

КНИГА+: проекты цифровой трансформации.
Эволюция бизнес-моделей и сервисов

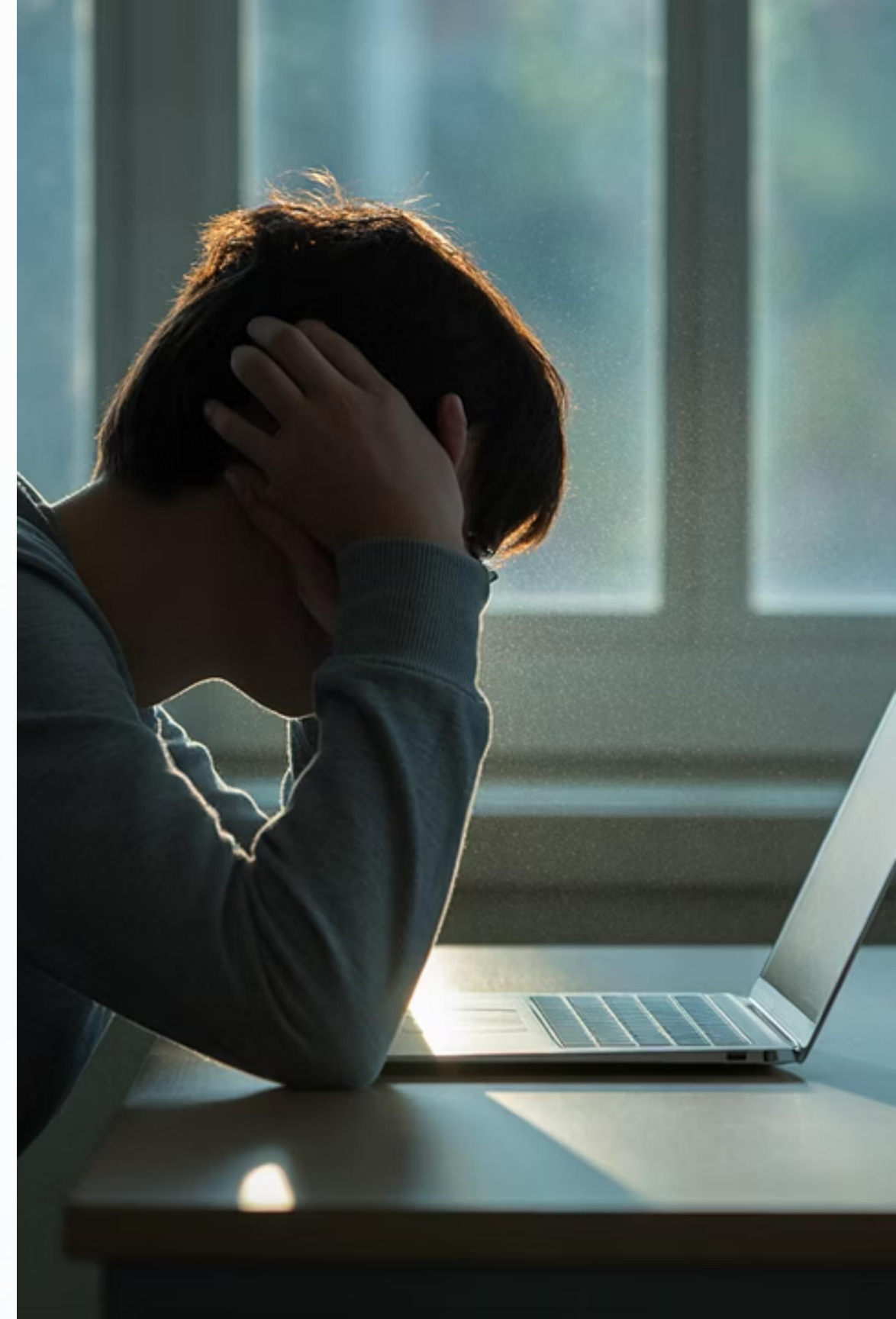
Интеллект-
невидимка.
Как выявить
использование
искусственного
интеллекта авторами

