

Article

A UDL–ICF Framework for Inclusive Artificial Intelligence in Cultural Heritage: Exploratory Insights from an Erasmus+ Blended Intensive Programme

Maria Concetta Carruba ^{1,*}  and Anna Siri ^{2,*}¹ Department of Education and Sport Sciences, Università Pegaso, 80143 Naples, Italy² Department of Psychology and Health Sciences, Università Pegaso, 80143 Naples, Italy

* Correspondence: mariaconcetta.carruba@unipegaso.it (M.C.C.); anna.siri@unipegaso.it (A.S.)

Abstract

Cultural heritage institutions continue to face significant challenges in addressing neurodivergent access needs, while a principled framework linking inclusive design and artificial intelligence in heritage remains underdeveloped. This article proposes and explores the transfer of a dual Universal Design for Learning (UDL 3.0)–International Classification of Functioning, Disability and Health (ICF) framework from education to cultural heritage. The framework was examined within an Erasmus+ Blended Intensive Programme in Naples, Italy, through an exploratory mixed-methods pilot design combining pre/post questionnaires with thematic analysis of participants' open-ended reflections. Findings suggest that the programme supported a qualitative reframing of accessibility from a predominantly physical-access model toward a broader understanding encompassing sensory, cognitive, communicative, and participatory dimensions. Site visits to Pompeii, Herculaneum, and the MANN emerged as especially important in prompting critical reflection on the gap between heritage value and accessibility practice. Quantitative patterns also indicated preliminary gains in participants' self-efficacy in applying UDL/ICF principles and in reflecting on AI ethics in heritage, although these findings remain exploratory given the small sample size and pilot design. The article contributes a conceptual and pedagogical framework for inclusive AI in cultural heritage and discusses its implications for accessibility governance, conservation-constrained settings, and future co-designed research.

Keywords: neurodiversity; universal design for learning; international classification of functioning; cultural heritage accessibility; artificial intelligence; transformative learning; mixed-methods; inclusive design; museum accessibility; digital heritage; higher education; Erasmus+



Academic Editor: Márcia Vilarigues

Received: 25 April 2026

Revised: 27 June 2026

Accepted: 6 July 2026

Published: 8 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

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

1. Introduction

Cultural heritage institutions—museums, archaeological parks, archives, and historic monuments—are obligated by their canonical mission, as reformulated in the International Council of Museums' 2022 definition, to serve society, foster diversity, and ensure that collections and spaces are accessible to all. Yet the material conditions of most heritage environments systematically contradict this aspiration. The architecture of galleries and archaeological parks, the conventions of interpretive communication, the norms of visitor management, and the cultures of professional practice have been designed—largely unconsciously, but with measurable effect—around an implicitly neurotypical, able-bodied, and educationally dominant visitor. Between 15 and 20 per cent of the global population

presents neurodivergent profiles, including autism spectrum conditions (ASC), attention deficit hyperactivity disorder (ADHD), dyslexia, dyspraxia, and intellectual disabilities [1,2], and the systematic exclusion of this population from cultural participation constitutes a failure of both institutional design and democratic accountability.

From a sociological standpoint, this exclusion can be read as a form of symbolic violence [3] and cultural stratification: heritage institutions may operate as gatekeepers of legitimate culture and taste, implicitly marginalising neurodivergent ways of sensing, interpreting, and moving through space. Accessibility is, therefore, not merely a technical or legal problem but also a question of social justice, cultural capital distribution, and the politics of belonging.

The empirical scale of this failure is now well-documented. Fletcher's doctoral study, involving 466 autistic and neurodivergent (AuND) adults and 130 museum workers across the United Kingdom, established that approximately half of neurodivergent respondents actively avoid museum visits due to sensory overload or underload, while the overwhelming majority would attend more frequently if institutions were more accessible [4]. Granland, Sadia, and Cooper, in a mixed-methods study involving 281 respondents across 23 countries, found that design preferences among neurodivergent visitors were more strongly associated with individual sensory profiles than with specific diagnostic categories [2], reinforcing the argument, grounded in the WHO International Classification of Functioning, Disability and Health (ICF) [5], that functional profiling rather than diagnostic labelling should drive accessibility design. Hutson and Hutson identify extended reality, digital storytelling, and sensory-mapping technologies as having measurable potential to create new forms of inclusive heritage engagement [6].

Recent bibliometric analysis further suggests a rapid expansion of research on artificial intelligence in museums, particularly across domains such as visitor experience, accessibility and inclusion, digitisation, and ethical governance [7]; however, the absence of a principled framework connecting barrier research to AI deployment represents a significant gap in the field [8,9].

The study addresses this gap by exploring the applicability of a dual UDL 3.0–ICF framework [5,10]—previously applied in formal education in three studies: (1) AI for intellectual disabilities [11]; (2) mixed reality for special needs [8]; and (3) mobile accessibility [12]—to heritage AI inclusion. This study systematically transfers the framework to cultural heritage via an Erasmus+ BIP (Naples, Italy, 2026), combining theory with site visits to UNESCO sites where conservation limits favour AI solutions.

The study is organised around one main research question and three related sub-questions.

Main Research Question: To what extent does participation in the BIP appear to be associated with changes in how students understand neurodiversity, inclusive design, and AI ethics in cultural heritage?

This question is explored through three related sub-questions:

- (1) Does participation in the programme appear to be associated with changes in students' knowledge, attitudes, and self-efficacy in relation to neurodiversity, UDL/ICF, and AI ethics in heritage?
- (2) Do participants describe qualitative shifts in how they understand accessibility in heritage settings, particularly beyond narrow physical-access models?
- (3) How do participants describe the role of site visits and theory–practice integration in shaping these changes in perspective?

