

A Professional Investigative Portal for Manipulation, Institutional Dynamics, and Evidence: IA, UX, Legal, and Implementation Blueprint

Introduction and Objectives

Mission and scope This project will design and deliver a professional, modern investigative portal that organizes and presents nine Polish source files, preserving every file and its full content without removal or alteration while adding metadata, navigation, and contextual layers that improve comprehension and ethical communication. The site will present clear categories, predictable navigation, accessible audio for each article, and strong provenance to earn reader trust on sensitive topics of manipulation, toxicity, and institutional dynamics ^[1]. The homepage and category structure will foreground purpose and priority user actions with visible, high-contrast navigation to minimize interaction costs and maintain credibility for investigative reading ^[2].

Taxonomy and thematic categories To strengthen information scent and support readers' mental models, we will organize the corpus using a controlled taxonomy and allow polyhierarchy (a document can appear in multiple thematic contexts while maintaining one canonical path). This ensures consistent retrieval and narrative coherence across the portal ^[3]. The five categories are:

- **Manipulacje i Toksyczność:** psychology of manipulation, toxic relationships, narcissism; this lens surfaces patterns of coercion and harm in an analytical, trauma-informed tone ^[1].
- **Instrumentalne wykorzystanie instytucji:** materials showing how institutions supported or amplified manipulation; this category contextualizes systemic dynamics alongside individual actors ^[3].
- **Dokumenty i Fakty:** key facts, official documents, evidence appraisals, and factual briefs; this category anchors claims in verifiable artifacts and normalized metadata ^[3].
- **Raporty i Analizy:** media-facing reports and critical analyses; this category consolidates interpretive work and cross-links to primary evidence ^[3].
- **Symbolika i Narracja:** "Klątwa," chronicles, and context-building materials that articulate the broader psychological and cultural framing ^[3].

This taxonomy clarifies scope and provenance while avoiding sensationalism, which is essential for reader trust and ethical investigative communication ^[4].

Corpus to be integrated (verbatim file list and category positioning) All nine files will be preserved in full and incorporated into the portal. Each item will have a primary placement for its canonical breadcrumb path and curated cross-links into related categories for lateral exploration, sustaining narrative coherence without duplicating content ^[5].

Title (verbatim file name)	Primary category	Related-category links	Document type (normalized)
raport dla mediów Kłątwa.txt	Raporty i Analizy	Symbolika i Narracja; Manipulacje i Toksyczność	media report
kronika ósmego kręgu.txt	Symbolika i Narracja	Raporty i Analizy; Manipulacje i Toksyczność	chronicle/longform
kluczowe fakty.txt	Dokumenty i Fakty	Raporty i Analizy; Instrumentalne wykorzystanie instytucji	fact brief
jak instytucje wspierały manipulację.txt	Instrumentalne wykorzystanie instytucji	Manipulacje i Toksyczność; Raporty i Analizy	institutional analysis
5 faktow.txt	Dokumenty i Fakty	Raporty i Analizy	fact brief
krytyczna analiza materiału dowodowego.txt	Dokumenty i Fakty	Raporty i Analizy	evidence analysis
sluzebnosc a dozywocie.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy	legal analysis
Raport dla mediów.txt	Raporty i Analizy	Dokumenty i Fakty; Instrumentalne wykorzystanie instytucji	media report
akt darowizny.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy	official document

Principles for cross-linking and provenance

- Each item will display standardized metadata (category, document type, summary, source/author, publication date, tags, provenance notes, related items) to improve scannability, filtering, and trust [3].
- Polyhierarchy will allow documents to appear in multiple thematic groupings, while a single canonical breadcrumb prevents confusion and preserves a coherent sense of place [3].
- Provenance notes and transparent sourcing reinforce trust, given audiences’ reliance on structured, credible brands to assess information quality [4].

Nonfunctional requirements Accessibility

- The portal will conform to WCAG 2.2 Level AA across navigation, content, and media, including perceivability, operability (keyboard), understandability, robust semantics, adequate contrast, and target sizes [6].
- An in-article table of contents (nonsticky by default) will improve scannability and direct access to sections for long-form pieces [7].
- Accordions, where used, will have button semantics, aria-expanded states, clear expand/collapse signifiers, and multiple-open capability [8].

Audio module

- Each article will feature a “► Posłuchaj analizy” button that plays a corresponding analysis from /audio/ with accessible controls (play/pause, seek, volume, mute, rate) and keyboard operation, without autoplay [9].
- Prerecorded audio-only content will include downloadable transcripts to ensure equivalent access and to support search and SEO [10].

Privacy and legal compliance (EU/Poland)

- We will document and apply a legitimate interests assessment (LIA), meeting the three cumulative conditions of lawful interest, necessity, and balancing, and inform data subjects of the legitimate interests relied upon ^[11].
- Under Polish law, we will implement Press Law safeguards: avoid publishing personal data and images of persons in proceedings without consent or authority permission; avoid pre-judgment opinions; publish corrections (sprostowanie) per statutory format and deadlines ^[12].
- Governance measures will include Polish-language notices, DPO appointment/notification (if applicable), DPIA where required, breach readiness, cookies consent for non-essential storage, and careful evaluation of cross-border transfers and vendors ^[13].
- We will respect data-subject rights and explain exceptions applicable to certain public tasks, including complaint routes to UODO ^[14].
- Editorial risk controls will address defamation and permissible criticism, with documentation of verification steps and public-interest rationale ^[15].

Evidence integrity and provenance

- The portal will store canonical originals and surface structured provenance metadata; future work will detail evidence-integrity measures (e.g., hashing, timestamping) not covered in the current source materials ^[13].

Performance and security

- The portal will prioritize predictable layouts, efficient rendering, and secure handling of documents and audio, favoring EU-based infrastructure and least-privilege access to reduce risk; non-essential analytics will be opt-in and documented in the cookies/privacy policies ^[13].

Information architecture and UX patterns (preview)

- Global navigation: visible top-level categories in the header on desktop for immediate orientation; avoid hiding them behind a hamburger on large screens ^[2].
- Mega menu or routing pages: for rich category overviews, provide a click-activated mega menu with an accessible landing-page fallback to ensure full keyboard and screen-reader access ^[16].
- Breadcrumbs: canonical, hierarchical trails beginning with Home; the current page appears as non-link and trails are shortened appropriately on mobile ^[5].
- Category landing pages: descriptive summaries, subcategory links where relevant, and structured lists/cards with consistent metadata (title, date, type, tags) ^[3].
- Lists/cards: front-loaded titles and concise summaries to optimize scannability, supplemented by tags and related links ^[3].
- Expand/collapse: accordions used selectively with descriptive headings, caret/plus icons, and accessible behavior for content-heavy pages ^[8].
- Search and filtering: a persistent search box; faceted filtering to refine results and prevent overload; facets ordered by relevance and utility ^[17]. Mobile facets will support interdependent options and batch apply to limit reloads ^[18]. Users will have “clear all” and within-facet “see more” patterns for long lists ^[19].
- Table of contents: nonsticky “On This Page” modules within long articles, plus back-to-top links, to improve discoverability of lower-page content ^[7].
- Iconography and tone: supportive icons paired with clear text labels; conventional caret/plus indicators for collapsible elements; minimal motion to reduce cognitive load on sensitive investigative content ^[2].

- Bilingual accessibility: the interface will identify language and any changes via proper semantics and provide a clear language toggle to support assistive technologies and indexing ^[6].

Trauma-informed editorial stance The portal will maintain a serious, analytical voice, minimize motion and surprise interactions, and use clear contextual warnings where necessary, helping readers engage with sensitive material without undue distress ^[1]. Press Law-aligned presentation and correction workflows, alongside structured provenance and transparent sourcing, further reduce harm and promote accountability ^[12].

Objectives and intended outcomes

- Deliver a cohesive knowledge system that preserves and elevates all nine documents through clear taxonomy, polyhierarchical cross-linking, and canonical breadcrumbs ^[3].
- Provide accessible reading and listening experiences, including per-article “► Posłuchaj analizy,” transcripts, and WCAG 2.2-aligned interactions ^[10].
- Embed privacy and legal compliance into the editorial and technical workflows, including LIA documentation, Press Law safeguards, Polish-language notices, cookies consent, and rights management ^[11].
- Strengthen trust via provenance metadata, consistent IA patterns, and visible, well-structured navigation that supports discovery and orientation ^[4].
- Prepare the foundation for subsequent sections that will specify detailed IA, design system, component behaviors, compliance artifacts, and implementation plans, including audio delivery and operational governance ^[1].

This introduction establishes the mission, constraints, and standards for the investigative portal so that the subsequent design, compliance, and implementation sections can proceed with shared goals, traceable rationale, and measurable outcomes grounded in recognized IA/UX, accessibility, and legal frameworks ^[1].