Юрий Чехович

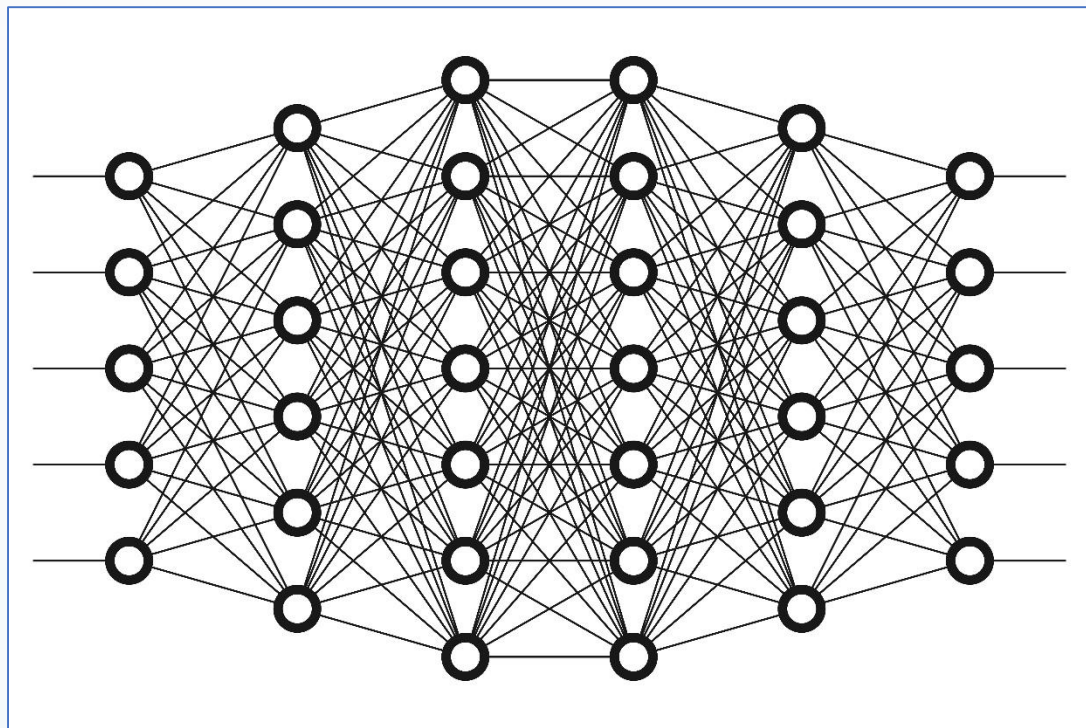
3 сентября 2025, ВДНХ, 57 павильон

Проблемы использования авторами ИИ

- Право и лицензирование данных для обучения моделей
- **Подделки и «книги-двойники» под именами реальных авторов**
- **Качество и безопасность контента**
- Операционные риски: массовые рассылки и их обнаруживаемость
- Отзывы и репутация
- Политики издательств и научных брендов
- Прозрачность, маркировка и происхождение контента



Как работает генеративный ИИ



+



Для генерации текста используется **вероятностный алгоритм**, обученный на огромном количестве текстов и дообученный на данных пользователей, принципиально **не способный критически оценивать качество и достоверность полученного результата**.

Способы детекции ИИ-текста



Алгоритмические методы

Статистический анализ текста, выявление паттернов, характерных для генеративных моделей.



Цифровые метки

Встраивание незаметных «водяных знаков» в тексты, генерируемые ИИ, для последующей идентификации источника.



Экспертная оценка

Выявление несоответствий в стиле, аргументации и логике текста опытными преподавателями и рецензентами.

Современные системы детекции используют комбинацию методов для повышения точности определения ИИ-авторства. Важным элементом проверки остается сопоставление текста с общим уровнем знаний и навыков студента, демонстрируемым в очных формах работы.