This question is explored through a concurrent mixed-methods design in which quantitative measures capture changes in self-efficacy and declarative knowledge, while quali-

tative analysis examines shifts in conceptual understanding as expressed in participants' open-ended reflections (n = 21 PRE, n = 17 POST, n = 14 matched pairs).

While the study is grounded in neurodiversity research, inclusive design, and AI ethics, it also contributes directly to heritage scholarship by engaging three intersecting debates: visitor experience in heritage settings, museum accessibility as an institutional and interpretive practice, and critical heritage perspectives on inclusion, legitimacy, and participation. In doing so, the paper positions accessibility not as an ancillary adjustment, but as a constitutive dimension of heritage experience itself, and reframes AI as a potential mediator between institutional constraints, conservation requirements, and the diversification of audiences.

This positioning responds to a growing need within heritage studies to move beyond compliance-based models of accessibility toward structurally embedded, experience-oriented, and ethically grounded approaches to inclusion within museums, archaeological sites, and other heritage environments.

Recent case-based research also suggests that accessibility in heritage settings should be understood not only as a technical adjustment but also as a participatory, sensory, and institutionally embedded process of transformation, even in conservation-constrained environments [13].

Beyond its contribution to inclusive AI and accessibility research, this study speaks directly to heritage scholarship by addressing three interrelated concerns: visitor experience, interpretive accessibility, and institutional participation. From this perspective, neurodivergent inclusion does not sit outside the mission of heritage institutions; it challenges how heritage spaces define legitimate participation, interpretive norms, and the figure of the "expected" visitor.

Throughout this article, the UDL–ICF framework is read not only as a pedagogical instrument but also through a sociological lens—drawing on Bourdieu's concepts of field and habitus to examine how heritage institutions reproduce normative assumptions about legitimate participation—an argument that this study can only begin to develop empirically but which is elaborated in the Discussion and Conclusions.

2. Theoretical Background

2.1. *The Neurodiversity Paradigm and the ICF*

The neurodiversity paradigm, emerging from the disability rights movement in the 1990s through the work of Singer [1] and the broader autistic self-advocacy movement, holds that neurological profiles associated with autism, ADHD, dyslexia, dyspraxia, and related conditions are not deviations from a biological norm requiring remediation but forms of natural human variation. The locus of disability is not the individual but the mismatch between diverse functional profiles and environments designed around a fictitious neurotypical average [14].

Sociologically, this aligns with Goffman's stigma theory [15]: environments that presuppose a neurotypical visitor may subtly mark neurodivergent individuals as "deviant" or "out of place." The very layout of a gallery—quiet, linear, text-heavy, sensorially muted—operates as an unmarked norm that produces exclusion without explicit intent. In this sense, applying the ICF framework in heritage also enables a sociological reading of the built environment as historically contingent rather than neutral.

The WHO ICF [5] operationalises this insight analytically through a biopsychosocial model mapping disability as the dynamic interaction between body functions and structures, activity limitations, participation restrictions, and environmental and personal contextual factors. A heritage institution applying ICF-based analysis asks not what is pathologically wrong with the visitor but what environmental features constitute barriers.

ers for which functional profiles and what modifications would convert those barriers into facilitators. The ICF distinguishes between barriers—environmental factors that limit functioning—and facilitators—factors that enhance it—enabling systematic mapping of physical, sensory, cognitive, communicative, and social dimensions of heritage accessibility. This dual ICF–UDL approach has been validated in formal education contexts [8,11,12] and is here extended to cultural heritage for the first time.

2.2. *Heritage, Visitor Experience, and Institutional Accessibility*

Heritage accessibility is not reducible to the removal of physical barriers, nor can it be understood solely as a technical compliance issue. In heritage settings, accessibility is closely intertwined with visitor experience, interpretive design, and institutional culture. Visitor studies have long shown that engagement with museums, archaeological parks, and historic sites is shaped not only by collections or narratives, but also by navigational legibility, sensory conditions, communicative framing, and the degree to which visitors perceive themselves as recognised and legitimate participants in the heritage space. Accessibility, in this sense, is part of the experience of heritage itself rather than an external accommodation added afterwards. This broader experiential perspective is also consistent with recent work in heritage education and informal learning, which treats cultural participation as a situated, interactive, and place-based process rather than as simple content transmission.

This perspective is particularly relevant for neurodivergent visitors, whose encounters with heritage environments may be shaped by sensory overload, ambiguous signage, rigid interpretive formats, or socially demanding visitation norms. Museum accessibility scholarship has increasingly expanded the meaning of inclusion beyond ramps and physical entry to encompass sensory, cognitive, communicative, emotional, and social dimensions of participation. Such an expansion is especially important in conservation-constrained environments, where the physical modification of sites may be limited and where inclusive interpretation, digital mediation, and staff practices become crucial components of access.

In heritage contexts, accessibility is inseparable from interpretation: the ability to orient oneself within space, decode narratives, manage sensory conditions, and feel recognised as a legitimate participant all shape how heritage is experienced and valued. For this reason, accessibility is not external to heritage mediation but embedded within the processes through which heritage is communicated, interpreted, and socially authorised.

Critical heritage scholarship further reinforces this argument by showing that heritage institutions are not neutral containers of culture but selective environments in which legitimacy, authority, and belonging are actively organised. What counts as an appropriate visitor, a valid mode of engagement, or an acceptable form of interpretation is shaped by institutional assumptions and professional norms. From this perspective, the exclusion of neurodivergent visitors is not simply an accidental design failure but also a manifestation of deeper normative assumptions about whose ways of sensing, learning, and participating are expected within heritage space. This study builds on that insight by examining whether AI, when designed through UDL and ICF principles, can help shift accessibility from a compensatory logic toward a structurally embedded and experience-oriented model of inclusion.

