых моделями искусственного интеллекта (LLM), с точки зрения компьютерной лингвистики и семантики.

*Хотя создаваемые тексты в большинстве случаев грамматически правильны и беглы, в них наблюдаются фактологические и семантические искажения. Исследуются проявления галлюцинаций на лексико-семантическом, пропозициональном, референциальном и концептуальном уровнях.*

**Ключевые слова:** *генеративный искусственный интеллект, большие языковые модели, галлюцинация, компьютерная лингвистика, семантический анализ, пропозиция, референция, когерентность.*

### **COMPUTATIONAL-LINGUISTIC AND SEMANTIC ANALYSIS OF HALLUCINATION PHENOMENON IN GENERATIVE ARTIFICIAL INTELLIGENCE TEXTS**

**Abstract.** *This article analyzes the hallucination phenomenon in texts generated by generative artificial intelligence (LLM) from a computational-linguistic and semantic perspective. Although the generated texts are mostly grammatically correct and fluent, factual and semantic distortions occur. The study investigates hallucinations across lexical-semantic, propositional, referential, and conceptual levels.*

**Keywords:** *generative AI, large language models (LLM), hallucination, computational linguistics, semantic analysis, proposition, reference, coherence, semantic entropy.*

**Kirish.** *Generativ sun'iy intellekt va katta til modellari (LLM – Large Language Models) so'nggi yillarda matn yaratish va tabiiy tilni qayta ishlash (NLP) sohasida mislsiz yutuqlarga erishdi. GPT-4, Llama, Claude kabi zamonaviy modellar grammatik jihatdan mukammal, sintaktik va uslubiy jihatdan ravon matnlarni shakllantirish qobiliyatiga ega. Biroq kompyuter lingvistikasi va tabiiy til semantikasi oldida turgan eng dolzarb muammolardan biri – bu neyron tarmoqlar tomonidan yaratilgan matnlardagi gallyutsinatsiya (hallucination) hodisasidir.*

*Gallyutsinatsiya – bu model tomonidan haqiqatga mos kelmaydigan,*

mantiqan asossiz, ob'ektiv voqelikda mavjud bo'lmagan faktlarni yoki semantik jihatdan ziddiyatli mazmunni qat'iylik va ravonlik bilan taqdim etilishi hodisasidir. Dastlabki bosqichlarda generativ modellar faqat sintaktik xatoliklar yoki grammatik nomuvozanatlik bergan bo'lsa, zamonaviy modellar sirtqi sintaktik va kogeziyaviy izchillikni saqlagani holda, chuqur semantik sathda ziddiyatli va yolg'on axborotlarni shakllantirmoqda.

Ushbu tadqiqotning maqsadi generativ sun'iy intellekt matnlaridagi gallyutsinatsiya hodisasini kompyuter-lingvistik va semantik tahlil doirasida o'rganish, uning leksik, propozitsional, referensial hamda matn-kogerentlik sathlarida namoyon bo'lish mexanizmlarini va ularni baholash prinsiplarini ochib berishdan iborat.

**Material va metodlar.** Tadqiqotda generativ sun'iy intellekt (GPT-4o, Claude 3.5 Sonnet, Llama 3) modellari tomonidan yaratilgan turli janrdagi (ilmiy, ommabop, tahliliy) matnlar hamda kompyuter lingvistikasining zamonaviy empirik ma'lumotlar bazalari tahlil ob'ekti sifatida olindi. Tadqiqot metodologiyasi quyidagi kompleks yondashuvlarni o'z ichiga oladi:

- Kompyuter-lingvistik tahlil: Neyron tarmoqlarning ehtimollik taqsimoti, ehtimollik entropiyasi (semantic entropy) hamda matn komponentlarining vektorli ifodalarini (word/sentence embeddings) tahlil qilish.
- Propozitsional-semantik tahlil: Matndagi har bir gap va hukm ifodalagan propozitsional tuzilmalarni ajratib olish va ularning haqiqiylik qiymatini (truth-value) hamda logiko-semantik izchilligini baholash.
- Referensial va kogerentlik tahlili: Matndagi olmoshlar, otlar va tushunchalarning haqiqiy referentlarga to'g'ri bag'ishlanganligi hamda abzatslararo ma'no rivojlanishini qiyosiy tekshirish.

**Natijalar.** O'tkazilgan kompyuter-lingvistik va semantik tahlillar generativ matnlardagi gallyutsinatsiya hodisasi bir necha asosiy sathlarda namoyon bo'lishini ko'rsatdi:

1. Leksik-semantik va konseptual siljish: Model tushunchalarning semantik chegaralarini matn davomida saqlab qolmaydi. Masalan, matn 'generativ sun'iy

intellekt' haqida boshlanib, keyingi gaplarda 'algoritmik modellar' va 'neyron tarmoqlar' tushunchalariga o'tadi. garchand bu tushunchalar yaqin bo'lsa-da, ularning semantik hajmi (extension) bir xil emas. Natijada konseptual chegaralarning yuvilishi va noaniqligi yuzaga keladi.

2. Propozitsional va mantiqiy uzilishlar: Sintaktik jihatdan to'g'ri va kirish so'zlari (demak, shu sababli, natijada) bilan bog'langan gaplar o'rtasida mantiqiy-semantik sababiyat mavjud bo'lmaydi. Masalan, 'Alisher Navoiy XV asrda yashagan. Uning ijodi turkiy adabiyot taraqqiyotiga katta ta'sir ko'rsatgan. Demak, XV asrda turkiy til butun Markaziy Osiyoning yagona rasmiy tili bo'lgan.' gaplaridagi uchinchi propozitsiya dastlabki ikkisidan mantiqan kelib chiqmaydi.