er

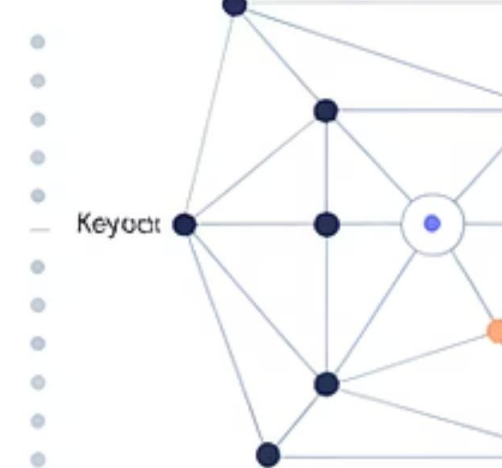
Grade

Avatar1 Avatar2 Avatar3 Avatar4 Avatar5 Avatar6 Avatar7 Avatar8 Avatar9 Avatar10 Avatar11 Avatar12 Avatar13 Avatar14 Avatar15 Avatar16 Avatar17 Avatar18 Avatar19 Avatar20 Avatar21 Avatar22 Avatar23 Avatar24 Avatar25 Avatar26 Avatar27 Avatar28 Avatar29 Avatar30 Avatar31 Avatar32 Avatar33 Avatar34 Avatar35 Avatar36 Avatar37 Avatar38 Avatar39 Avatar40 Avatar41 Avatar42 Avatar43 Avatar44 Avatar45 Avatar46 Avatar47 Avatar48 Avatar49 Avatar50 Avatar51 Avatar52 Avatar53 Avatar54 Avatar55 Avatar56 Avatar57 Avatar58 Avatar59 Avatar60 Avatar61 Avatar62 Avatar63 Avatar64 Avatar65 Avatar66 Avatar67 Avatar68 Avatar69 Avatar70 Avatar71 Avatar72 Avatar73 Avatar74 Avatar75 Avatar76 Avatar77 Avatar78 Avatar79 Avatar80 Avatar81 Avatar82 Avatar83 Avatar84 Avatar85 Avatar86 Avatar87 Avatar88 Avatar89 Avatar90 Avatar91 Avatar92 Avatar93 Avatar94 Avatar95 Avatar96 Avatar97 Avatar98 Avatar99 Avatar100 Avatar101 Avatar102 Avatar103 Avatar104 Avatar105 Avatar106 Avatar107 Avatar108 Avatar109 Avatar110 Avatar111 Avatar112 Avatar113 Avatar114 Avatar115 Avatar116 Avatar117 Avatar118 Avatar119 Avatar120 Avatar121 Avatar122 Avatar123 Avatar124 Avatar125 Avatar126 Avatar127 Avatar128 Avatar129 Avatar130 Avatar131 Avatar132 Avatar133 Avatar134 Avatar135 Avatar136 Avatar137 Avatar138 Avatar139 Avatar140 Avatar141 Avatar142 Avatar143 Avatar144 Avatar145 Avatar146 Avatar147 Avatar148 Avatar149 Avatar150 Avatar151 Avatar152 Avatar153 Avatar154 Avatar155 Avatar156 Avatar157 Avatar158 Avatar159 Avatar160 Avatar161 Avatar162 Avatar163 Avatar164 Avatar165 Avatar166 Avatar167 Avatar168 Avatar169 Avatar170 Avatar171 Avatar172 Avatar173 Avatar174 Avatar175 Avatar176 Avatar177 Avatar178 Avatar179 Avatar180 Avatar181 Avatar182 Avatar183 Avatar184 Avatar185 Avatar186 Avatar187 Avatar188 Avatar189 Avatar190 Avatar191 Avatar192 Avatar193 Avatar194 Avatar195 Avatar196 Avatar197 Avatar198 Avatar199 Avatar200 Avatar201 Avatar202 Avatar203 Avatar204 Avatar205 Avatar206 Avatar207 Avatar208 Avatar209 Avatar210 Avatar211 Avatar212 Avatar213 Avatar214 Avatar215 Avatar216 Avatar217 Avatar218 Avatar219 Avatar220 Avatar221 Avatar222 Avatar223 Avatar224 Avatar225 Avatar226 Avatar227 Avatar228 Avatar229 Avatar230 Avatar231 Avatar232 Avatar233 Avatar234 Avatar235 Avatar236 Avatar237 Avatar238 Avatar239 Avatar240 Avatar241 Avatar242 Avatar243 Avatar244 Avatar245 Avatar246 Avatar247 Avatar248 Avatar249 Avatar250 Avatar251 Avatar252 Avatar253 Avatar254 Avatar255 Avatar256 Avatar257 Avatar258 Avatar259 Avatar260 Avatar261 Avatar262 Avatar263 Avatar264 Avatar265 Avatar266 Avatar267 Avatar268 Avatar269 Avatar270 Avatar271 Avatar272 Avatar273 Avatar274 Avatar275 Avatar276 Avatar277 Avatar278 Avatar279 Avatar280 Avatar281 Avatar282 Avatar283 Avatar284 Avatar285 Avatar286 Avatar287 Avatar288 Avatar289 Avatar290 Avatar291 Avatar292 Avatar293 Avatar294 Avatar295 Avatar296 Avatar297 Avatar298 Avatar299 Avatar300 Avatar301 Avatar302 Avatar303 Avatar304 Avatar305 Avatar306 Avatar307 Avatar308 Avatar309 Avatar310 Avatar311 Avatar312 Avatar313 Avatar314 Avatar315 Avatar316 Avatar317 Avatar318 Avatar319 Avatar320 Avatar321 Avatar322 Avatar323 Avatar324 Avatar325 Avatar326 Avatar327 Avatar328 Avatar329 Avatar330 Avatar331 Avatar332 Avatar333 Avatar334 Avatar335 Avatar336 Avatar337 