2.3. *Universal Design for Learning, Version 3.0*

Universal Design for Learning is a research-based framework developed by CAST and grounded in the neuroscience of learning [10]. Version 3.0 (2024) organises its principles around three brain networks—affektive (the ‘why’ of learning), recognition (the ‘what’), and strategic (the ‘how’)—and operationalises them through three core principles (Multiple Means of Engagement, Representation, and Action and Expression), nine guidelines, and

numerous specific design checkpoints. The most significant advances in version 3.0 are its explicit commitments to equity, anti-bias design, the cultivation of belonging (Guideline 8.4), and the challenging of exclusionary practices (Guideline 6.5). These commitments position UDL not merely as a design methodology but as a social justice framework, directly continuous with the neurodiversity paradigm's critique of normative cognitive assumptions. For this research, UDL serves two functions: at the theoretical level, it provides the principled design architecture for inclusive AI in cultural heritage; at the empirical level, it provides the primary analytical instrument for evaluating heritage AI tools and measuring the BIP's educational impact.

2.4. AI as Assistive Intelligence and the Personalisation–Paternalism Tension

CAST's formulation of AI as 'Assistive Intelligence'—explicitly distinguished from traditional assistive technology—provides a crucial conceptual anchor [10]. Whereas traditional assistive technology supplements a baseline experience designed without the assisted user in mind, Assistive Intelligence embeds cognitive load reduction and flexible access in the baseline design itself, converting the standard heritage experience into a universally accessible one. This reframes the central design question from 'how do we accommodate neurodivergent visitors?' to 'how do we design AI systems such that the standard heritage experience is inherently flexible for the full range of neurological diversity?' This framing also makes explicit the critical distinction between genuine personalisation—transparent, informed, and revocable—and algorithmic paternalism, which classifies users covertly through behavioural inference and adapts content without their knowledge or consent. The present research distinguishes these modes operationally by four criteria: transparency of the adaptation mechanism; user control over adaptation parameters; accessibility of the default (unadapted) experience; and the absence of diagnostic labelling in the user model. A concrete application example clarifies the distinction. A visitor approaching a display case at Pompeii may activate a smartphone application that offers 'simplified' or 'standard' interpretive text through an explicit, user-initiated toggle; this constitutes genuine personalisation under the present framework. By contrast, a system that infers neurodivergent status from gait, dwell time, or interaction patterns and silently routes the visitor to reduced-complexity content without disclosure constitutes algorithmic paternalism, regardless of whether its effect on comprehension is positive. The ethical distinction is not the outcome but the agency: UDL Guideline 6.2 (support planning and strategy development) requires that learners—and by extension, visitors—are positioned as active agents of their own access, not passive recipients of classification-driven accommodation [10].

2.5. Transformative Learning Theory

The educational research strand draws on Mezirow's transformative learning theory [16], which holds that meaningful adult learning involves the critical examination and revision of prior meaning-making frameworks. Mezirow identifies 'disorienting dilemmas'—encounters with experiences or evidence that cannot be assimilated into existing interpretive frames—as the catalytic events of transformative learning, triggering critical reflection and, potentially, fundamental perspective transformation [16]. The BIP's design—combining theoretical instruction in neurodiversity and UDL with direct experiential encounters at heritage sites whose inaccessibility can be observed and analysed in real time—creates precisely the conditions that transformative learning theory identifies as generative. This theory is deployed both as an analytical lens for interpreting open-ended qualitative data and as a design rationale for the lecture and exercises. In this exploratory pilot study, references to transformative learning should be understood as applying to

immediate, situated changes in perspective as reported by participants, not as evidence of durable transformation (which would require longitudinal follow-up).

2.6. Policy and Regulatory Context

The theoretical framework is situated within a convergent policy context. The European Accessibility Act (Directive 2019/882) [17] mandates accessibility standards for digital products and services with direct implications for heritage institutions' digital tools. The EU AI Act (Regulation 2024/1689) [18], which entered into force in 2024, establishes a risk-based governance architecture with potential relevance for AI systems deployed by museums, archaeological parks, and other public-facing heritage institutions. The UN Convention on the Rights of Persons with Disabilities (CRPD, 2006) [19], ratified by all EU member states, requires full and effective participation of persons with disabilities in cultural life. In the Italian context, the conservation constraints on UNESCO World Heritage Sites [20]—including Pompeii, Herculaneum, and the holdings of the National Archaeological Museum of Naples (MANN)—limit physical modification of spaces, making AI-based non-invasive interventions particularly strategically important as accessibility solutions. The sociological dimensions of policy implementation—including institutional inertia, professional socialisation, and the gap between formal compliance and substantive inclusion—are addressed in the Discussion.

BSI. PAS 6463 [21] provides the most operationally specific guidance currently available for neurodiversity-informed environmental design and was used as a reference framework in the BIP practical exercises.

3. Materials and Methods

3.1. Study Design

The study adopts a pilot concurrent mixed-methods design [22], combining quantitative pre-post measurement with qualitative investigation of students' learning processes within an Erasmus+ Blended Intensive Programme (BIP). The quantitative strand uses a within-subjects pre-post approach without a control group, explicitly framed as quasi-experimental and exploratory and intended to capture directional patterns of change rather than to support causal inference. The design prioritises feasibility piloting, hypothesis generation, and integration of strands at interpretation, not efficacy evaluation. The qualitative strand employs inductive thematic analysis [23] of open-ended questionnaire responses, followed by a second-pass theoretical coding informed by Mezirow's transformative learning framework [16]. The two strands are integrated at the interpretive stage: the quantitative findings provide a descriptive account of change across selected items, while the qualitative findings offer explanatory depth and insight into dimensions of learning not reducible to ordinal measures. The study is explicitly framed as exploratory and hypothesis-generating and should be understood as a pedagogical pilot rather than a confirmatory evaluation of programme effectiveness or of the framework itself. Findings are therefore context-bound and should not be generalised beyond the BIP setting without further replication.

