

## Wstęp

---

*Change by design or change by disaster*

Na współczesnym etapie rozwoju technologii cyfrowych potrzebujemy wypracowania nowych paradygmatów metodologicznych związanych z badaniem istoty, sposobów wytwarzania oraz właściwości kontentu medialnego. Wynika to nie tylko z rekonfiguracji miejsca, które we współczesnej komunikacji zajmuje język, i rodzących się na tym tle „wyzwań multimodalności” (Kress 2016), lecz przede wszystkim z faktu, że do grona agentów komunikacji coraz częściej są włączane generatywne modele AI, których użycia nie da się sprowadzić jedynie do instrumentalnego poszerzenia ludzkich możliwości. O ile jeszcze trzy lata temu agentywność AI – w szczególności dużych modeli językowych (Large Language Models, LLM) – wydawała się całkiem umowna (Andreas 2022), o tyle dziś ich status w ekosystemie komunikacyjnym uległ radykalnej zmianie. Przykładowo generatywne modele neuronowe typu ChatGPT Agent są w stanie autonomicznie inicjować wykorzystanie narzędzi do rozwiązywania złożonych zadań, w tym zadań multimodalnych obejmujących przetwarzanie tekstu, dźwięku i obrazu.

Przejawy agentywności i intencjonalności systemów generatywnej AI wymagają refleksji nie tyle technologicznej, ile humanistycznej, w tym przede wszystkim rewizji formuły „twórca znaku – znak”. Jak zauważa Marcus Tomalin (2025), gdy zautomatyzowany system generuje treść na podstawie polecenia (promptu) użytkownika, intencjonalność człowieka jako autora zapytania nie uczestniczy bezpośrednio w samym

procesie wytwarzania. Prompt (ang. *prompt* – komunikat wejściowy aktywujący mechanizm generatywny modelu AI i wyznaczający parametry odpowiedzi w ramach zadanego problemu) określa specyfikację projektu wypowiedzi. Jednak realizacji tego projektu dokonuje system generatywnej AI, który funkcjonuje jako aktywny agent, samodzielnie dobierający nie tylko tryby reprezentacji, lecz także statystyczne wzorce korpusu, na których jest tworzona odpowiedź.

Pojęcie designu (projektowania) jest tu rozumiane – za Guntherem Kressem – jako proces świadomego i celowego (a więc powiązane go z ideą agentywności/sprawczości) wyboru i zestawienia zasobów służących „projektowaniu sensu” (Kress 2016: 88). W tym kontekście inżynieria multimodalna, obejmująca cały zespół praktyk kształtowania sensów realizowanych we współczesnej cyfrowej przestrzeni informacyjno-komunikacyjnej, wchodzi na nowy poziom złożoności i wymaga wnikliwej uwagi badawczej.

W oddawanej do rąk Czytelnika monografii kierujemy tę uwagę m.in. na:

- (1) ogólne wyzwania multimodalności w epoce „przebudzenia AI” (*AI awakening*), gdy do grona agentów komunikacji coraz częściej są włączane generatywne modele sztucznej inteligencji. Temu jest poświęcony pierwszy rozdział monografii. Koncentrujemy się w nim na humanistycznych trajektoriach inżynierii multimodalnej, kształtujących się na tle ewolucji technologii medialnych oraz poszerzania repertuaru środków oddziaływania na układ zmysłowo-percepcyjny – sensorium człowieka. W tym kontekście – za Kressem – praktyki inżynierii multimodalnej są ujmowane jako formy działalności semiotycznej, rozumianej w kategoriach (re)designu/(re)projektowania treści multimodalnych, natomiast kategorie agentywności/sprawczości i potencjału reprezentacyjnego zasobów znakowych zostają przemyślane na nowo z uwzględnieniem pojawienia się nowych aktorów oraz nowych form dla treści multimodalnych powstających we współczesnym cyfrowym

środowisku medialnym. Szczególna uwaga jest poświęcona opisowi zasad działania i podatności na zakłócenia generatywnej AI, w tym wyjaśnieniu przyczyn powstawania halucynacji oraz zniekształceń kulturowych w generowanym przez AI kontencie multimodalnym. Ponadto analizowane są manipulacyjne praktyki nowego typu, w tym LLM grooming – kondycjonowanie promptami modelu LLM (systematyczne ustawianie zachowania modelu przez sekwencje komunikatów, aby utrwalać pożądane wzorce odpowiedzi) oraz instytucjonalne formy cenzury, których głównym obiektem staje się nie odbiorca – człowiek, lecz model AI, celowo dostrajany do transmitowania zadanych narracji. Na tym tle wprowadza się pojęcie nowej kompetencji medialnej, obejmującej kompetencje informacyjną, medialną, multimodalną oraz kompetencję AI (*AI literacy*) jako wzajemnie powiązane umiejętności niezbędne do odpowiedzialnego i krytycznego korzystania z praktyk inżynierii multimodalnej w warunkach współczesnej przestrzeni medialnej;