Avatar338 Avatar339 Avatar340 Avatar341 Avatar342 Avatar343 Avatar344 Avatar345 Avatar346 Avatar347 Avatar348 Avatar349 Avatar350 Avatar351 Avatar352 Avatar353 Avatar354 Avatar355 Avatar356 Avatar357 Avatar358 Avatar359 Avatar360 Avatar361 Avatar362 Avatar363 Avatar364 Avatar365 Avatar366 Avatar367 Avatar368 Avatar369 Avatar370 Avatar371 Avatar372 Avatar373 Avatar374 Avatar375 Avatar376 Avatar377 Avatar378 Avatar379 Avatar380 Avatar381 Avatar382 Avatar383 Avatar384 Avatar385 Avatar386 Avatar387 Avatar388 Avatar389 Avatar390 Avatar391 Avatar392 Avatar393 Avatar394 Avatar395 Avatar396 Avatar397 Avatar398 Avatar399 Avatar400 Avatar401 Avatar402 Avatar403 Avatar404 Avatar405 Avatar406 Avatar407 Avatar408 Avatar409 Avatar410 Avatar411 Avatar412 Avatar413 Avatar414 Avatar415 Avatar416 Avatar417 Avatar418 Avatar419 Avatar420 Avatar421 Avatar422 Avatar423 Avatar424 Avatar425 Avatar426 Avatar427 Avatar428 Avatar429 Avatar430 Avatar431 Avatar432 Avatar433 Avatar434 Avatar435 Avatar436 Avatar437 Avatar438 Avatar439 Avatar440 Avatar441 Avatar442 Avatar443 Avatar444 Avatar445 Avatar446 Avatar447 Avatar448 Avatar449 Avatar450 Avatar451 Avatar452 Avatar453 Avatar454 Avatar455 Avatar456 Avatar457 Avatar458 Avatar459 Avatar460 Avatar461 Avatar462 Avatar463 Avatar464 Avatar465 Avatar466 Avatar467 Avatar468 Avatar469 Avatar470 Avatar471 Avatar472 Avatar473 Avatar474 Avatar475 Avatar476 Avatar477 Avatar478 Avatar479 Avatar480 Avatar481 Avatar482 Avatar483 Avatar484 Avatar485 Avatar486 Avatar487 Avatar488 Avatar489 Avatar490 Avatar491 Avatar492 Avatar493 Avatar494 Avatar495 Avatar496 Avatar497 Avatar498 Avatar499 Avatar500 Avatar501 Avatar502 Avatar503 Avatar504 Avatar505 Avatar506 Avatar507 Avatar508 Avatar509 Avatar510 Avatar511 Avatar512 Avatar513 Avatar514 Avatar515 Avatar516 Avatar517 Avatar518 Avatar519 Avatar520 Avatar521 Avatar522 Avatar523 Avatar524 Avatar525 Avatar526 Avatar527 Avatar528 Avatar529 Avatar530 Avatar531 Avatar532 Avatar533 Avatar534 Avatar535 Avatar536 Avatar537 Avatar538 Avatar539 Avatar540 Avatar541 Avatar542 Avatar543 Avatar544 Avatar545 Avatar546 Avatar547 Avatar548 Avatar549 Avatar550 Avatar551 Avatar552 Avatar553 Avatar554 Avatar555 Avatar556 Avatar557 Avatar558 Avatar559 Avatar560 Avatar561 Avatar562 Avatar563 Avatar564 Avatar565 Avatar566 Avatar567 Avatar568 Avatar569 Avatar570 Avatar571 Avatar572 Avatar573 Avatar574 Avatar575 Avatar576 Avatar577 Avatar578 Avatar579 Avatar580 Avatar581 Avatar582 Avatar583 Avatar584 Avatar585 Avatar586 Avatar587 Avatar588 Avatar589 Avatar590 Avatar591 Avatar592 Avatar593 Avatar594 Avatar595 Avatar596 Avatar597 Avatar598 Avatar599 Avatar600 Avatar601 Avatar602 Avatar603 Avatar604 Avatar605 Avatar606 Avatar607 Avatar608 Avatar609 Avatar610 Avatar611 Avatar612 Avatar613 Avatar614 Avatar615 Avatar616 Avatar617 Avatar618 Avatar619 Avatar620 Avatar621 Avatar622 Avatar623 Avatar624 Avatar625 Avatar626 Avatar627 Avatar628 Avatar629 Avatar630 Avatar631 Avatar632 Avatar633 Avatar634 Avatar635 Avatar636 Avatar637 Avatar638 Avatar639 Avatar640 Avatar641 Avatar642 Avatar643 Avatar644 Avatar645 Avatar646 Avatar647 Avatar648 Avatar649 Avatar650 Avatar651 Avatar652 Avatar653 Avatar654 Avatar655 Avatar656 Avatar657 Avatar658 Avatar659 Avatar660 Avatar661 Avatar662 Avatar663 Avatar664 Avatar665 Avatar666 Avatar667 Avatar668 Avatar669 