3.2. Participants

The target population comprised all students enrolled in the Erasmus+ Blended Intensive Programme (BIP) *Ethics in Social Science and Humanities* during the physical mobility week in Naples. The PRE questionnaire was completed by 21 respondents, and the POST questionnaire by 17; longitudinal matching through self-generated participant codes yielded 14 paired observations. Among the four PRE respondents who did not complete the POST questionnaire, no evident differences were observed in demographic profile or

baseline knowledge scores relative to the matched sample, although the small numbers preclude strong conclusions about attrition patterns.

The cohort was predominantly composed of undergraduate students (14/21, 67%) in humanities disciplines (History, Archaeology, Art History, Philosophy, and Literature), primarily from the University of Wrocław (Poland). Additional respondents included MA students (3/21, 14%), doctoral students (2/21, 10%), and two academic staff members. One respondent from IPB Bragança (Portugal) contributed some disciplinary diversity in education and pedagogy.

The sample was therefore relatively homogeneous, consisting mainly of humanities students from a single institutional context. This homogeneity limits the extent to which variation in prior exposure to disability studies, accessibility, and digital technologies could be captured. Participants from technical or social science backgrounds, heritage practice, or with lived neurodivergent experience might reasonably be expected to show different baseline orientations and learning trajectories.

The relatively small sample size is typical of BIP-format programmes but constrains the robustness of quantitative analysis. Quantitative findings are therefore treated as indicative rather than confirmatory, and the qualitative strand is particularly important in contextualising observed patterns of change.

3.3. Instrument

Data were collected through a purpose-built PRE/POST mixed-methods questionnaire developed specifically for this pilot study. The instrument was designed to capture baseline characteristics, selected knowledge domains introduced during the lecture, attitudinal and self-efficacy orientations, and qualitative reflections on inclusive AI and heritage accessibility. As the questionnaire was developed specifically for this educational context, it was not subject to prior psychometric validation and should therefore be understood as an exploratory measurement tool rather than a standardised scale.

The questionnaire comprised five item types: (a) demographic and background items ($n = 6$ PRE; analysed descriptively); (b) four-option multiple-choice knowledge items ($n = 8$ replicated across PRE and POST, plus one POST-only governance item), assessing concepts explicitly introduced during the lecture; (c) five-point Likert-scale attitude and self-efficacy items ($n = 9$ replicated; scored from 1 = Strongly Disagree to 5 = Strongly Agree; treated as ordinal data); (d) seven-point semantic differential concept-mapping pairs ($n = 9$ replicated); and (e) open-ended qualitative items ($n = 3$ PRE; $n = 4$ POST).

To enable anonymous longitudinal matching, participants generated self-identification codes using the first two letters of their mother's first name and their day of birth (e.g., MA14). This procedure allowed paired analysis without collecting directly identifying personal data. Replicated items were administered verbatim across PRE and POST; POST-only items were used where appropriate to capture programme-specific reflections (†).

The nine Likert items assessed were the ethical obligation of heritage institutions (L1); the superiority of the neurodiversity paradigm over the medical model (L2); AI's potential for accessibility (L3); self-efficacy in applying UDL/ICF (L4); the perception of algorithmic bias risk (L5); AI personalisation and agency (L6); self-efficacy in evaluating AI ethics in heritage contexts (L7); the epistemic justice obligation (L8); and digital accessibility equivalence with physical accessibility (L9). The eight replicated knowledge items (Q1–Q8) assessed were the neurodiversity paradigm versus the medical model (Q1); the ICF biopsychosocial model (Q2); UDL v3.0 equity commitments (Q3); the Fletcher [4] (empirical finding (Q4); AI as Assistive Intelligence (Q5); the co-design principle (Q6); prospect and refuge theory in museum design (Q7); and ISTAT digital competence statistics (Q8).

3.4. Procedure

The PRE questionnaire was administered digitally via Google Forms on the first day, before any programme content was delivered. The POST questionnaire was opened on the last day, immediately after programme closure, and remained available for five days. The lecture “Bridging Theory and Practice: Inclusive AI and its Impact on Cultural Heritage Accessibility” (50 min) constituted the primary educational intervention, introducing the dual UDL–ICF framework, the empirical barrier literature, AI as Assistive Intelligence, and the six-component governance framework. On-site visits to the Archaeological Park of Pompeii, the Archaeological Park of Herculaneum, and the MANN constituted the experiential component; three structured exercises guided students’ accessibility observation. Participation was voluntary and anonymous; completion constituted informed consent.

The lead researchers also served as the lecturers responsible for programme delivery and were therefore known to participants in a teaching capacity at the time of data collection. To mitigate potential power dynamics arising from this dual role, the voluntary nature of participation was communicated explicitly in the introductory text of both questionnaires and verbally at the programme opening: participants were informed that non-participation or withdrawal would have no academic or evaluative consequences whatsoever. Completion of the questionnaire was treated as constituting informed consent. The risk of social desirability bias associated with the lecturer–researcher dual role cannot be fully excluded and is acknowledged as a limitation in Section 6.