Information Architecture, Taxonomy, and Content Inventory Mapping

1) Purpose and principles

- IA goal: The portal must preserve every file intact while enabling comprehension, discovery, and trust through clear navigation, strong information scent, and transparent provenance across all pages. ^[4]
- Polyhierarchy with a single canonical breadcrumb path: Each item has one canonical category for breadcrumbs to avoid confusion, while editorially curated cross-links place it in multiple thematic contexts to support lateral exploration and narrative coherence. This approach aligns with user-centered taxonomy design, card sorting/tree testing, and predictable labeling. ^[3]
- Wayfinding predictability: Visible global navigation, clear labels, and canonical breadcrumbs orient users who often arrive via deep links, reducing pogo-sticking and supporting trust. ^[2]
- Provenance clarity: For investigative content, normalized metadata and per-item provenance notes improve credibility and retrieval, sustaining reader trust in sensitive domains. ^[4]

2) Taxonomy and category definitions (operationalized)

- Manipulacje i Toksyczność
- Scope: Psychology of manipulation, toxic relationships, narcissism, and patterns of coercion/harm; analysis-focused and trauma-informed.
- Editorial note: Use non-stigmatizing labels and maintain analytical tone.
- Instrumentalne wykorzystanie instytucji
- Scope: Materials evidencing how institutional processes or actors support or amplify manipulation; systemic dynamics and accountability.

- Editorial note: Highlight procedural mechanisms and institutional behaviors.
- Dokumenty i Fakty
- Scope: Verifiable artifacts and factual summaries (official documents, evidence analysis, fact briefs); emphasis on source integrity.
- Editorial note: Normalize document types and sourcing details for consistent retrieval.
- Raporty i Analizy
- Scope: Interpretive outputs aimed at public/press (media reports, critical analyses, and chronicles when serving analytical functions).
- Editorial note: Provide concise abstracts and links to underlying evidence.
- Symbolika i Narracja
- Scope: Framing constructs and narratives that contextualize experiences (e.g., “Klątwa”), including chronicles as narrative scaffolding.
- Editorial note: Pair narrative framing with analytical cross-links to evidence.
- Polyhierarchy policy: An item can appear in multiple thematic contexts via related links, but retains one canonical category for breadcrumbs to preserve orientation and avoid label drift. [3]

3) Page types and templates

- Homepage
- Elements: Mission statement; five-category grid with supportive icons and clear labels; featured items; persistent search box; latest additions; editorial note on provenance and verification practices. [1]
- Behavior: Visible global navigation on desktop; language toggle; serious, analytical tone with minimal motion; high-contrast links. [2]
- Category landing pages (five)
- Elements: Category description; list of items rendered as cards with title, date (if available), document type, concise summary, tags; facet filters (category, document_type, tags, date); “View All” and pagination; cross-links to related categories. [3]
- Expand/collapse pattern: Card summaries use accessible accordions for “Read more”; headings are buttons with aria-expanded state, caret icons, multiple panels can be open, and collapsed content is programmatically hidden. [8]
- Item (document) pages (nine items)
- Elements: Immutable, verbatim .txt viewer; metadata panel; “► Posłuchaj analizy” per-article audio control wired to an accessible player (play/pause, seek, volume, mute, playback rate, keyboard operability); transcript toggle; provenance note; canonical download link to the original .txt; related items; breadcrumb trail. [10]
- Accessibility: Player supports keyboard and screen reader users; provide transcript for audio-only and captions for synchronized media when applicable; disable autoplay or provide immediate pause/stop/volume controls. [9]
- Long documents: Provide a nonsticky in-article table of contents labeled “On This Page,” plus a back-to-top link for scannability. [7]
- Utility pages
- About/Editorial Standards: Mission, methodology, sourcing, and verification; transparency commitments.
- Corrections/Right of Reply: Workflow and timelines aligned with Polish Press Law (sprostowanie), non-alteration rules, and publication placement. [12]
- Privacy and Cookies: GDPR-compliant notice in Polish, cookie/analytics consent, DPO contact, rights, transfers, and safeguards. [13]
- Methodology and Provenance: Hashing, timestamps, version history; redaction policy when legally necessary, documented without replacing canonical originals. [12]

- Audio and Accessibility: WCAG 2.2 commitments, transcripts/captions, keyboard access, focus visibility, target size, non-text contrast, and motion control. [10]
- Contact: Secure intake for queries and right-of-reply submissions.

4) Complete inventory mapping of the nine files (canonical placement, related-category links, normalized document type)

- All items are displayed in full, unaltered text and provide an original-file download link on their item pages.

Filename	Canonical category	Related-category links	Document type	Verbatim_display_and_download
raport dla mediów Kłątwa.txt	Raporty i Analizy	Symbolika i Narracja; Manipulacje i Toksyczność	media report	Yes
kronika ósmego kręgu.txt	Symbolika i Narracja	Raporty i Analizy; Manipulacje i Toksyczność	chronicle/longform narrative	Yes
kluczowe fakty.txt	Dokumenty i Fakty	Raporty i Analizy; Instrumentalne wykorzystanie instytucji	fact brief	Yes
jak instytucje wspierały manipulację.txt	Instrumentalne wykorzystanie instytucji	Manipulacje i Toksyczność; Raporty i Analizy	institutional analysis	Yes
5 faktow.txt	Dokumenty i Fakty	Raporty i Analizy	fact brief	Yes
krytyczna analiza materiału dowodowego.txt	Dokumenty i Fakty	Raporty i Analizy	evidence analysis	Yes
sluzebnosc a dozywocie.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy	legal analysis	Yes
Raport dla mediów.txt	Raporty i Analizy	Dokumenty i Fakty; Instrumentalne wykorzystanie instytucji	media report	Yes
akt darowizny.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy	official document	Yes

5) Metadata schema for items and lists (to support search, filtering, and provenance)

- Item-level schema (controlled vocabularies and normalized fields)
- title: Verbatim filename used as display title; optional display title alias.
- canonical_category: Single category for breadcrumb path.
- related_categories: Array of additional thematic contexts.
- document_type: Controlled vocabulary (e.g., media report, fact brief, evidence analysis, legal analysis, official document, chronicle).

- author/source: If known; else “unknown/unspecified.”
- date: ISO 8601 if available; else “undated.”
- summary: Editorial abstract (50–120 words), analytical and non-judgmental.
- tags: Controlled and free-form, Polish terms (e.g., “manipulacja”, “narcyzm”, “instytucje”, “dowody”, “raport”).
- provenance_note: How obtained; authenticity caveats; editorial handling.
- hash_sha256: Computed on original file for integrity verification.
- original_filename and canonical_url: For retrieval and citation.
- audio_asset: Path to /audio/ if available; transcript_url for downloadable transcript.
- rights and sensitivity_flags: Content warnings (e.g., distressing material).
- Rationale: Normalized metadata and controlled vocabularies improve information scent and consistent retrieval in lists, search, and facets. ^[3]
- List-level schema for category pages and search results
- Facets: category; document_type; tags; date (ISO); popularity/coverage ordering and availability filters; interdependent facets shown contextually post-selection; batch apply on mobile. ^[17]
- Sort options: newest first; A–Z by title; (optional) type priority.
- Pagination: 12–24 items per page with “View All” link; persistent search box outside mega menus. ^[16]

6) Navigation model and wayfinding aids

- Global header
- Visible top-level categories on desktop; utility links (About, Corrections, Privacy); language toggle (PL/EN); search box always visible. ^[2]
- Mobile: Hamburger labeled “Menu” opens a drawer with the same categories and utility links; maintain adequate contrast, size, and redundant navigation paths. ^[20]
- Breadcrumbs
- Structure: Home > Category > Item; current page not a link; canonical hierarchical path; truncate to avoid wrapping on mobile. ^[5]
- Related content
- Hybrid approach: Algorithmic similarity (shared tags/metadata) plus editorial curation to sustain narrative coherence and lateral exploration within categories. ^[2]
- Search
- Persistent input; results page with facets, clear-all control, and accessible labels; keyboard navigation and status messages; avoid hiding search inside menus. ^[16]

7) Preservation and integrity commitments (“no deletion” requirement)

- Source preservation
- No removal or alteration of any original file; all nine .txt files are preserved verbatim and remain downloadable as canonical sources to sustain trust. ^[4]
- Reader-facing provenance module
- Display SHA-256 hash, file size, last-modified timestamp, and version history; any legally necessary redactions are documented with clear editorial notes and do not replace the canonical original, consistent with Polish Press Law constraints. ^[12]
- Archive and backup strategy

- Versioned storage with periodic integrity checks (hash revalidation), redundant backups, and disaster recovery procedures; public-note updates for any format migrations to maintain transparency. ^[4]

8) Accessibility and multilingual hooks (IA-level)

- Language metadata
- Assign lang attributes sitewide; detect per-item language and tag for search; provide a clear PL/EN toggle with consistent layout and readable typography for both languages. ^[6]
- Accessible expand/collapse and focus order
- Accordions implemented as buttons with aria-expanded; multiple open panels allowed; programmatically hide collapsed content; keyboard support (Enter/Space, arrow keys, Home/End). ^[8]
- Keyboard operability and targets
- All interactive controls are keyboard operable; focus visible; touch targets $\geq 24 \times 24$ CSS pixels; high-contrast category icons with text labels. ^[6]
- Content warnings
- Provide a togglable content-warning pattern with respectful language for distressing material; minimal motion; pause/stop controls for any media exceeding five seconds. ^[1]