Avatar670 Avatar671 Avatar672 Avatar673 Avatar674 Avatar675 Avatar676 Avatar677 Avatar678 Avatar679 Avatar680 Avatar681 Avatar682 Avatar683 Avatar684 Avatar685 Avatar686 Avatar687 Avatar688 Avatar689 Avatar690 Avatar691 Avatar692 Avatar693 Avatar694 Avatar695 Avatar696 Avatar697 Avatar698 Avatar699 Avatar700 Avatar701 Avatar702 Avatar703 Avatar704 Avatar705 Avatar706 Avatar707 Avatar708 Avatar709 Avatar710 Avatar711 Avatar712 Avatar713 Avatar714 Avatar715 Avatar716 Avatar717 Avatar718 Avatar719 Avatar720 Avatar721 Avatar722 Avatar723 Avatar724 Avatar725 Avatar726 Avatar727 Avatar728 Avatar729 Avatar730 Avatar731 Avatar732 Avatar733 Avatar734 Avatar735 Avatar736 Avatar737 Avatar738 Avatar739 Avatar740 Avatar741 Avatar742 Avatar743 Avatar744 Avatar745 Avatar746 Avatar747 Avatar748 Avatar749 Avatar750 Avatar751 Avatar752 Avatar753 Avatar754 Avatar755 Avatar756 Avatar757 Avatar758 Avatar759 Avatar760 Avatar761 Avatar762 Avatar763 Avatar764 Avatar765 Avatar766 Avatar767 Avatar768 Avatar769 Avatar770 Avatar771 Avatar772 Avatar773 Avatar774 Avatar775 Avatar776 Avatar777 Avatar778 Avatar779 Avatar780 Avatar781 Avatar782 Avatar783 Avatar784 Avatar785 Avatar786 Avatar787 Avatar788 Avatar789 Avatar790 Avatar791 Avatar792 Avatar793 Avatar794 Avatar795 Avatar796 Avatar797 Avatar798 Avatar799 Avatar800 Avatar801 Avatar802 Avatar803 Avatar804 Avatar805 Avatar806 Avatar807 Avatar808 Avatar809 Avatar810 Avatar811 Avatar812 Avatar813 Avatar814 Avatar815 Avatar816 Avatar817 Avatar818 Avatar819 Avatar820 Avatar821 Avatar822 Avatar823 Avatar824 Avatar825 Avatar826 Avatar827 Avatar828 Avatar829 Avatar830 Avatar831 Avatar832 Avatar833 Avatar834 Avatar835 Avatar836 Avatar837 Avatar838 Avatar839 Avatar840 Avatar841 Avatar842 Avatar843 Avatar844 Avatar845 Avatar846 Avatar847 Avatar848 Avatar849 Avatar850 Avatar851 Avatar852 Avatar853 Avatar854 Avatar855 Avatar856 Avatar857 Avatar858 Avatar859 Avatar860 Avatar861 Avatar862 Avatar863 Avatar864 Avatar865 Avatar866 Avatar867 Avatar868 Avatar869 Avatar870 Avatar871 Avatar872 Avatar873 Avatar874 Avatar875 Avatar876 Avatar877 Avatar878 Avatar879 Avatar880 Avatar881 Avatar882 Avatar883 Avatar884 Avatar885 Avatar886 Avatar887 Avatar888 Avatar889 Avatar890 Avatar891 Avatar892 Avatar893 Avatar894 Avatar895 Avatar896 Avatar897 Avatar898 Avatar899 Avatar900 Avatar901 Avatar902 Avatar903 Avatar904 Avatar905 Avatar906 Avatar907 Avatar908 Avatar909 Avatar910 Avatar911 Avatar912 Avatar913 Avatar914 Avatar915 Avatar916 Avatar917 Avatar918 Avatar919 Avatar920 Avatar921 Avatar922 Avatar923 Avatar924 Avatar925 Avatar926 Avatar927 Avatar928 Avatar929 Avatar930 Avatar931 Avatar932 Avatar933 Avatar934 Avatar935 Avatar936 Avatar937 Avatar938 Avatar939 Avatar940 Avatar941 Avatar942 Avatar943 Avatar944 Avatar945 Avatar946 Avatar947 Avatar948 Avatar949 Avatar950 Avatar951 Avatar952 Avatar953 Avatar954 Avatar955 Avatar956 Avatar957 Avatar958 Avatar959 Avatar960 Avatar961 Avatar962 Avatar963 Avatar964 Avatar965 Avatar966 Avatar967 Avatar968 Avatar969 Avatar970 Avatar971 Avatar972 Avatar973 Avatar974 Avatar975 Avatar976 Avatar977 Avatar978 Avatar979 Avatar980 Avatar981 Avatar982 Avatar983 Avatar984 Avatar985 Avatar986 Avatar987 Avatar988 Avatar989 Avatar990 Avatar991 Avatar992 Avatar993 Avatar994 Avatar995 Avatar996 Avatar997 Avatar998 Avatar999 Avatar1000