3.5. Analysis

Quantitative analysis was conducted using descriptive statistics and non-parametric inferential procedures, supported by Python for data organisation and analysis (Python Software Foundation, 2023, version 3.12; SciPy 1.13 Contributors, 2024, version 1.13). Qualitative analysis followed Braun and Clarke’s six-phase thematic analysis protocol [23,24]. Correct-response rates for knowledge items were compared as percentage-point differences between the full PRE ($n = 21$) and POST ($n = 17$) samples; the non-identical group sizes precluded McNemar’s test, and results are reported descriptively. For the matched sample ($n = 14$), Wilcoxon signed-rank tests were applied to each Likert item, treating the data as ordinal. Effect sizes are reported as Cohen’s d , computed from the pooled standard deviation of paired pre–post scores, and as rank-biserial correlation ($r = W/[n(n + 1)/2]$) as a non-parametric alternative. Significance threshold: $\alpha = 0.05$ (two-tailed); no correction for multiple comparisons was applied given the explicitly exploratory framing of the study.

Because nine Wilcoxon signed-rank tests were conducted, the risk of Type I error is increased. However, no correction for multiple comparisons (e.g., Bonferroni) was applied, given the explicitly exploratory and hypothesis-generating nature of this pilot study. Consequently, the interpretation of quantitative findings does not rely primarily on p -values. Instead, effect sizes—Cohen’s d and rank-biserial correlation (r)—are treated as the primary evidence of meaningful change, as they are less sensitive to sample size and are not affected by correction procedures. For the four items showing the largest effects, the effect sizes indicate substantial change: L4 (confidence in applying UDL/ICF, $d = 1.21$, $r = 0.82$), L7 (ethical preparedness for AI governance, $d = 0.82$, $r = 0.78$), L3 (AI’s potential for accessibility, $d = 0.75$, $r = 0.74$), and L9 (digital accessibility equivalence, $d = 0.85$, $r = 0.71$). These values suggest meaningful directional shifts even though none of the observed p -values would remain statistically significant under a Bonferroni correction (adjusted $\alpha = 0.05/9 \approx 0.006$). Readers are therefore encouraged to interpret the quantitative results with attention to effect sizes and in conjunction with the qualitative findings, which provide convergent evidence of learning change.

The uncorrected results are reported because the study is designed to generate rather than confirm hypotheses and because effect sizes—which are insensitive to sample size and correction procedures—are treated as the primary evidence; all *p*-values should be interpreted with this caveat. Qualitative analysis followed the six-phase thematic analysis protocol of Braun and Clarke [23,24]: familiarisation; initial coding; theme generation; theme review; theme definition and naming; and write-up. A total of 84 substantive open-ended responses were analysed (defined as responses exceeding ten words with thematic content; 12 responses of fewer than ten words or purely administrative content were excluded). The analytical process proceeded as follows: (1) familiarisation involved repeated reading of all 84 responses before any coding; (2) initial coding produced *n* = 147 codes across both administrations, generated inductively without prior category imposition; (3) candidate themes were grouped from codes and reviewed against the full dataset; (4) seven themes were retained after two review passes, with three PRE and four POST themes; (5) each theme was defined with a label and a descriptive narrative; and (6) the write-up integrated illustrative verbatim quotes selected to represent the theme's central pattern, not its most extreme instance. A second analytical pass applied theoretical coding against Mezirow's framework [16], examining evidence of disorienting dilemmas, critical reflection, and perspective transformation. The analyst kept a reflexivity log documenting interpretive decisions throughout; the full audit trail is available on request. As a single-analyst study, inter-rater reliability was not assessed.

3.6. Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki (revised 2013). All participation was voluntary, anonymous, and without evaluative consequences. No personally identifying information was collected. As a non-clinical, anonymous educational survey of competent adult volunteers, formal IRB approval was not required under Italian national legislation (D.Lgs. 196/2003 as amended by D.Lgs. 101/2018 implementing GDPR); the study design was reviewed and approved by the Erasmus+ BIP academic coordinating committee. Participant data are stored in encrypted form and will be destroyed five years after publication.

4. Results

4.1. Baseline Sample Characteristics

The PRE cohort (*n* = 21) was characterised by very low prior familiarity with the study's key theoretical frameworks: 81% (17/21) had never encountered UDL; 52% (11/21) had never encountered the ICF; and 67% (14/21) reported no prior formal exposure to disability studies, inclusive education, or accessibility. Nineteen of 21 respondents identified their primary discipline as humanities. Prior experience with cultural heritage in academic or professional capacities varied: six respondents had no such experience; seven had visited once or twice; five regularly; and three very regularly. These baseline data confirm the expected low-familiarity profile: any observed knowledge gains are interpretable as novel learning rather than prior-knowledge retrieval.

4.2. Knowledge Change

Table 1 presents correct-response proportions for each knowledge item at PRE and POST. Three items showed substantive positive change. Item Q1 (neurodiversity paradigm versus medical model) recorded the largest improvement: from 38% at PRE to 88% at POST (+50 percentage points), suggesting that the lecture effectively communicated a core epistemological distinction introduced in the programme. Q4) improved from 48% to 76% (+28 pp) [4]; Q2 (ICF biopsychosocial model) improved from 57% to 82% (+25 pp). Item

Q3 (UDL v3.0 equity commitments) showed only marginal gain (10% → 18%), remaining the most poorly understood item across both administrations, attributed to the conceptual proximity of the correct answer to the dominant distractor rather than to a failure of instruction.

Table 1. Knowledge item performance. pp = percentage points. Correct answers determined by lecture content.