- (2) konkretne praktyki konstruowania kontentu multimedialnego, w tym praktyki konstruowania immersyjnych przestrzeni wirtualnych w ramach relaksacyjnego segmentu mediów cyfrowych (na przykładzie m.in. kontentu medytacji uważności, gatunku afirmacji, dyskursu OS – Oddly Satisfying, ASMR – Autonomous Sensory Meridian Response i in.) oraz praktyki konstruowania multimodalnej przestrzeni paratekstowej na platformach streamingowych w ramach rozrywkowego segmentu nowych mediów (na przykładzie konstruowania gatunku opisu filmowego).

Praktyki pierwszego typu są obiektem rozważań w rozdziale drugim. Instytucjonalizacja i medializacja praktyk refleksyjności i spowolnienia staje się wielkoskalowym procesem przeciwdziałającym przyspieszeniu społecznemu. Praktyki medytacyjne, afirmacyjne, jak również kontemplacja nagrań ASMR wpisują się w trend spokojnego życia, którego filarami są idee minimalizmu, wellness i slow life. W rozdziale tym opisujemy nowy paradygmat społeczno-kulturowy zwany rewolucją uważności,

przejawiający się wieloaspektowo w obszarze aplikacji multimedialnych i obszernych zasobów internetowych, a także analizujemy techniki inżynierii multimodalnej w sferze mediów relaksacyjnych. Z perspektywy lingwistyki mediów omawiamy mechanizmy tworzenia immersyjnych przestrzeni wirtualnych i doświadczeń, w których obrazy, słowa, szepty czy też dźwięki niefonematyczne stają się nośnikami pragmatycznie określonych relaksacyjnych treści multisensorycznych i wywoływaczami reakcji emocjonalnych oraz psychosomatycznych. Cyfrowy dyskurs medytacji czy ASMR to przyjazne środowisko immersyjne, do którego adresat jest niejako „wciągany” i przez empatyczne oglądanie i słuchanie rekonstruuje quasisensoryczne obrazy bez udziału bezpośredniego bodźca zewnętrznego. Wszystkie elementy semiotyczne biorące udział w procesie konceptualizacji dyskursywnej i konstruowaniu tekstów są podporządkowane globalnej strategii polegającej na wyciszeniu i harmonizacji wewnętrznego stanu człowieka. Strategie fatyczności, alterocentryczności, a także pozytywnej fascynacji i estetyzacji przejawiają się na wszystkich etapach inżynierii relaksacyjnych tekstów medialnych. Jednocześnie odejście od dominacji reprezentacji abstrakcyjnych, symbolicznych na rzecz komponentów sensorycznych określa na nowo technologie mediów konwergentnych, których użytkownicy szukają sposobów na zaspokajanie potrzeby bliskości i cielesności w społeczeństwie informacyjnym (co w pierwszej dekadzie funkcjonowania mediów cyfrowych nie było priorytetem). W miarę rozwoju technologii obserwujemy znaczne rozszerzenie funkcjonalności mediów – od prezentacji treści i telekomunikacji do konstruowania przestrzeni komunikacji inter- i intrapersonalnej, będącej obecnie swoistą substytucją kontaktów fizjologicznych i doświadczeń sensorycznych.

W rozdziale trzecim przyglądamy się nowym sposobom komunikowania się z odbiorcą (wizualnym/użytkownikiem/subskrybentem), które zostały opracowane przez nowego gracza na rynku kontentu rozrywkowego – platformy streamingowe. Opisujemy fenomen serwisów VOD (video on demand) przez pryzmat innowacji, które zdecydowały

o sukcesie nowego medium i wyznaczyły nową jakość w dostarczaniu treści audiowizualnych, przechodząc następnie do dyskursu platform streamingowych i jego specyfiki. Jedną z najważniejszych cech tego dyskursu jest jego charakter paratekstowy, został on bowiem zbudowany – po stronie nadawcy – prawie w całości z tekstów eskortujących. We współczesnej przestrzeni informacyjnej, przesyconej narracjami i widowiskami, akt wstępnej recepcji (przed-odbioru) ma duże znaczenie, ale szczególną wagę zyskuje w środowisku nowego medium, które jest biblioteką filmów i seriali. Platformy budują swój dyskurs na podstawie tekstów, których głównym zadaniem jest ułatwienie wyboru w gęstej siatce katalogu. Czyniąc to, bazują na już istniejących wzorcach gatunkowych połączonych w multimodalne kolekcje, które dostosowują do swojej przestrzeni dyskursywnej, odpowiednio je strukturyzując i hierarchizując. W rozdziale opisujemy zarówno zmiany wzorców gatunkowych uwarunkowane zanurzeniem w dyskursie streamingu, jak i ich role oraz funkcje w rozbudowanym systemie multimodalnych komunikacji, m.in. w kreowaniu hipotezy interpretacyjnej i kształtowaniu procesu decyzyjnego odbiorcy.