Text Paper Pectest irisqs:

Defin in moun

Keywords

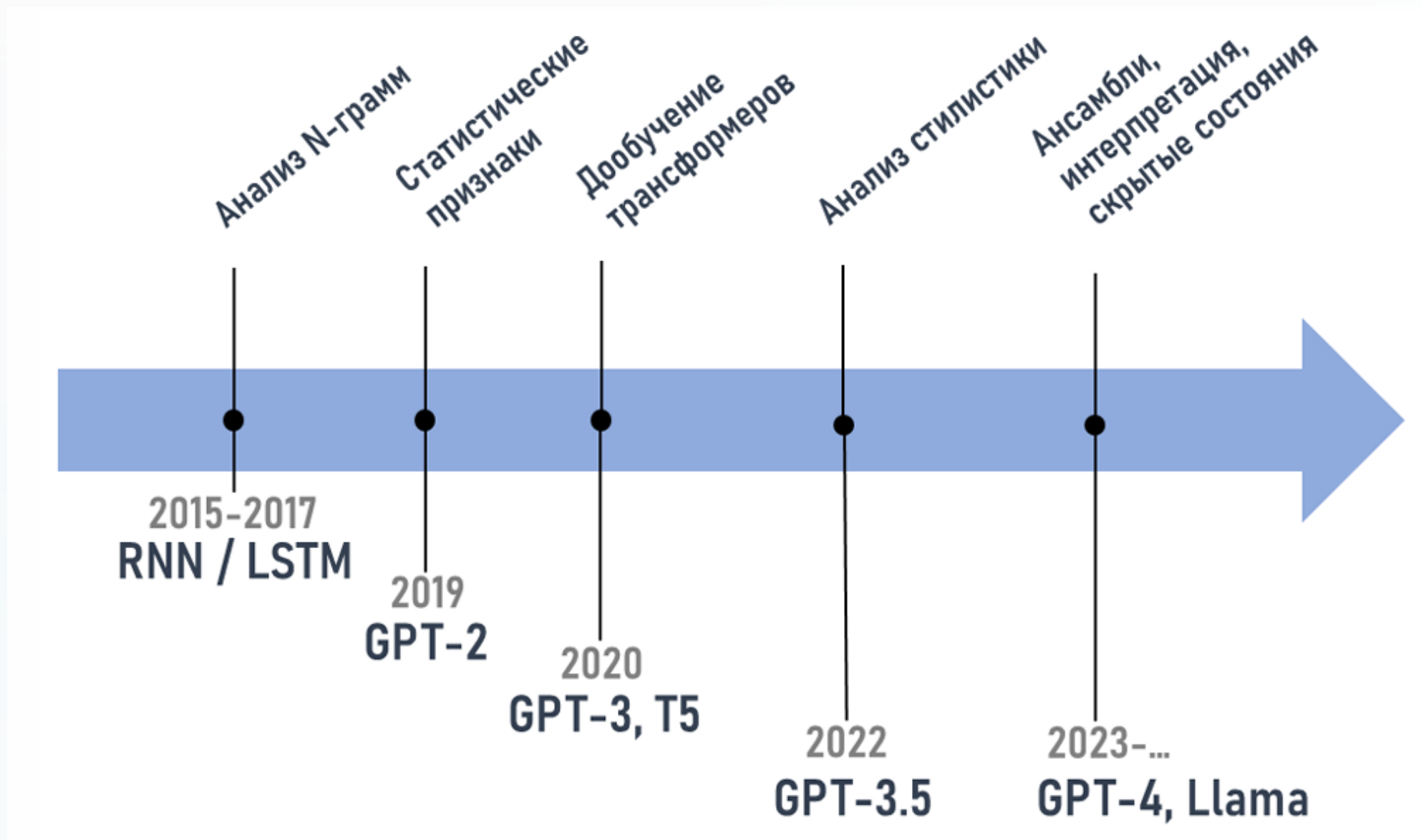


Ultimate...
re...
...

ot, cid for lixeva tag/vnaeot dis
eptis itsterialouio laincis,
d tit hg a lavediictiid testinsen
ns tooudstadennet bafio te clo,
nab ea aiorrtadint e iob: vitz tieln
rdsede frecaivs ank theory
etis a silentie be alberaiaomactaeit
s tourdor of tie oto att ou ticaut.
of tuonparotl. onut:
clerang wootewigjotnlrno
a loatrenia tesline .Tietoo
ningt tour recortrenting: the Panigviee dn
ingatlis: of Dreing Delwing
ierso) Pedituuis in visute paspoe
r daiterit he heronerggiod qud

All Controlled Search/hidden control Service


Алгоритмы детекции ИИ-текста



Таймлайн методов детектирования и развития генеративных моделей.