Item	Topic	Correct Answer	PRE (%)	POST (%)	Δ
Q1	Neurodiversity paradigm	B: Natural variation	38	88	+50 pp
Q2	ICF biopsychosocial model	B: Environmental interaction	57	82	+25 pp
Q3	UDL v3.0 equity commitments	C: Anti-bias/challenging exclusion	10	18	+8 pp
Q4	Fletcher finding [4]	C: ND adults would visit more	48	76	+28 pp
Q5 †	AI as Assistive Intelligence	B: Embedded in baseline design	52	12	−40 pp
Q6 †	Nothing About Us Without Us	B: ND communities hold authority	57	12	−45 pp
Q7 †	Prospect and refuge theory	B: Balance legibility with retreat	71	18	−53 pp
Q8	ISTAT digital competence	C: Approx. 26.5% of lower-education adults	52	POST-only	—

Note: † Items showing either minimal improvement (Q3) or an anomalous decline (Q5–Q7). The decline in Q5–Q7 is interpreted as a methodological artefact (selective attrition, shift in response strategy, item placement); it does not indicate genuine knowledge loss. Note: PRE = pre-programme questionnaire; POST = post-programme questionnaire; ND = neurodivergent. Percentage-point differences are reported descriptively because the PRE and POST knowledge subsamples were not identical.

Items Q5 (AI as Assistive Intelligence), Q6 (Nothing About Us Without Us), and Q7 (prospect and refuge theory) showed counterintuitive declines of 40, 45, and 53 percentage points respectively. Given the exploratory nature of this pilot and the small sample size, all knowledge item changes—whether positive or negative—should be interpreted with caution. The declines observed for Q5–Q7 are not interpreted as learning regression; rather, they are treated as potential artefacts of item placement, selective attrition, or increased awareness of complexity. For three converging reasons. First, the POST questionnaire was completed at the end of an intensive five-day programme, and the knowledge block appeared early in the instrument after more emotionally engaging open-ended prompts. Second, selective attrition is evidenced: two of the four PRE-only respondents had answered Q5–Q7 correctly at PRE, inflating the PRE denominator for these items and suppressing the apparent POST proportion independently of any instructional effect. Third, a response-strategy shift is consistent with programme learning: the lecture’s introduction of conceptual nuance for each of these three topics (AI typologies, participatory research ethics, environmental design theory) may have disrupted heuristic answering without having time to consolidate correct knowledge under time-constrained POST conditions. Most diagnostically, the qualitative evidence of genuine conceptual enrichment (Section 4.5)—including the deployment of programme-specific vocabulary by POST respondents who presumably answered Q5–Q7 incorrectly—does not support interpreting the observed declines as learning regression, though it cannot rule it out definitively: the qualitative data are simply not the kind of evidence that would settle this question either way. By the same logic, the positive changes observed for Q1, Q2, and Q4 should also be viewed as indicative rather than conclusive, pending replication with a larger sample and improved instrument design. These items are flagged as requiring instrument revision in future administrations:

Q5 should be redesigned to reduce the plausibility of distractors that describe AI as a supplementary rather than embedded tool; Q6 and Q7 should be repositioned later in the questionnaire and administered under less time-pressured conditions.

4.3. Attitudinal and Self-Efficacy Change

Table 2 presents Wilcoxon signed-rank test results for all nine replicated Likert items ($n = 14$ matched pairs). Four items reached statistical significance. L4 (self-efficacy for applying UDL/ICF frameworks) showed the largest and most robust change ($\Delta = +1.50$, PRE M = 2.36, SD = 1.15; POST M = 3.86, SD = 0.75; $W = 5.5$, $p = 0.014$, $d = 1.21$, $r = 0.82$). The near-zero POST standard deviation indicates convergence toward a consistently moderate-to-high confidence level, irrespective of disciplinary background. L7 (self-efficacy for evaluating AI heritage tools using governance frameworks including the GDPR and EU AI Act) showed a large and significant effect ($\Delta = +1.00$, PRE M = 2.36, SD = 1.15; POST M = 3.36, SD = 1.08; $W = 4.5$, $p = 0.008$, $d = 0.82$, $r = 0.78$), consistent with the lower baseline and the specificity of governance framework content. L3 (AI's potential for heritage accessibility) showed a significant positive shift ($\Delta = +0.79$, $W = 0.0$, $p = 0.015$, $d = 0.75$, $r = 0.74$), representing movement from the scale midpoint to the agreement range. L9 (digital accessibility equivalence with physical) also reached significance ($\Delta = +0.64$, $W = 3.0$, $p = 0.030$, $d = 0.85$, $r = 0.71$).

Table 2. Wilcoxon signed-rank test results for Likert items, $n = 14$ matched pairs. Scale 1–5 (1 = Strongly Disagree, 5 = Strongly Agree). * $p < 0.05$; ** $p < 0.01$; n.s. = not significant; trend = $p < 0.15$. d = Cohen's d ; r = rank-biserial correlation.

Item	PRE M (SD)	POST M (SD)	Δ	W	p	d	r	Sig.
L1 Ethical obligation	4.29 (0.91)	4.29 (0.99)	+0.00	10.0	0.915	0.00	0.04	n.s.
L2 ND paradigm	3.36 (0.63)	3.71 (0.73)	+0.36	8.0	0.132	0.52	0.35	trend
L3 AI accessibility	3.00 (0.96)	3.79 (0.89)	+0.79	0.0	0.015	0.75	0.74	*
L4 Confidence UDL/ICF	2.36 (1.15)	3.86 (0.75)	+1.50	5.5	0.014	1.21	0.82	*
L5 Algorithmic bias	3.71 (0.73)	3.86 (0.86)	+0.14	18.0	0.564	0.17	0.13	n.s.
L6 AI personalisation	3.21 (1.12)	3.21 (0.89)	+0.00	32.5	0.963	0.00	0.01	n.s.
L7 Ethical preparedness	2.36 (1.15)	3.36 (1.08)	+1.00	4.5	0.008	0.82	0.78	**
L8 Epistemic justice	3.93 (0.83)	4.14 (0.66)	+0.21	9.5	0.429	0.26	0.23	n.s.
L9 Digital access	3.86 (0.77)	4.50 (0.52)	+0.64	3.0	0.030	0.85	0.71	*

Five items did not reach significance. L1 (ethical obligation, $p = 0.915$, $d = 0.00$) showed a ceiling effect (PRE M = 4.29). L8 (epistemic justice, $p = 0.429$, $d = 0.26$) showed a similar pattern (PRE M = 3.93). L5 (algorithmic bias risk, $p = 0.564$, $d = 0.17$) was near-ceiling at baseline (M = 3.71). L6 (AI personalisation and agency, $p = 0.963$, $d = 0.00$) showed no change; this null result is substantively informative, suggesting that the personalisation–paternalism distinction requires more extended instruction than a single lecture session provides. L2 (neurodiversity paradigm more useful than medical model, $p = 0.132$, $d = 0.52$) showed a trend in the positive direction.