9) Cross-references and transition

- IA enables UX and visual components: The defined categories, canonical breadcrumbs, metadata schema, faceted lists, accessible accordions, and audio hooks provide the structural backbone for menus, cards, icons, and expanders in the next section. ^[2]
- Forward pointer: The upcoming UX, Navigation, and Visual Design System blueprint will detail layouts, grid systems, iconography, typography, and the accessible audio player implementation consistent with these IA and taxonomy rules. ^[1]

UX, Navigation, Visual Design System, and Audio Accessibility

1) Orientation and flow (transition from IA to UX)

Section SW2 established five top-level categories, canonical breadcrumb paths, polyhierarchical placement rules, and an inventory of nine verbatim .txt documents that must be preserved and displayed without edits. ^[3]

This section operationalizes that information architecture into concrete screens, interactions, and design system elements that improve readability, trust, and discoverability across the portal. ^[2]

Because audiences increasingly rely on clear structuring and transparent provenance to evaluate information, the UX prioritizes visible navigation, predictable layouts, and scannable metadata. ^[4]

The five categories defined in SW2 (Polish labels used sitewide):

- Manipulacje i Toksyczność
- Instrumentalne wykorzystanie instytucji
- Dokumenty i Fakty
- Raporty i Analizy
- Symbolika i Narracja

Canonical breadcrumbs will use a single authoritative path for each item, while polyhierarchy enables the same document to appear in multiple relevant categories via cross-links and curated related content. ^[5]

Transition principle: IA decisions become screen-level models—home, category, item, search/results—backed by global navigation, local navigation, and on-page TOCs to orient direct-entry visitors. [2]

2) Page layouts and navigation patterns

Homepage

- Mission header with concise tagline and portal purpose; place logo top-left and minimize motion to avoid distraction on sensitive topics. [1]
- Five-category grid with supportive icons and one-sentence descriptions to expose breadth at a glance and strengthen information scent. [2]
- Featured items strip (editorially curated) followed by a Latest section with timestamped cards using consistent metadata (title, date, document type, tags). [3]
- Persistent search box in the header; do not hide search inside a menu. [16]
- Visible global navigation on desktop with the five categories; avoid hiding them behind a hamburger to preserve context. [2]
- Language toggle (PL/EN) in the header, retaining user context when switching. [6]

Rationale: Making primary options visible on desktops reduces interaction costs and improves orientation, particularly for users arriving mid-site. [2]

Category landing pages

- Layout: Card-based lists (12–24 items per page) with consistent metadata: title (front-loaded keywords), publication date, normalized document type, concise summary, and controlled tags. [3]
- Breadcrumb trail “Home > Category” plus a brief category explainer to set expectations. [5]
- Faceted filtering on the left (desktop) and a top “Filters” drawer (mobile) with batch-apply: document_type, tags, and date range. [17]
- On mobile, apply selected facets in a batch to reduce reloads and friction. [18]
- Each card offers an accessible “Read more” expander to preview a few paragraphs without losing page context; implement expanders with proper button semantics. [8]

Item pages (nine verbatim documents)

- Immutable text viewer presenting the original .txt content verbatim, plus a canonical download link to the source file and a provenance note. [3]
- Metadata panel showing document type, date, author/source (if known), category placement, and normalized tags for reliable retrieval. [3]
- “► Posłuchaj analizy” button above the article body launches or focuses an accessible audio player; no autoplay. [10]
- Transcript toggle reveals a searchable, downloadable transcript below the player for all prerecorded audio-only analyses. [10]
- Related items: algorithmic by shared tags plus editorial curation to sustain narrative coherence and surface polyhierarchical relationships. [3]
- Canonical breadcrumb: “Home > Category > Item” with the current page as non-link. [5]
- For long texts, provide an “On This Page” table of contents and back-to-top links for fast intra-page navigation. [7]

Global navigation

- Header: category links (visible on desktop), utility links (About/Editorial, Corrections/Right of Reply, Methodology/Provenance, Privacy/Cookies, Contact), and a persistent search field. [2]
- Mega menu (optional on desktop): click-activated groups with supportive labels and short descriptions; provide a simple fallback by routing to landing pages to maximize accessibility. [16]
- Mobile: a labeled drawer (“Menu”) containing all primary and utility links; use a standard hamburger icon labeled “Menu” for clarity and redundant access paths. [20]
- Current location states are reflected in the global nav and breadcrumbs to orient direct-entry visitors. [2]

3) Interaction details and accessibility (WCAG 2.2 AA)

Accordions/expanders

- Use native button elements for headers; apply aria-expanded and aria-controls; hide collapsed content programmatically. [8]
- Allow multiple panels open; provide clear caret/plus icons reflecting state; support Enter/Space keys and optional Arrow navigation. [8]
- Avoid hiding essential content; ensure print styles expand accordions for offline review. [8]

Keyboard and focus

- Ensure every function is operable via keyboard with a logical focus order and visible focus indicators meeting non-text contrast. [6]
- Touch targets are at least 24×24 CSS pixels to support motor accessibility. [6]
- Provide skip-to-content and landmark navigation (header, main, nav, footer). [6]
- Prefer native HTML semantics and only add ARIA when needed to convey name/role/value. [21]

Breadcrumbs

- Implement canonical breadcrumb trails with the current page rendered as plain text and no link. [5]
- On mobile, truncate intermediate levels to avoid wrapping, while preserving the parent context. [5]
- Announce breadcrumbs as a navigation landmark for assistive technologies. [5]

Content warnings

- For potentially distressing material, display an optional, respectful disclosure with a clear “Reveal content” control and predictable, non-sensational behavior with minimal motion. [1]

4) Visual design system (serious, analytical tone)

Color system

- Base: neutral grays for backgrounds and containers, reserving color for emphasis and category cues; ensure text contrast $\geq 4.5:1$ for normal text and $\geq 3:1$ for large text. [6]
- Accent palette by category (used for labels, icons, and small highlights—not large surfaces):
- Manipulacje i Toksyczność: deep aubergine (measured, non-alarming)
- Instrumentalne wykorzystanie instytucji: steel blue
- Dokumenty i Fakty: graphite
- Raporty i Analizy: teal
- Symbolika i Narracja: muted amber

- Use red exclusively for error and destructive actions to avoid sensational signaling. ^[1]

Typography

- UI: system sans-serif stack for clarity in navigation and labels; Body: a highly legible serif or humanist sans for long-form reading with a consistent type scale. ^[1]
- Target line length of 60–80 characters and generous line-height to reduce cognitive load on long investigative texts. ^[1]

Grid and spacing

- 12-column responsive grid with consistent spacing increments (4/8 px scale) to maintain predictable card layouts that support quick scanning. ^[3]

Iconography and narrative signaling

- Pair category icons with explicit text labels; keep icons conventional and supportive, not primary carriers of meaning. ^[2]
- Accordions use caret/plus icons with clear state changes and enlarged click targets including the heading text. ^[8]
- Provide alt text for icons indicating category function rather than decorative semantics. ^[2]

Category-to-icon and color mapping: | Category (PL) | Icon cue | Accent color | | --- | --- | --- | | Manipulacje i Toksyczność | Interlocking hands/threads | Deep aubergine | | Instrumentalne wykorzystanie instytucji | Civic building/scales | Steel blue | | Dokumenty i Fakty | Document/checkmark | Graphite | | Raporty i Analizy | Report/chart | Teal | | Symbolika i Narracja | Spiral/mask | Muted amber |

Imagery

- Prefer abstract/editorial illustrations; avoid literal depictions of identifiable individuals; minimize or avoid motion and never autoplay animations. ^[1]

5) Search, filtering, and cross-linking

- Persistent, labeled search input in the header; do not bury search within menus. ^[16]
- Results page: facets for category, document_type, tags, and date; provide a prominent “Clear all” and keyboard-accessible, high-contrast controls. ^[19]
- Standardize taxonomy values to avoid duplicates and strengthen retrieval across facets. ^[18]
- Combine tag-based similarity with editorially curated “Related” lists to sustain narrative coherence and expose polyhierarchical relationships without content duplication. ^[3]

6) Audio module: player, playlist, and transcripts

Per-article controls

- Each item page includes “► Posłuchaj analizy” that opens/focuses an accessible player with play/pause, seek, volume, mute, playback rate, and time elapsed/remaining; do not autoplay audio. ^[10]
- All controls are keyboard operable with visible focus and clear labels/ARIA for screen readers. ^[9]

Player technology

- Use a privacy-respecting, accessible open-source player (e.g., Plyr or MediaElement.js) self-hosted with local assets; configure accurate roles/names/values and avoid third-party trackers or cookies by default. ^[13]