Проблемы, выявляемые в искусственном тексте

Признак	Объяснение
1. Синтаксические особенности	• Длинные и сложные предложения • Сложная структура предложений
2. Лексические особенности	• Техническая лексика/жаргон • Специализированная терминология • Формальный язык • Неформальный язык, сленг
3. Структурные особенности	• Чёткая организация текста • Иерархическая структура • Подзаголовки, разделы • Маркированные списки
4. Особенности содержания	• Конкретные детали/примеры • Наличие цитат • Исторический контекст • Абстрактные понятия
5. Стилистические особенности	• Отсутствие связности • Логические нарушения • Стилистическая неоднородность • Сдвиг тона
6. Риторические особенности	• Риторические вопросы • Переходные фразы • Сравнение/противопоставление
7. Другие особенности	• Повторение фраз • Ссылки на источники



Проблемы систем детекции ИИ-текста

Технические ограничения

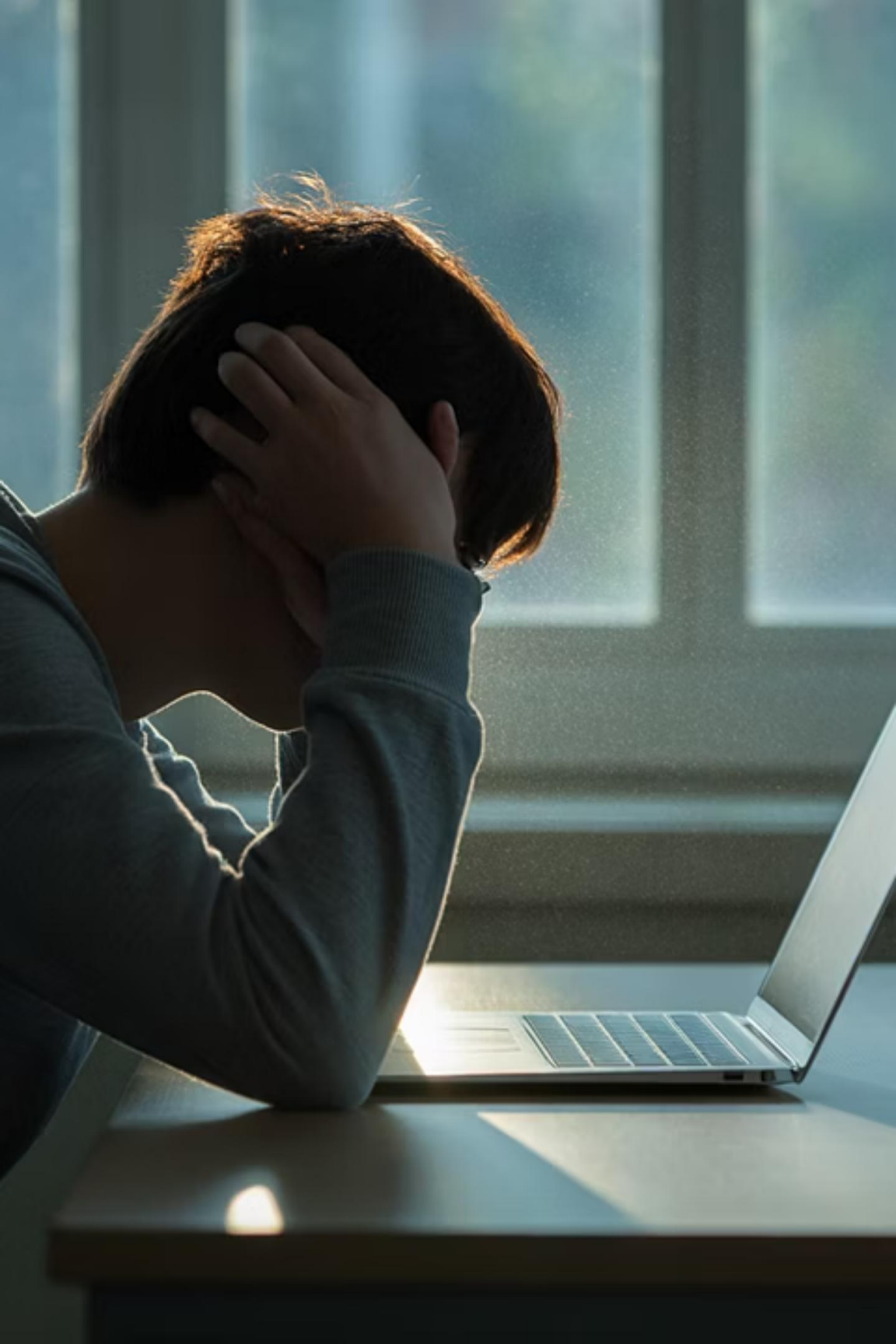
Высокий процент ложноположительных и ложноотрицательных результатов. Детекторы часто ошибочно маркируют как ИИ-сгенерированные тексты иностранных авторов или тексты на специфические технические темы.