Stosowane współcześnie szerokie spektrum narzędzi multimodalnej kreacji, wielość koncepcji badawczych zaczerpniętych z różnych dyscyplin naukowych zajmujących się komunikacją medialną, trudna do zbadania, płynna i zmieniająca się materia mediów cyfrowych wymagały od recenzentów wielkiego nakładu pracy, wnikliwości i cierpliwości. Wyrazy szczególnej wdzięczności kierujemy do Pana Profesora Bogusława Skowronka oraz Pani Profesor Jolanty Lubochy-Kruglik za rzeczowe, wysoce kompetentne zaopiniowanie projektu monografii.

Mamy nadzieję, że przygotowana przez nas publikacja spotka się z Państwa życzliwym przyjęciem i stanie się przyczynkiem, a jednocześnie zachętą do dalszego zgłębiania zagadnień inżynierii multimodalnej we współczesnych mediach.



## Введение



*Change by design or change by disaster.*

Современная стадия развития цифровых технологий требует выработки новых методологических парадигм исследования природы, способов создания и свойств медиаконтента. Это связано не только с реконfigurацией места, занимаемого языком в современной коммуникации, и возникающими на этом фоне «вызовами мультимодальности» (Kress 2016) – но, прежде всего, с тем, что сегодня в число агентов коммуникации всё чаще вовлекаются генеративные модели ИИ, использование которых нельзя свести лишь к инструментальному расширению возможностей человека. Если еще три года назад агентивность ИИ – в частности, больших языковых моделей (Large Language Models, LLM) – казалась весьма условной (Andreas 2022), то сегодня их статус в коммуникативной экосистеме радикально изменился. Так, нейросетевые модели наподобие ChatGPT Agent демонстрируют способность автономно инициировать использование инструментов для решения сложных задач, в том числе мультимодальных, связанных с обработкой текста, аудио, изображения.

Проявления агентивности и интенциональности систем генеративного ИИ требуют не столько технологической, сколько гуманитарной рефлексии – в частности, пересмотра формулы «создатель знака – знак». Как отмечает Маркус Томалин (Tomalin 2025), когда автоматизированная интеллектуальная система генерирует контент по запросу (промпту) пользователя, интенциональность

человека как автора запроса напрямую не задействована в этом процессе. Промпт (от англ. *prompt* — входное сообщение, активирующее генеративный механизм ИИ-модели и задающее параметры формируемого ответа в рамках поставленной задачи) определяет спецификацию дизайна сообщения. Однако реализуется этот дизайн системой генеративного ИИ, которая функционирует как активный агент, самостоятельно выбирающий не только режимы репрезентации, но и статистические паттерны корпуса, на которых строится ответ.

Дизайн (*design*) здесь, вслед за Г. Крессом, понимается как осознанный и целенаправленный, то есть связанный с идеей агентивности, процесс выбора и комбинирования ресурсов «проектирования смысла» (Kress 2016:88). На этом фоне мультимодальный инжиниринг, охватывающий весь комплекс практик смыслообразования, реализуемых в современном цифровом информационно-коммуникационном пространстве, выходит на новый уровень сложности и требует пристального исследовательского внимания.

В данной монографии рассматриваются следующие вопросы: (1) общие вызовы мультимодальности в эпоху «пробуждения ИИ» (*AI awakening*), когда генеративные модели искусственного интеллекта всё активнее включаются в ряды агентов коммуникации. Этому посвящена первая глава монографии. Мы фокусируемся в ней на гуманитарных траекториях мультимодального инжиниринга, формирующихся на фоне эволюции медиатехнологий и расширения репертуара средств воздействия на человеческий сенсориум. В этом контексте, вслед за Г. Крессом, практики мультимодального инжиниринга трактуются как формы семиотической работы, понимаемой в терминах (ре-)дизайна мультимодального контента, а категории агентивности и репрезентационного потенциала знаковых ресурсов переосмысляются с учетом появления новых акторов и новых

форм мультимодального контента, возникающих в современной цифровой медиасреде. Особое внимание уделяется описанию принципов работы и уязвимостей генеративного ИИ, включая описание причин возникновения «галлюцинаций» и культурных искажений в генерируемом ИИ мультимодальном контенте. Кроме того, рассматриваются манипулятивные практики нового типа, включая LLM grooming и институциональные формы цензуры, основным объектом которых оказывается не адресат-человек, а ИИ-модель, целенаправленно «настраиваемая» на трансляцию заданных нарративов. На этом фоне вводится понятие «новой грамотности», включающей информационную, медийную, мультимодальную и ИИ-грамотность как взаимосвязанные компетенции, необходимые для ответственного и критического использования практик мультимодального инжиниринга в условиях современного медиапространства;

- (2) конкретные практики конструирования мультимедийного контента, включая: практики создания иммерсивных виртуальных пространств в релаксационном сегменте цифровых медиа (на примере в т.ч. контента медитаций осознанности, жанра аффирмаций, дискурс OS – Oddly Satisfying, ASMR – Autonomous Sensory Meridian Response и др.) и практики создания мультимодальных паратекстовых пространств на стриминговых платформах в развлекательном сегменте новых медиа (на примере конструирования жанра киноописаний).