- Provide strong non-text contrast for controls and ensure robust behavior across assistive technologies. [6]

Playlist from /audio/

- Auto-generate a playlist view by scanning /audio/ and mapping tracks to their corresponding item pages; allow “play next,” queueing, and per-track metadata display. [9]

Transcripts

- Provide on-page toggleable transcripts and downloadable text for all prerecorded audio-only analyses, and clearly label when a transcript is auto-generated and edited for accuracy. [10]

Formats and delivery

- Offer modern and fallback formats (e.g., AAC/MP4 or HLS plus MP3/OGG), support byte-range requests and progressive loading, and expose an explicit “Download” option when appropriate. [9]
- Do not embed third-party analytics in the player; if optional analytics exist elsewhere, gate them behind consent and keep audio functional without it. [13]

Accessibility specifics

- Keyboard shortcuts (e.g., Space/Enter for play/pause, Arrow for seek), status message announcements for state changes, and WCAG-conformant contrast for icons and text. [9]
- Provide user media preferences (e.g., reduced motion toggles) and never autoplay previews. [1]

7) Multilingual UX (Polish primary, optional English)

- Provide a clear PL/EN toggle in the header that preserves the user’s current context and sets lang attributes at the document root and at inline language switches. [6]
- Localize UI labels and date/number formats per locale, and ensure search accommodates Polish diacritics with tolerant matching. [6]
- Avoid mixing languages within a single UI element to maintain clarity for assistive technologies. [6]

8) Performance and stability notes relevant to UX

- Favor predictable, low-motion layouts for sensitive content, and keep the homepage free from auto-advancing carousels to reduce cognitive load. [1]
- Use robust, validated semantic markup and expose accessible names/roles/values to ensure compatibility with assistive technologies. [6]
- Keep primary navigation and search server-rendered for speed and reliability in text-heavy views. [2]

9) Compliance and ethical guardrails embedded in UX

- Prominent, scannable pages linked from header/footer: Editorial Standards, Methodology/Provenance, Corrections/Right of Reply, and Privacy/Cookies, all available in Polish and describing user rights and complaint routes. [13]
- Implement a Corrections/Right of Reply (“Sprostowanie”) workflow with intake, deadlines, and placement rules compliant with Polish Press Law, and present links prominently on all pages. [12]
- Present a non-blocking, accessible cookie banner for non-essential analytics with clear opt-in choices; audio playback and core content should function without analytics consent. [13]

10) Acceptance criteria checklist

- Visible top navigation with five categories on desktop; accessible mobile drawer labeled “Menu.” [2]
- Category pages: card lists with consistent metadata, accessible expanders, and facets (document_type, tags, date) with batch apply on mobile. [3]
- Item pages: verbatim text viewer, canonical download link, provenance notes, related items, and canonical breadcrumbs. [5]
- Audio on every item: “► Posłuchaj analizy,” accessible player (no autoplay), and transcript toggle with downloadable text. [10]
- WCAG 2.2 AA: keyboard operability, visible focus, target sizes $\geq 24 \times 24$, adequate contrast, and accessible expand/collapse behaviors. [6]
- Visual system: consistent icon set paired with text, category accent colors with WCAG-compliant contrast, and no sensational imagery or motion. [2]
- Search: persistent header search; results with facets, a “Clear all,” and robust empty states. [16]
- Multilingual: PL/EN toggle preserves context and sets appropriate lang semantics; localized formats for dates and numbers. [6]

Implementation notes tied to the nine .txt documents

- All nine source files are preserved verbatim on their item pages with immutable displays and canonical downloads; polyhierarchy is achieved through cross-links and curated “Related” lists while preserving a single canonical breadcrumb path per item. [3]
- Where an item reasonably belongs to multiple categories (e.g., a report relevant to both “Raporty i Analizy” and “Symbolika i Narracja”), use editorial curation and tags to surface lateral links without duplicating the content. [3]

Legal, Privacy, and Ethical Framework (EU/Poland) for a Document-Heavy Investigative Portal

1) Purpose and scope — This framework sets binding guardrails for publishing nine Polish .txt source documents and per-article audio in a portal examining manipulation, toxicity, and institutional dynamics, balancing rigorous public-interest reporting with legal, privacy, and accessibility duties to build trust and minimize harm in a fragmented information environment [4]. The framework ensures compliance with Polish Press Law constraints and GDPR, while preserving evidentiary integrity and enabling responsible disclosure of sensitive material [12].

2) GDPR compliance (EU/Poland)

- Lawful basis (legitimate interests) and LIA: We rely on Article 6(1)(f) GDPR (legitimate interests) for investigative publication, subject to a documented three-part test: (i) a lawful, clearly articulated, real/present interest; (ii) strict necessity—no equally effective, less intrusive alternative; and (iii) a balancing test showing that data subjects’ rights and freedoms do not override our interests, all assessed case-by-case and documented before processing [11]. Controllers must inform data subjects of the legitimate interests relied upon and conduct the balancing test with methodology considering the nature/source of interest, reasonable expectations, impacts, and safeguards [11].
- Necessity, minimization, and transparency: Where legitimate interests are used, we confirm the interest cannot reasonably be achieved just as effectively by less intrusive means; if such means exist, Article 6(1)(f) cannot be relied upon [11]. We provide GDPR-compliant privacy notices (including the legal basis, recipients, retention, transfers, rights, and UODO complaint routes) in Polish for consumer-facing communications, with additional Polish-language expectations applicable in proceedings [13]. Reasonable expectations are evaluated considering

relationship, proximity, collection context, service nature, and confidentiality obligations; common practice alone does not establish expectations [22].

- Redaction standards and evidentiary integrity: To avoid doxxing and unlawful processing, we redact personal data and images of suspects, witnesses, and victims unless consent or competent authority permission is obtained and avoid publishing private-life information unless directly related to public activity, while annotating redactions to preserve evidentiary context [12]. In light of the portal's constraint to preserve all files, originals are retained in a secure, immutable archival store with provenance notes, while public-facing versions apply lawful redactions and contextualization [12].
- Data-subject rights and objections: Rights of access, rectification, erasure, and objection are honored with standard GDPR timelines; when a data subject objects on grounds related to their particular situation, processing ceases unless compelling legitimate grounds exist that override the data subject's interests [11]. Certain public-task controllers may refuse access when necessary for task performance and if rights/freedoms are not superior or if classified information protection would be infringed, which is noted for interactions with public institutions [14].
- DPIA triggers and breach readiness: We conduct a DPIA where UODO lists the processing operation (e.g., monitoring or profiling) or otherwise presents high risk, maintain records of processing, and implement data protection by design/default [13]. Breaches are notified to the supervisory authority without undue delay and, where feasible, within 72 hours, with "awareness" defined as reasonable certainty that a breach occurred; preliminary notifications are filed if investigations exceed 72 hours, with withdrawal if ultimately not notifiable [23].
- Special-category data and safeguards: Sensitive material (e.g., health, sexuality) is treated with heightened safeguards, risk assessment, and minimization, and may require DPIA and lawful redaction before publication where public-interest necessity is not met [13].
- Vendors, cross-border transfers, and EU hosting: International transfers require appropriate safeguards (e.g., SCCs) and Transfer Impact Assessments; vendors are evaluated for privacy/security commitments, and EU-based hosting/CDNs are preferred to minimize transfer risk [13]. Controllers should scrutinize third-country authority requests case-by-case, as mere avoidance of foreign sanctions may not override data subjects' rights [11].

3) Polish Press Law and defamation safeguards

- Right of reply (sprostowanie): Editors must publish a factual correction of inaccurate or untrue information upon request by an interested person/entity, submitted within 21 days, including identity/address, concise formatting (not exceeding twice the size/time of the original fragment), and prompt publication with specified placement and typography; text cannot be altered without consent, and no commentary may accompany it in the same issue or on the same day electronically (a simple notice of forthcoming response is allowed) [12]. Electronic corrections must be published within 3 working days in e-form and labeled clearly (e.g., "Sprostowanie") [12].
- Publication constraints: It is prohibited to publish personal data and images of persons under preparatory or court proceedings, witnesses, or victims unless they consent or competent authorities permit disclosure; opinions on case outcomes are prohibited before a first-instance judgment [12]. Publishing information about the private sphere requires consent unless directly related to public activity; audio/visual publication requires consent from persons providing information [12].
- Defamation due diligence: Criminal defamation (Article 212) protects external honor/good name and can be committed via text or mass media; civil remedies (injunctions, apologies, immaterial/material damages) also apply; permissible criticism and public-interest context can limit liability but require verification, contextual accuracy, and non-stigmatizing presentation, which must be documented in editorial files [15].

4) Audio, cookies, and consent

- Accessible audio and transcripts: Provide an alternative for time-based media for prerecorded audio-only content (transcripts), ensure keyboard operability, visible focus, and controls for play/pause, seek, volume, mute, and rate; if

any audio plays automatically for more than 3 seconds, provide immediate mechanisms to pause/stop or control volume ^[10]. Media players must be accessible to keyboard and screen reader users and support captions/transcripts and user customization ^[9].