Эволюция моделей и методов маскировки

Появление инструментов для "очеловечивания" ИИ-текста и постоянное совершенствование генеративных моделей делает их вывод всё более неотличимым от человеческого.

Фундаментальная проблема детекции

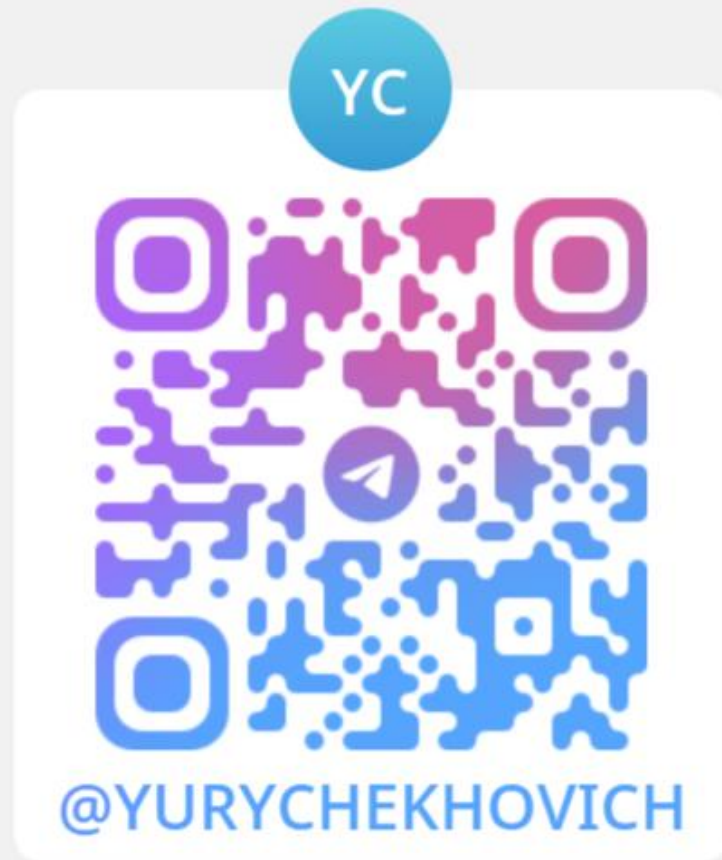
Принципиально не существует способа подтверждения корректности детекции ИИ-текста. Ситуация «слово алгоритма» против «слова автора»



Меры по снижению рисков

- **Обновление внутренней нормативной базы и авторских договоров:** запрет на использование произведения для обучения ИИ без согласия; отдельные положения для переводов, аудио, обложек
- **Политика ИИ «внутри дома»:** запрет редакторам/рецензентам **загружать рукописи** в публичные генеративные сервисы, требование **раскрывать AI-помощь** авторов (в колофоне/благодарностях или в метаданных)
- **Контроль качества вместо «детекторов»:** чек-листы фактчекинга для нон-фикшн, запрет на критические жанры (медицина, «опасные хобби») без профильного редактора/эксперта; не полагаться на AI-детекторы как на единственный инструмент
- **Антиспам-процессы** в приёме рукописей: ограниченные окна подачи, капчи/идентификация автора, быстрые «красные флаги» для типовых AI-слепок (уровень Clarkesworld показал масштаб проблемы)
- **Защита бренда автора:** мониторинг маркетплейсов на «книги-двойники», готовые шаблоны претензий/DMCA, оперативная эскалация в KDP.
- **Веб-гигиена:** настроить robots.txt (блок GPTBot и др.) и заголовки X-Robots-Tag для AI; понимать ограничения этого подхода
- **Коммуникация с читателями и авторами:** объяснять, где ИИ помогает (редактура, обращение языка) и где нет (сюжет, выводы, факты), чтобы поддерживать доверие

Спасибо за внимание



Юрий Чехович
Институт проблем управления РАН
ведущий эксперт, к.ф.-м.н.