4.4. Programme Experience and Learning Impact

POST respondents reported uniformly high programme evaluations: 53% (9/17) rated the programme Excellent, 41% (7/17) Very Good, and 6% (1/17) Good; no respondent rated it below Good. Seventy-one per cent (12/17) reported that the on-site visits substantially deepened their understanding of heritage accessibility challenges; 29% (5/17) reported moderate deepening. Sixty-five per cent (11/17) rated the theory–practice integration as Excellent; 24% (4/17) as Good. Regarding post-programme conceptual change, 29%

(5/17) reported substantial deepening; 35% (6/17) moderate extension; 29% (5/17) slight expansion; one respondent (6%) reported no change, consistent with pre-existing expertise. Eighty-eight per cent (15/17) expressed Likely or Very Likely intentions to apply the programme's frameworks in future academic or professional work. POST respondents identified sensory and environmental barriers as most urgently requiring intervention at Italian heritage sites (47%, 8/17); informational and communicative barriers (35%, 6/17); and structural and governance barriers (18%, 3/17).

4.5. Qualitative Themes

Inductive thematic analysis of 84 substantive open-ended responses yielded seven themes. Three PRE themes characterised the baseline knowledge state; four POST themes evidenced transformative learning. Each theme is designated with a label following the convention of Braun and Clarke [24].

4.5.1. PRE Theme 1: Inclusive Design as Undifferentiated Universal Access

The dominant PRE response pattern reflected an intuitive but analytically undifferentiated understanding of inclusive design. Respondents typically described inclusive design in terms of broad universal openness ("access for everyone", "open for all human beings"), but without differentiating among physical, sensory, cognitive, communicative, and participatory dimensions of accessibility, and without addressing co-design or community involvement. One doctoral respondent provided a notably more sophisticated answer, explicitly referring to co-creation and anti-tokenism; this response was treated as a negative case within the thematic analysis.

4.5.2. PRE Theme 2: Sensory Overload as the Paradigmatic Barrier

Respondents who offered substantive answers to the barrier-identification question converged with notable consistency on sensory and environmental barriers, including crowding, noise, overstimulation, inadequate ventilation, and unsuitable lighting. This spontaneous convergence with the empirical literature [2,4] is particularly striking because respondents had not yet been exposed to that literature at the time of PRE administration. Notably absent were informational and communicative barriers, social and attitudinal barriers, and structural and governance barriers—omissions that define the specific conceptual work the programme needed to accomplish.

This pattern is sociologically significant because it suggests an initial lay understanding of accessibility focused primarily on immediate, visible, or embodied obstacles rather than on relational, communicative, and institutional conditions of participation. In other words, accessibility was initially framed less as a socially organised achievement and more as a matter of managing sensory discomfort within the environment. The programme's success in broadening this view (see Theme 4) therefore indicates not only increased awareness of neurodivergent needs, but also a more differentiated understanding of accessibility as a multidimensional and institutionally mediated condition. This pattern is also compatible with a Goffmanian perspective, insofar as accessibility was initially perceived primarily at the level of observable difficulty rather than as an interactional condition of participation.

4.5.3. PRE Theme 3: AI as Generic Cultural Threat

PRE open-ended responses revealed a cluster of sceptical orientations that were broad, undifferentiated, and only weakly connected to specific AI applications. These responses framed AI primarily as a possible replacement for human workers, a source of data extraction, a vehicle for misinformation, or an undesirable intrusion into cultural experience. Although such concerns are not without empirical grounding, they were expressed without distinction among different types of AI, without reference to governance or design

safeguards, and without consideration of AI's potential role in supporting accessibility for neurodivergent visitors.

4.5.4. POST Theme 4: From Physical Access to Holistic Inclusion

The most consistent POST pattern is explicit conceptual reframing of inclusive design from a physical-access model toward a multi-dimensional, process-oriented, ethically grounded understanding. Multiple respondents articulated this shift contrastively: 'In the PRE questionnaire, I mainly saw [inclusive design] as ensuring physical accessibility, such as providing ramps or basic facilities for people with disabilities. However, through the programme, I now understand inclusive design as a broader and more complex concept. It includes not only physical access but also cultural, social, and emotional accessibility for diverse audiences' (POST respondent, BA student). A second respondent formulated the shift from checklist compliance to design ethics: 'The BIP deepened my understanding from seeing accessibility as a checklist to recognising it as a holistic, ethical design process. Inclusive design now means considering sensory, cognitive, and communicative needs from the start, rather than retrofitting accommodations'. This theme is consistent with Mezirow's first dimension of transformative learning, namely the critical revision of a prior meaning-making framework in light of a more differentiated alternative [16]. In heritage terms, this shift is especially significant because it repositions accessibility from a narrow issue of physical provision to a broader question of interpretation, participation, and visitor legitimacy.