- UI targets and motion: Maintain minimum 24×24 CSS pixel target sizes for controls and adequate non-text contrast; avoid flashing and excessive motion, and provide pause/stop for any autoplay media ^[24].
- Cookies and telemetry: Obtain opt-in consent for non-essential analytics or media cookies and present clear prior information; do not rely solely on browser settings for cookie consent under Polish law ^[13]. Configure the audio player to function without non-essential tracking by default and only enable preference storage or telemetry after consent ^[13].
- Children's data: Apply heightened diligence for minors' data, minimizing collection and applying clear consent and redaction safeguards where applicable under Polish data protection rules ^[13].

5) Ethical and trauma-informed presentation

- Language and tone: Use non-stigmatizing, precise language when addressing manipulation, toxic relationships, and narcissism; maintain a serious, analytical tone with predictable layouts and minimal motion to reduce cognitive load for sensitive content ^[1]. Provide content warnings where distressing material may appear and avoid sensationalism to support reader safety and trust ^[1]. Publicly disclose editorial standards, verification methods, and limitations to strengthen transparency and credibility in a low-trust environment ^[4].

6) Operational checklists and templates

- GDPR/LIA checklist
- Define and document a lawful, clearly articulated, real/present legitimate interest tied to public-interest reporting ^[11].
- Necessity test: confirm no equally effective, less intrusive alternative exists for the same purpose ^[11].
- Balancing test: assess impacts, reasonable expectations, and safeguards; record the outcome and inform data subjects of the reliance on legitimate interests ^[11].
- Provide Polish-language privacy notices detailing rights, legal bases, recipients, retention, transfers, and UODO complaint routes ^[13].
- Conduct DPIA where UODO requires or risk is high; implement data protection by design/default; maintain records of processing ^[13].
- Breach readiness: define awareness thresholds, trigger 72-hour internal clock, and use preliminary notifications if investigations exceed 72 hours ^[23].
- Vendors/transfers: execute SCCs/TIAs for non-EU services and prefer EU hosting/CDNs to mitigate transfer risk ^[13].
- Press Law/right-of-reply workflow checklist
- Intake: verify requester's standing, identity, address, and timeliness (≤ 21 days); ensure the correction is concise ($\leq 2 \times$ original fragment's size/time) ^[12].
- Editorial handling: publish within prescribed timelines (including 3 working days for e-form), in the same section with specified typography, labeled "Sprostowanie," without alterations or same-day commentary ^[12].
- Constraints: enforce consent/authority permissions for publishing personal data/images in proceedings, avoid pre-judgment opinions, and obtain consent for private-life information and audio/visual publication ^[12].
- Minimal sample texts
- Privacy notice highlights
 - We process personal data in the legitimate interests of public-interest investigative reporting, after a case-by-case balancing test and necessity analysis ^[11].

- You have rights of access, rectification, erasure, and objection; contact our DPO and you may lodge a complaint with the Polish DPA (UODO) ^[13].
- We may use EU-based service providers and, where transfers outside the EU occur, apply SCCs and conduct TIAs ^[13].
- This notice is available in Polish and English; the Polish version governs for Polish consumers ^[13].
- Corrections/Right-of-Reply notice
 - If you believe a published statement is inaccurate or untrue, you may submit a factual correction (“sprostowanie”) within 21 days, including your identity and address; concise formatting rules apply, and we will publish promptly with prescribed placement and typography ^[12].
 - We will not alter your correction without your consent and will refrain from commentary in the same issue or day in electronic form; a notice of forthcoming response may be provided ^[12].
- Audio transcript disclosure
 - A text transcript is provided for each prerecorded audio analysis to ensure accessible, non-audio access for all users ^[10].
- Retention and archival approach
- Original .txt files are preserved in a secure, immutable archive with provenance notes; public-facing versions apply necessary lawful redactions and annotations while maintaining contextual integrity ^[12].

7) Governance and accountability

- Roles and gates: The Editor-in-Chief is accountable for editorial standards and Press Law compliance; the Data Protection Lead (DPO) oversees GDPR governance, DPIA/LIA documentation, breach readiness, and UODO liaison; a Legal Review Gate signs off on publication decisions involving high-risk personal data or sensitive allegations ^[13]. Entities appointing a DPO must notify UODO within 14 days and make the DPO's name and contact publicly available ^[13].
- Logging and audits: Maintain audit logs of corrections/right-of-reply requests, redaction decisions, LIA/DPIA outcomes, and breach notifications to demonstrate accountability and enable supervisory review ^[13].
- Policy cadence: Review privacy, corrections, and audio accessibility policies at least annually or upon regulatory change; ensure Polish-language versions are updated in parallel and accessible to users ^[13].

Implementation notes: All nine source documents must be preserved and utilized; the portal presents them within categories while respecting redaction and consent constraints, and each article provides an accessible “► Listen to analysis” control with transcripts and consent-respecting media behavior ^[10].

Technical Architecture, Security, Performance, and Implementation Blueprint

1) Architecture recommendation (trade-offs and final choice)

- WordPress newsroom stack (pros/cons): WordPress offers familiar editorial workflows, rich plugin ecosystems for media, and quick setup; however, it introduces performance variability, plugin security risks, and limited guarantees for immutable verbatim rendering of sensitive .txt evidence without custom development and strict governance. Discoverability can be well supported via established navigation patterns, but ensuring privacy-first operation (no third-party trackers, EU-only hosting/CDNs, strict CSP) is harder in a plugin-heavy stack.

- Pure static generator (e.g., Hugo/Jekyll): Delivers excellent performance and simplicity; however, editorial workflows are developer-centric, versioning and provenance require bespoke tooling, and multilingual navigation with faceted search and accessible audio often demands significant custom code and maintenance.
- Jamstack + headless CMS (recommended): Pair a headless CMS (e.g., Strapi or Sanity) with a static or hybrid framework (e.g., Next.js with static export/incremental static regeneration). This approach provides:
- Immutable, verbatim rendering of .txt evidence with canonical downloads and provenance panels.
- Predictable, visible category navigation, canonical breadcrumbs, and routing pages, aligning with investigative portal UX and trust-building needs. Visible global navigation on desktop and clear category labels preserve reader orientation and reduce interaction costs ^[2]. Breadcrumbs should follow a canonical hierarchical path with the current page as non-link and start with Home ^[5]. Mega menus can reveal options at a glance and should be click-activated with routing-page fallback for accessibility ^[16].
- Privacy-first delivery with EU-based hosting/CDN, strict CSP, and self-hosted analytics, supporting GDPR and Polish-law constraints ^[13].
- Strong provenance features (hashes, timestamps, version history) and bilingual PL-first UI with optional EN, presented in a serious, analytical design with minimal motion ^[1].
- Faceted search and in-page table of contents patterns to support discovery and scannability at scale ^[17]. Final choice: Jamstack + headless CMS with static export for category and item pages, serverless functions where needed (search indexing, audit), and EU-based hosting/CDN. This architecture balances performance, control, evidence integrity, and accessibility while meeting legal/privacy constraints and trust requirements for investigative publishing ^[4].

2) Core content model and categories

- Categories (Polish labels):
- Manipulacje i Toksyczność
- Instrumentalne wykorzystanie instytucji
- Dokumenty i Fakty
- Raporty i Analizy
- Symbolika i Narracja
- Item schema fields (polyhierarchy via related_categories; canonical breadcrumb path via canonical_category). Use controlled vocabularies and taxonomy rules to ensure consistent retrieval and strong information scent ^[3]. Breadcrumbs show a single canonical path per item ^[5].

Field	Type	Required	Description
title	string	yes	Display title; front-load key terms; use the verbatim filename for clarity.
canonical_category	enum	yes	One of the five categories for breadcrumb path and canonical placement.
related_categories	array	no	Additional categories to support polyhierarchy and cross-linking ^[3] .
document_type	enum	yes	Normalized types (report, analysis, official document, media statement) to power facets ^[3] .
author/source	string	no	Author or source organization for transparency.
date	ISO 8601	yes	Publication date for sorting and schema.org metadata ^[3] .
summary	string	yes	Concise abstract to feed list/cards and TOC anchors ^[3] .
tags	array	yes	Controlled vocabulary (topics, institutions, actors, locations) ^[3] .
provenance_note	string	yes	Brief origin/sourcing/verification notes to reinforce trust ^[3] .
hash_sha256	string	yes	Hash of original .txt for evidence integrity.
original_filename	string	yes	Original filename displayed and preserved.
canonical_url	URL	yes	Stable item URL; included in schema.org and sitemaps.
audio_asset	path	no	/audio/ mapping to the item's track (if available).
transcript_url	URL	no	Downloadable transcript of the audio (if audio-only).
sensitivity_flags	array	no	Flags for content warnings, privacy/redaction notes (if needed).