3. Referensial noaniqlik va almashinuv: Matndagi olmosh va anaphorik birliklarning muayyan denotat yoki referentga bog'lanishida xatoliklar yuzaga keladi. Model kontekstdagi obyektlarni chalkashtirib qo'yadi va natijada matn o'quvchisi uchun mazmuniy noaniqlik tug'iladi.

4. Informatsion tautologiya va psevdoyangi axborot: Model sintaktik konstruksiyalarni o'zgartirgan holda bir xil semantik propozitsiyani qayta-qayta takrorlaydi. Bu holatda matn hajmi kengayadi, biroq uning informatsion entropiyasi (yangi axborot hajmi) ortmaydi.

Kompyuter-lingvistik nuqtai nazardan gallyutsinatsiya generativ modellar keyingi so'z yoki token ehtimolligini hisoblashda eng yuqori ehtimollikka ega bo'lgan, lekin mazmunan noto'g'ri bo'lgan vektorni tanlashi natijasida yuzaga keladi. Semantik entropiya (Farquhar et al., 2024) ko'rsatkichi yuqori bo'lgan nuqtalarda gallyutsinatsiya darajasi keskin ortishi empirik tarzda tasdiqlandi.

1-jadval. Generativ matnlardagi gallyutsinatsiyalarning semantik va lingvistik tipologiyasi

| Gallyutsinatsiya darajasi | Lingvistik va semantik belgilari                                                    | Yuzaga kelish sababi / Mexanizmi                                                     |
|---------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| Leksik-semantik           | Tushunchalar va terminlar semantik doirasining o'zgarishi, tushunchalar almashinuvi | Vektorlar fazosidagi leksik yaqinlik va tushunchalar chegarasining yuvilishi         |
| Propozitsional            | Hukmlar va gaplar o'rtasidagi sabab-oqiyat va mantiqiy aloqalarning uzilishi        | Ehtimollik taqsimotida mantiqiy bog'liqlikdan ko'ra sintaktik ravonlikning ustunligi |
| Referensial               | Olmoshlar va anaphorik vositalarning noto'g'ri obyekt yoki shaxsga yo'naltirilishi  | Uzoq kontekstli xotirada (context window) referentlar indeksatsiyasining yo'qolishi  |
| Matn / Kogerentlik        | Matn boshlanishi va xulosasi o'rtasidagi ziddiyat, informatsion tautologiya         | Generatsiya jarayonida dastlabki tezisning global nazorat qilinmasligi               |

**Xulosa.** Generativ sun'iy intellekt matnlarida gallyutsinatsiya hodisasini o'rganish shuni ko'rsatadiki, matnning grammatik to'g'riligi va sintaktik ravonligi uning semantik va faktologik mukammalligini kafolatlamaydi. Model sirtqi kogeziyani saqlagan holda, chuqur semantik va propozitsional ziddiyatlarni yuzaga keltirishi mumkin.

Gallyutsinatsiyalarni faqat faktologik xatolar sifatida emas, balki kompyuter-lingvistik va semantik sathdagi tizimli hodisa sifatida baholash lozim. Ularni kamaytirish uchun modellarni o'qitishda faqat sintaktik va ehtimollik mezonlaridan tashqari, semantik entropiya, propozitsional mantiq va referensial izchillikni nazorat qiluvchi lingvistik modullarni joriy etish talab etiladi.

Kelgusi tadqiqotlar o'zbek tili korpuslari va lingvistik resurslari bazasida generativ modellarning gallyutsinatsiya darajasini avtomatik baholash va filtrlash algoritmlarini yaratishga qaratilishi lozim.

**Foydalanilgan adabiyotlar:**

1. Søgaard A. Do Language Models Have Semantics? On the Five Standard Positions // Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics. – Vienna : Association for Computational Linguistics, 2025. – P. 25910–25922. – DOI: 10.18653/v1/2025.acl-long.1258.
2. Chen H., Yang X., Zhu J., Wang W. Quantifying Semantic Emergence in Language Models // Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics. – Vienna : Association for Computational Linguistics, 2025. – P. 12041–12054. – DOI: 10.18653/v1/2025.acl-long.588.
3. Meconi D., Stirpe S., Martelli F., Lavallo L., Navigli R. Do Large Language Models Understand Word Senses? // Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. – Suzhou : Association for Computational Linguistics, 2025. – P. 33897–33916. – DOI: 10.18653/v1/2025.emnlp-main.1720.
4. Saglam B., Kassianik P., Nelson B., Weerawardhena S., Singer Y., Karbasi A. Large Language Models Encode Semantics and Alignment in Linearly Separable Representations // Proceedings of the 14th International Joint Conference on Natural Language Processing. – Mumbai : Association for Computational Linguistics, 2025. – P. 2282–2303. – DOI: 10.18653/v1/2025.ijcnlp-long.124.
5. Diera A., Scherp A. Do Language Models Encode Semantic Relations? Probing and Sparse Feature Analysis // Proceedings of LREC 2026. – ELRA Language Resource Association, 2026. – P. 2115–2126. – DOI: 10.63317/52znrwwkdog7.
6. Cassotti P., Baes N., De Pascale S., de Sá J. M. C., Periti F., Haslam N., Geeraerts D., Tahmasebi N. SenseRel: A Sense-Level Benchmark for Denotational and Connotational Meaning Relations // Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics. – San Diego : Association for Computational Linguistics, 2026. – P. 499–515. – DOI: 10.18653/v1/2026.acl-long.20.
7. Farquhar S., Kossen J., Kuhn L., Gal Y. Detecting Hallucinations in Large Language Models Using Semantic Entropy // Nature. 2024. Vol. 630. – P. 625–630.