Первый тип практик рассматривается во второй главе. Институционализация и медиатизация практик рефлексивности и замедления становится масштабным процессом, противодействующим «социальному ускорению». Практики медитаций и аффирмаций, а также созерцание роликов ASMR соответствуют тренду размеренной жизни, столпами которого являются идеи минимализма, веллнесс и неспешности. В связи с этим в фокус внимания помещается новая социокультурная

парадигма, называемая «революцией осознанности», которая многогранно проявляется в сфере мультимедийных приложений и обширных онлайн-ресурсов, а также анализируются методы мультимодального инжиниринга в сфере релаксационных медиа. С позиций медиалингвистики описываются механизмы создания иммерсивных виртуальных пространств и переживаний, в которых изображения, слова, шепот и нефонематические звуки становятся носителями прагматически обусловленного мультисенсорного релаксационного контента и триггерами эмоциональных и психосоматических реакций. Цифровой дискурс медитации, или ASMR, представляет собой дружественную иммерсивную среду, в которую реципиент «втягивается» и посредством эмпатического созерцания и слушания реконструирует квазисенсорные образы без участия прямого внешнего стимула. Все семиотические элементы, участвующие в процессе дискурсивной концептуализации и конструирования текста, подчинены глобальной стратегии успокоения и гармонизации внутреннего состояния человека. Стратегии фатичности, альтероцентричности, позитивной фасцинации и эстетизации проявляются на всех этапах создания релаксационных медиатекстов. В то же время переход от доминирования абстрактных, символических представлений к сенсорным компонентам по-новому определяет конвергентные медиатехнологии, пользователи которых ищут способы удовлетворить потребность в близости и телесности в информационном обществе (что не было приоритетом в первое десятилетие существования цифровых медиа). По мере развития технологий мы наблюдаем значительное расширение функциональности медиа – от представления содержания и телекоммуникаций до создания пространств для межличностного и внутриличностного общения, которые теперь могут быть субститутами физиологических контактов и сенсорно осязаемых переживаний.

Третья глава посвящена способам коммуникации с адресатом (зрителем/ пользователем/ подписчиком), которые были разработаны новым игроком на рынке развлекательного контента – онлайн-кинотеатрами (стриминговыми платформами). Описывается феномен сервисов VOD (video on demand) через призму инноваций, которые определили успех новых медиа и задали новый стандарт в доставке аудиовизуального контента. Анализируется также дискурс стриминговых платформ и его специфика, в частности паратекстуальный характер этого дискурса, построенного – со стороны адресанта – практически полностью на сопроводительных текстах. В современном информационном пространстве, насыщенном нарративами и зрелищами, акт первичной рецепции (пререцепции) становится особенно важным в новой медиасреде, представляющей собой библиотеки фильмов и сериалов. Онлайн-кинотеатры строят свой дискурс на основе текстов, основная цель которых – облегчить выбор в плотной сетке кинокаталога. При этом они опираются на существующие жанровые образцы, объединенные в мультимодальные коллекции, которые адаптируют к своему дискурсивному пространству, структурируя и иерархизируя его соответствующим образом. В главе описываются как изменения жанровых образцов, обусловленные погружением в стриминговый дискурс, так и их роль и функции в сложной системе мультимодальной коммуникации, в т.ч. в создании интерпретативных гипотез и процессе принятия решений реципиентом.

Широкий спектр современных инструментов мультимодального инжиниринга, применяемых сегодня, многообразие исследовательских концепций, заимствованных из различных научных дисциплин, занимающихся медиакоммуникацией, а также текучая и постоянно меняющаяся материя цифровых медиа потребовали от рецензентов значительных усилий, глубокой вдумчивости и терпения. Мы хотели бы выразить особую благодарность профессору Богуславу Сковронеку и профессор Йоланте Любохе-Круглик за

содержательное и высокопрофессиональное рецензирование проекта монографии.

Авторы надеются, что предлагаемая вниманию читателя книга будет принята благосклонно и послужит вкладом и одновременно стимулом для дальнейшего изучения вопросов мультимодального инжиниринга в современных медиа.