3) Document rendering and preservation

- Immutable .txt viewer: Render the .txt payload verbatim, preserving whitespace and line breaks, with no editorial changes. Provide a prominent “Download original” link to the canonical file to guarantee unaltered access.
- Provenance panel: Display SHA-256 hash, file size, last-modified timestamp, and version history (diff summaries with timestamps). Any lawful redactions (if required by Press Law) appear as transparent editorial notes; originals remain preserved and downloadable, avoiding substitution ^[12].
- Category mapping and publication of all nine files (canonical category, with cross-links):

Filename	Canonical Category	Cross-links
raport dla mediów Kłątwa.txt	Raporty i Analizy	Symbolika i Narracja; Manipulacje i Toksyczność
kronika ósmego kręgu.txt	Symbolika i Narracja	Raporty i Analizy; Manipulacje i Toksyczność
kluczowe fakty.txt	Dokumenty i Fakty	Raporty i Analizy; Instrumentalne wykorzystanie instytucji
jak instytucje wspierały manipulację.txt	Instrumentalne wykorzystanie instytucji	Manipulacje i Toksyczność; Raporty i Analizy
5 faktow.txt	Dokumenty i Fakty	Raporty i Analizy
krytyczna analiza materiału dowodowego.txt	Dokumenty i Fakty	Raporty i Analizy
sluzebnosc a dozywocie.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy
Raport dla mediów.txt	Raporty i Analizy	Dokumenty i Fakty; Instrumentalne wykorzystanie instytucji
akt darowizny.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji; Raporty i Analizy

4) Audio module: per-article “► Posłuchaj analizy”

- Player: Integrate a self-hosted, accessible HTML5 player (e.g., Plyr or MediaElement.js) with play/pause, seek, volume/mute, playback rate, and time display; ensure keyboard operability, visible focus, proper ARIA names/roles, and no autoplay to comply with WCAG 2.2 Level AA requirements for media ^[10]. The player must be usable via keyboard and screen readers, support transcripts, and maintain non-text contrast for controls ^[9]. Target sizes of controls should be ≥24×24 CSS pixels, and status messages should be announced to assistive tech ^[24].
- UI placement: A “► Posłuchaj analizy” button appears above the text viewer on each item page. Clicking reveals the player and transcript toggle. Transcripts are downloadable and indexable in search to support perceivable access and discovery ^[10].
- Playlist and mapping: Auto-generate a playlist from /audio/ by filename or item ID mapping; display per-item play controls (no global autoplay).
- Privacy: No third-party trackers; audio functions without analytics consent. If cookies store player preferences, they are non-essential and gated behind opt-in consent in line with Polish Telecommunication Law/ECA cookie requirements ^[13].

5) Search and discovery

- Full-text search: Build a compile-time client-side index (e.g., Lunr or FlexSearch) over .txt content, transcripts, and normalized metadata (title, tags, category, document_type). For larger scale, provide a server-side index behind rate limits.
- Facets: Category, document_type, tags, and date facets; values normalized via controlled vocabularies; batch apply on mobile, “Clear all,” and accessible labels/status messages. Design facets by popularity, coverage, and utility, and show only available values ^[17]. Use interdependent facets to show context-relevant options and standardize metadata to avoid duplicates; batch apply on mobile to reduce reloads ^[18]. Provide breadcrumb-like reset and truncation with “See more” or search-within-facet for long lists ^[19].

- Polish-language search: Normalize diacritics, apply stemming where feasible, and index transcripts to improve retrieval. Keep the search box visible in the header to reduce interaction cost and improve consistency ^[2].
- In-page TOC: For long articles and reports, include a nonsticky “On This Page” table of contents to enhance scannability and direct access to sections ^[7].

6) Schema.org and SEO

- Per-item structured data: Use schema.org CreativeWork or NewsArticle with headline (title), datePublished (date), author (author/source), about (summary), keywords (tags), inLanguage (pl or en), and url (canonical_url).
- BreadcrumbList: Emit structured breadcrumbs for the canonical trail (Home › Category › Item) ^[5].
- Sitemaps: Generate localized sitemaps for PL and EN routes. Use canonical URLs and robots directives to avoid indexing sensitive endpoints (e.g., back office, manifests).

7) Performance and Core Web Vitals

- LCP: Server-render pages (static export/ISR), inline critical CSS, and avoid heavy hero imagery; use EU-based CDN with HTTP/2 or HTTP/3, long cache lifetimes, and immutable asset hashing (cache-busting via file fingerprints) ^[13].
- CLS: Reserve fixed space for the player and accordions; load fonts with stable metrics; avoid late-inserting UI. Use predictable, supportive iconography with text labels to preserve clarity in investigative contexts ^[1].
- INP: Limit blocking JS, defer non-critical scripts, and lazy load related lists and transcript panels until user intent is expressed.
- Build pipeline: Use incremental static regeneration for timely updates; prefetch related content on hover/focus where appropriate; compress assets (Gzip/Brotli) and serve modern image formats for any illustrative iconography. Keep motion minimal and provide pause/stop for any moving content to reduce cognitive load on sensitive pages ^[1].

8) Security and hardening (OWASP-aligned)

- HTTP security headers:
- Content-Security-Policy (CSP): Restrict scripts and media sources; disallow unsafe-inline; explicitly allow self-hosted audio paths.
- Strict-Transport-Security (HSTS): Enforce HTTPS.
- X-Content-Type-Options: nosniff.
- X-Frame-Options/Frame-Ancestors: deny or restrict embedding.
- Referrer-Policy: strict-origin-when-cross-origin.
- Permissions-Policy: limit sensors, autoplay, and camera/microphone as appropriate.
- Input handling and abuse prevention: Sanitize queries; implement rate limiting on search endpoints; robust error handling and logs without PII leakage. Prefer native HTML semantics and use ARIA only when necessary, avoiding roles that hide descendant semantics ^[21].
- Editorial back office: Enforce MFA, least-privilege roles, and audit trails; logically separate the CMS from the public delivery layer. Appoint and publish DPO contact, maintain records of processing, and conduct DPIA where required under Polish law ^[13].
- Media/document handling: Validate uploads, verify MIME types, store originals in write-once, versioned storage, and run periodic integrity checks. Breach readiness includes defining “awareness” triggers and preliminary notifications where investigations exceed 72 hours ^[23].

Example CSP (adjust sources to deployment):

Content-Security-Policy: default-src 'self'; script-src 'self'; style-src 'self' 'unsafe-inline'; img-src 'self' data:; font-src 'self'; media-src 'self'; connect-src 'self'; frame-ancestors 'none';

9) Evidence integrity and provenance (technical features)

- Hashes and manifests: Compute SHA-256 per original .txt and store a signed manifest per release; display hash and last-verified date on item pages to reinforce transparency and reader trust ^[4].
- Trusted timestamps: Apply TSA-backed timestamps at publication; maintain version history with diff summaries and clear editorial notes for contextualization.
- Public integrity page: Publish a consolidated list of all files with hashes, timestamps, and verification cadence for independent checking.

10) Multilingual and accessibility implementation

- Language toggle: PL as primary with optional EN; preserves route structure; set lang attribute at document root and mark inline language changes to support assistive technologies ^[6].
- Navigation and components: Visible global navigation on desktop with click-activated mega menu or routing pages; avoid multilevel cascading menus; provide local navigation and breadcrumbs to orient users who often arrive on internal pages ^[2].
- Accordions and TOC: Use button semantics for accordion headers, proper aria-expanded state, caret/plus icons, expand/collapse all controls, and programmatic hiding of collapsed content; maintain consistent styling across pages ^[8].
- Player and controls: Keyboard-operable media player with visible focus, target sizes $\geq 24 \times 24$ CSS pixels, and status announcements; provide transcripts/captions and avoid autoplay beyond three seconds without immediate controls ^[10].
- Skip links and landmarks: Provide skip-to-content, header/main/nav/footer landmarks, and maintain high contrast and predictable behavior for sensitive content ^[6].
- Hamburger icon on mobile: Standard three-line icon labeled "Menu," adequate contrast/size/padding, and redundant navigation paths to accommodate diverse users ^[20].

11) Analytics and consent (privacy-respecting)

- Analytics: Prefer self-hosted Matomo in cookieless mode by default; offer explicit opt-in for cookies and enhanced analytics in a minimal, accessible CMP. Provide prior clear information and purpose-specific controls for non-essential analytics in alignment with Polish Telecommunication Law/ECA cookie requirements; avoid relying solely on browser defaults ^[13].
- Vendor/transfer controls: Prefer EU hosting/CDNs; if non-EU vendors are used, conduct TIAs and implement SCCs; update privacy and cookies pages with transfer disclosures and safeguards ^[13].
- Audio independence: Audio playback works without analytics consent; only player preferences (if stored) are gated behind opt-in consent and clearly documented.

12) Implementation roadmap (phased)

- Phase 1 (Foundation)
- Set up repo, headless CMS, and EU-hosted Jamstack deployment.
- Implement data model and controlled vocabularies for categories, document types, and tags.
- Build category landing pages, item templates with immutable .txt viewer, canonical "Download original," breadcrumbs, and metadata.
- Publish all nine files with canonical mapping and cross-links.
- Phase 2 (Media and Search)

- Integrate accessible audio player and per-item “► Posłuchaj analizy” button.
- Implement transcripts (toggleable and downloadable).
- Build full-text search with facets, mobile batch apply, “Clear all,” and status messages.
- Add related-content modules powered by tags and category proximity.
- Phase 3 (Trust and Compliance)
 - Add provenance panel (SHA-256 hashes, timestamps, version history); publish integrity page.
 - Implement Privacy/Cookies, Corrections/Right of Reply, and Editorial Standards pages aligned with Polish Press Law workflows ^[12].
 - Emit schema.org metadata and breadcrumb structured data ^[5].
- Phase 4 (Performance and Security)
 - Configure EU CDN, caching, immutable asset hashing, HTTP/2/3; optimize for LCP/CLS/INP.
 - Apply security headers, CSP, input sanitization, search rate limiting, audit logging; finalize DPIA/LIA documentation and DPO publication ^[13].
- Phase 5 (Polish/English UX)
 - Implement language toggle; localize UI; verify diacritics-aware search and indexing.
 - Conduct WCAG 2.2 AA accessibility QA; remediate focus, labels, ARIA, and contrast issues ^[6].

13) Acceptance criteria

- All nine .txt files are published verbatim with canonical downloads and provenance panels (hash, size, timestamp, version history).
- Five category landing pages present structured lists/cards with strong labels, accessible expanders, and consistent metadata; visible global navigation with canonical breadcrumbs.
- Item pages include a “► Posłuchaj analizy” button, accessible audio player, and transcript toggle; audio functions without analytics consent.
- Persistent header search delivers diacritics-aware Polish matching; results offer facets with batch apply on mobile and “Clear all.”
- WCAG 2.2 AA conformance for navigation, accordions, player, focus visibility, target sizes, and color contrast; language identification and skip links implemented ^[6].
- Security headers and CSP implemented; search rate-limited; audit logging without PII leakage; EU-based hosting/CDN and documented transfer safeguards ^[13].
- Schema.org metadata (CreativeWork/NewsArticle and BreadcrumbList) present; canonical URLs and localized sitemaps; robots directives exclude sensitive endpoints ^[5].
- Privacy/Cookies and Corrections/Right of Reply pages are available in Polish with compliant workflows and timelines ^[12].

Navigation and design system notes

- Use a serious, analytical tone with predictable layouts, high-contrast link text, and minimal motion to build credibility and reduce distraction; pair supportive icons with clear text labels across categories ^[1]. Mega menu entries should be clickable with routing pages to ensure accessibility and avoid hover-only behavior ^[16]. Accordions must use button semantics, proper ARIA states, and clear expand/collapse signifiers ^[8]. Breadcrumbs reflect the canonical hierarchy and include the current page as non-link ^[5].

Security/compliance and governance alignment

- GDPR lawful basis and governance: Document Legitimate Interests Assessments (LIA) with necessity and balancing tests; provide Polish-language privacy notices; maintain records of processing; appoint and publish DPO contact; and implement breach readiness aligned to 72-hour notification guidance ^[11]. Respect Polish Press Law constraints on publication of personal data/images and correction workflows ^[12]. Prefer EU hosting/CDNs and assess transfers with TIAs/SCCs as needed ^[13].

Transition to SW6 The final section (SW6) will consolidate evidence-integrity communication, analytics/consent disclosures, and operational governance into a concluding implementation and governance plan, preparing the article for assembly and launch ^[4].

Governance, Evidence Integrity, Multilingual Strategy, and Launch Roadmap

1) Binding constraints and alignment

- Non-removal and immutability: All nine .txt source files are published in full, verbatim, with canonical download links on their item pages; originals are immutable and stored in read-only archival storage. If strictly necessary for lawful compliance, public views may include minimal, clearly annotated redactions while preserving access to the unchanged canonical download and documenting rationale in provenance notes.
- Category organization and polyhierarchy: The five categories are the primary navigation and discovery lenses—Manipulacje i Toksyczność; Instrumentalne wykorzystanie instytucji; Dokumenty i Fakty; Raporty i Analizy; Symbolika i Narracja. Each item receives a single canonical breadcrumb path, with curated cross-links to related categories to support polyhierarchical discovery without duplicating content ^[3].
- Navigation and orientation: Primary categories are visible on desktop in the header, with either a click-activated mega menu or routing pages and canonical breadcrumbs to orient users who often arrive deep-linked from external sources ^[2]. Breadcrumbs reflect a single hierarchical path and end with a non-linked current page ^[5].
- Audio module: Every item page provides a “► Posłuchaj analizy” button that launches an accessible, self-hosted player. Keyboard operability, visible focus, and predictable controls comply with WCAG 2.2 AA ^[10]. The player is fully usable for keyboard and screen-reader users ^[9]. Transcripts are available via a toggle and as downloadable text.
- Tone and safety: The site uses a serious, analytical voice, minimal motion, high-contrast text, and predictable layouts to build credibility for sensitive topics ^[1]. Icons are supportive but always paired with clear text labels to preserve accessibility and information scent ^[2].

Nine-file preservation and canonical categorization

- All files remain intact, presented in an immutable viewer with canonical downloads and hashes. Cross-links allow lateral exploration without duplication.

File	Canonical Category	Cross-links
raport dla mediów Kłątwa.txt	Symbolika i Narracja	Raporty i Analizy; Manipulacje i Toksyczność
kronika ósmego kręgu.txt	Symbolika i Narracja	Raporty i Analizy
kluczowe fakty.txt	Dokumenty i Fakty	Raporty i Analizy
jak instytucje wspierały manipulację.txt	Instrumentalne wykorzystanie instytucji	Raporty i Analizy; Manipulacje i Toksyczność
5 faktow.txt	Dokumenty i Fakty	Raporty i Analizy
krytyczna analiza materiału dowodowego.txt	Raporty i Analizy	Dokumenty i Fakty
sluzebnosc a dozywocie.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji
Raport dla mediów.txt	Raporty i Analizy	Dokumenty i Fakty
akt darowizny.txt	Dokumenty i Fakty	Instrumentalne wykorzystanie instytucji

2) Evidence integrity and provenance (reader trust and transparency)

- Hashes and manifests
- Compute and display a SHA-256 hash, file size, and last-verified timestamp on every item page. Publish a signed Integrity page listing all items, canonical URLs, hashes, and timestamps per release.
- Versioning and chain of custody
- Maintain a time-stamped version history for contextual updates without replacing originals. Each item includes a provenance note describing source, validation steps, and known limitations to support informed interpretation and trust [4].
- Redaction policy
- If compelled by law or ethics, apply minimal, context-preserving redactions only to public views; annotate clearly in-page and never substitute the canonical original download. Retain originals in secure archival storage with least-privilege access controls.
- Citation and verification
- Provide a standardized citation block and a “Verify using SHA-256” prompt with instructions. Encourage independent verification by re-computing hashes and checking timestamps.

3) Multilingual and localization plan (Polish primary, optional English)

- Language identification and semantics
- The site defaults to lang="pl"; English routes use lang="en"; inline language changes are marked semantically to assist screen readers and indexing [6].
- UI and navigation
- A persistent PL/EN toggle in the header preserves context and deep links. All labels, category descriptions, system messages, and error states are localized, with identical layout and predictable behavior across languages for scannability [1].
- Search behavior and indexing

- Implement diacritics-aware matching for Polish queries. Index titles, verbatim .txt content, abstracts, tags, and transcripts. Provide relevant, well-ordered facets with batch-apply on mobile to reduce overload ^[17].
- Locale formats and metadata
- Localize dates and numbers per user language. Emit schema.org metadata with itemLanguage and inLanguage fields and declare page language via HTML attributes for accessibility and SEO ^[6].

4) Analytics, consent, and privacy-respecting measurement

- Minimal, cookieless analytics
- Use an EU-hosted, self-managed, cookieless analytics baseline (e.g., Matomo) for aggregated traffic insights. Offer opt-in cookies for enhanced metrics via a clear, accessible consent banner with purpose-specific controls ^[13].
- Consent and audio
- Core reading and the audio player work without analytics consent. If storing player preferences (e.g., playback speed), treat them as non-essential and gate them behind opt-in cookie consent per Polish e-privacy rules ^[13].
- Privacy policy and rights
- Publish a Polish-language privacy and cookies notice detailing lawful bases (including legitimate interests with necessity and balancing), minimization, retention, recipients, transfers, and data-subject rights, plus DPO contact and the UODO complaint route ^[13]. Where relying on legitimate interests, document the LIA before processing and inform data subjects of the interests relied upon ^[11].
- Breach readiness
- Define “awareness” triggers, notify the supervisory authority within 72 hours if required, and issue preliminary notifications when investigation exceeds that window, with withdrawal if not notifiable ^[23].

5) Operational governance and right-of-reply workflow

- Roles and responsibilities
- Editor-in-Chief: final editorial decisions, ethical guardrails, Press Law alignment, approval of sensitive disclosures.
- Data Protection Lead (DPO): LIA/DPIA stewardship, GDPR record-keeping, breach response, supervisory liaison, privacy notices.
- Legal Review: pre-publication sign-off on high-risk items, defamation risk assessment, and necessity/proportionality of any redactions.
- Right-of-reply (sprostowanie) and corrections
- Intake requests within statutory timelines, verify identity and concise formatting, and publish promptly with required placement and labeling (“Sprostowanie”), without alteration or same-day commentary in electronic form ^[12].
- Press Law safeguards
- Do not publish personal data or images of suspects, witnesses, or victims without consent or authority permission; avoid pre-judgment opinions; obtain consent for publication of private-life or audio/visual content unless it directly relates to public activity ^[12].
- Defamation due diligence
- Apply truthfulness checks, contextualization, and public-interest framing using non-stigmatizing language; prepare remedies and documentation to mitigate civil/criminal risk under Polish law ^[15].
- Audit and logging
- Maintain tamper-evident logs of provenance decisions, corrections, LIA/DPIA outcomes, access controls, and security events, without logging unnecessary personal data.

6) Category subpages and narrative signaling (practical setup)

- Manipulacje i Toksyczność
- Focus: manipulation psychology, toxic relationships, narcissism; use neutral, precise terms; include content warnings for distressing material. Pair supportive icons with explicit text labels for clarity [2].
- Instrumentalne wykorzystanie instytucji
- Focus: ways institutions supported or amplified manipulation; show clear procedural context and cross-link to evidence and reports.
- Dokumenty i Fakty
- Focus: key facts, official documents, evidence appraisal, “5 faktów”; provide immutable viewers, hashes, and canonical downloads.
- Raporty i Analizy
- Focus: media-facing reports and critical analyses (including “kronika” when synthetic), with abstracts and links back to underlying evidence.
- Symbolika i Narracja
- Focus: “Kłątwa,” chronicles, and broader psychological framing; avoid sensationalism, add content warnings, and link to primary documents for grounding. Maintain minimal motion and predictable layouts for sensitive content [1].

7) Audio module consolidation (per-article playback and playlist)

- Per-item player
- Place “► Posłuchaj analizy” above the body. Provide play/pause, seek, volume/mute, playback rate, and time display; no autoplay. Ensure keyboard operability and visible focus, and meet non-text contrast requirements [10].
- Transcripts
- Provide a toggleable, on-page transcript and a downloadable text file. Label if auto-generated and note human review status. Make transcripts searchable to aid discovery and accessibility [10].
- Playlist
- Auto-scan /audio/ to map tracks by filename or item ID and expose a tracker-free playlist view. The playlist uses the same accessible player and does not set third-party cookies [9].

8) Pre-launch QA and go-live checklist

- Accessibility (WCAG 2.2 AA)
- All controls keyboard-operable; skip links and landmarks present; visible focus; target sizes $\geq 24 \times 24$ CSS px; non-text contrast compliant; accordions use button semantics with aria-expanded; language attributes set correctly [6]. If accordions are used, ensure descriptive headings, multiple-open capability, and accessible state change announcements [8].
- Navigation and orientation
- Visible primary navigation on desktop, no hover-only menus, click-activated submenus or routing pages, canonical breadcrumbs starting with Home and ending with current page as non-link [2]. Breadcrumbs reflect hierarchy, not history [5].
- Search and filtering
- Persistent search box; relevant facets with batch apply on mobile; clear undo/reset; show only available facet values; diacritics-aware indexing [17].
- Performance and stability
- Server-rendered pages with critical CSS; EU CDN; Brotli/Gzip; cache and preconnect; reserve layout space; defer non-critical scripts; lazy load transcripts and related modules.
- Security hardening

- HTTPS with HSTS; strong CSP limiting scripts and media to self-hosted origins; X-Content-Type-Options, Referrer-Policy, frame-ancestors; sanitize inputs; rate-limit search; enforce MFA and least-privilege roles in CMS.
- Evidence integrity
- Compute and publish SHA-256 for all nine .txt files; show last-verified timestamps; release a signed Integrity manifest.
- Legal and ethics
- Publish Polish-language Privacy/Cookies, Corrections/Right-of-Reply, Editorial Standards, and Methodology/Provenance pages; complete LIA and any DPIA; validate consent requirements for any private-life data or audio ^[13].
- Schema.org and SEO
- Emit CreativeWork/NewsArticle per item, BreadcrumbList for canonical trails, localized sitemaps; robots excludes sensitive endpoints.
- Multilingual
- Verify PL/EN toggle preserves current context; check localized dates/numbers; confirm language attributes and screen reader announcements; test search with and without diacritics ^[6].

9) Launch roadmap (phased execution)

- Phase A: Content ingestion and integrity
- Import all nine files; assign canonical categories and breadcrumbs; compute hashes; draft provenance notes; configure related links.
- Phase B: UX and audio integration
- Build category landing pages and item views with immutable viewers; add “► Posłuchaj analizy” and transcripts; configure /audio/ playlist mapping.
- Phase C: Compliance and trust scaffolding
- Publish Privacy, Cookies, Editorial Standards, Corrections, and Methodology/Provenance pages; complete LIA/DPIA; appoint and disclose DPO contact ^[13].
- Phase D: Performance, security, and analytics
- Apply security headers and CSP; configure EU CDN; deploy cookieless analytics; implement consent banner for optional cookies; run WCAG/performance audits ^[6].
- Phase E: Multilingual and final QA
- Implement PL/EN localization; validate search and facets; finish schema.org; conduct user testing with Polish-speaking audiences; remediate and sign off.

10) Acceptance criteria (go/no-go)

- Preservation: All nine .txt files are published verbatim with canonical downloads, visible SHA-256 hashes, and last-verified timestamps on item pages and the Integrity page.
- Navigation: Visible desktop categories; accessible mobile drawer; canonical breadcrumbs; persistent header search; relevant facets with clear reset.
- Audio: “► Posłuchaj analizy” on every item; accessible player; transcript toggle and download; /audio/ playlist without third-party trackers ^[9].
- Accessibility and ethics: WCAG 2.2 AA conformance verified; content warnings present; serious, analytical tone; icons always paired with text labels ^[1].
- Legal/privacy: Polish-language governance pages live; right-of-reply workflow operational per Press Law; LIA/DPIA documented; minimal, opt-in analytics and cookie consent in place ^[12].

- Performance/security: Core Web Vitals within targets; CSP and security headers active; rate-limited search; EU hosting/CDN; audit logs enabled without unnecessary PII.
- Multilingual: Polish primary with optional English; language semantics correct; diacritics-aware search; localized formats and schema.

Outcome This blueprint operationalizes governance, evidence integrity, multilingual UX, and EU/Polish compliance while honoring the no-deletion constraint. It organizes the nine source files across five categories with canonical breadcrumbs and curated cross-links, delivers accessible per-article audio with transcripts, and specifies trust, consent, and security practices for a responsible public launch of a modern investigative portal ^[2].

References

- [1] [Homepage Design: 5 Fundamental Principles - NN/G](#)
- [2] [Menu-Design Checklist: 17 UX Guidelines - NN/G](#)
- [3] [Information Architecture Articles & Videos - Niels...](#)
- [4] [\[PDF\] Reuters Institute Digital News Report 2025](#)
- [5] [Breadcrumbs: 11 Design Guidelines for Desktop and ...](#)
- [6] [WCAG 2.2 Guidelines - Penn State | Accessibility](#)
- [7] [Table of Contents: The Ultimate Design Guide - NN/...](#)
- [8] [Accordions on Desktop: When and How to Use - NN/G](#)
- [9] [Audio and Video | Web Accessibility Checklist - De...](#)
- [10] [Web Content Accessibility Guidelines \(WCAG\) 2.2 - ...](#)
- [11] [\[PDF\] Guidelines 1/2024 on processing of personal ...](#)
- [12] [\[PDF\] Translation of Polish Press Law Act - Sciend...](#)
- [13] [Data Protection Laws and Regulations in Poland - C...](#)
- [14] [How can data subjects exercise their rights in the...](#)
- [15] [Defamation in Poland | Dudkowiak & Putyra](#)
- [16] [Mega Menus Work Well for Site Navigation - NN/G](#)
- [17] [How Faceted Search Improves Customer Experience \[+...](#)
- [18] [Faceted search: 9 best practices to improve UX and...](#)
- [19] [Faceted Search: An Overview - Algolia](#)
- [20] [The Hamburger-Menu Icon Today: Is it Recognizable?...](#)
- [21] [WAI-ARIA Authoring Practices 1.2 - W3C](#)
- [22] [Threats to Media Freedom and Pluralism in Poland A...](#)
- [23] [Personal data breaches and notification – viewpoin...](#)
- [24] [WCAG Checklist: A Simplified Guide to WCAG 2.2 AA ...](#)