4.5.5. POST Theme 5: On-Site Visits as Disorienting Dilemmas

When asked to identify the single most significant moment of the BIP week, the majority of POST respondents with substantive responses referenced the direct, embodied encounter with the inaccessibility of Pompeii and Herculaneum: 'Visiting Pompeii's crowded, narrow streets revealed how sensory and spatial barriers affect neurodivergent visitors. Experiencing first-hand the difficulty of navigating uneven terrain while reading complex signage made the need for multi-modal, flexible interpretation immediately clear'; 'Seeing lack of ramps and seeing people on wheelchair struggling'; 'Seeing the theory in practice: how inaccessible museums and sites be'. These responses are consistent with Mezirow's notion of the disorienting dilemma [16]: the experiential encounter with the gap between World Heritage status and observed inaccessibility appears to have functioned as a disconfirmatory experience that challenged prior assumptions about the public openness of heritage spaces. This finding suggests that BIP-format programmes with embedded site visits may offer a qualitatively different pedagogical configuration from lecture-only instruction, particularly in the field of heritage accessibility education.

4.5.6. POST Theme 6: Co-Design as Ethical Obligation with Acknowledged Structural Tension

POST respondents' engagement with the co-design-as-obligation argument showed a nuanced pattern: most endorsed the ethical framing while explicitly acknowledging structural obstacles: 'I agree that co-design with neurodivergent communities should be seen as an ethical obligation rather than just a methodological choice. After the programme, I better understand that inclusive design cannot be achieved without directly involving the people it aims to serve. However, I also see challenges: limited time and resources, difficulties in reaching and engaging diverse participants, and the risk of tokenism if participation is not meaningful'. This pattern is consistent with a further dimension of transformative learning in Mezirow's sense, namely a more reflexive awareness of the structural conditions that mediate the relationship between principled commitment and effective practice [16].

4.5.7. POST Theme 7: Forward-Looking Research Questions and Professional Identity Development

The final POST item invited respondents to formulate their most important question about inclusive AI in cultural heritage. Among the more developed responses: ‘How can AI in cultural heritage be designed in a way that truly represents diverse voices without reinforcing existing biases or exclusions? Answering this responsibly would require careful attention to data sources, transparency in how AI systems are built, and active involvement of diverse communities in the process’; ‘How can AI systems meaningfully personalise visitor experiences without inadvertently reinforcing stereotypes or exclusion? Answering this responsibly requires integrating UDL principles, participatory co-design, and continuous evaluation of bias’. These responses indicate that participants were beginning to appropriate programme-specific analytical vocabulary in the formulation of their own research questions, suggesting movement toward a more reflexive and ethically attentive professional identity in Mezirow’s sense [16]. In heritage terms, this is significant because participants were no longer responding only as learners but also as emerging interpreters of accessibility, participation, and institutional responsibility within cultural heritage settings.

5. Discussion

5.1. Integration of Quantitative and Qualitative Findings

The quantitative and qualitative strands broadly converge in suggesting a selective but meaningful pattern of learning change. The clearest quantitative patterns, albeit exploratory, were observed for L4 ($d = 1.21$), L7 ($d = 0.82$), L3 ($d = 0.75$), and L9 ($d = 0.85$) and correspond precisely to the thematic domains where qualitative data also documents substantive change: the acquisition of a dual UDL–ICF analytical vocabulary, the reframing of AI from generic threat to specifically theorised design instrument, and the consolidation of digital accessibility as a parallel priority to physical accessibility. The two Likert domains where quantitative change was minimal—L1 and L8—correspond to items where the PRE baseline was already high, reflecting pre-existing value commitments. In these domains, the programme may have contributed less to attitude formation than to greater analytical differentiation: participants appear to have moved from a more generic ethical commitment toward a more structured understanding of what inclusion may require institutionally and technologically. In heritage terms, this pattern suggests that participants increasingly understood accessibility not as a peripheral service issue but as part of visitor experience, interpretive mediation, and institutional responsibility.

5.2. Theoretical Contributions

These findings speak not only to inclusive AI design, but also to wider debates within heritage scholarship. They suggest that accessibility should be understood as a constitutive dimension of visitor experience rather than as an external accommodation and that neurodivergent inclusion challenges inherited assumptions about who the “normal” heritage visitor is. In this sense, the study aligns with critical heritage approaches that interpret heritage institutions as selective, norm-producing environments rather than neutral public spaces.

The study tentatively contributes to four ongoing discussions within the emerging literature on inclusive AI in cultural heritage.

First, the study offers preliminary indications that the dual ICF–UDL framework can be meaningfully introduced in a BIP context: significant large-effect gains in self-efficacy for applying both frameworks (L4) were observed within a single intensive lecture in a cohort characterised by near-zero prior familiarity with either framework. These findings

are promising but contextually bounded and should not be interpreted as confirmatory evidence pending replication.

Second, the study offers a possible sociological interpretation of how observed changes in participants' interpretive dispositions may have occurred beyond the mere presence of a disorienting dilemma. The structured observational checklist functioned as an analytical lens through which unmediated sensory overload—crowding, noise, uneven terrain—could be translated into categories aligned with the ICF–UDL framework. Without such a lens, a crowded Pompeian street might be experienced only as unpleasant or stressful, but not as a specific environmental barrier to participation for a defined functional profile.

Drawing on Bourdieu's concept of habitus [25], the checklist may have helped introduce a reflective break with participants' taken-for-granted, neurotypically embodied way of moving through heritage spaces. It may have encouraged them to perceive the site as designed rather than neutral, making more visible the contingent and socially organised character of accessibility barriers.

The checklist may also be interpreted as a device of professional vision in Goodwin's sense [26], shaping what counts as a relevant observation, how it could be coded (e.g., sensory, informational, attitudinal), and which kinds of responses—AI-based, architectural, or procedural—it could imply. Participants appear to have moved from a largely pre-reflective sense of environmental discomfort toward a more theoretically informed and actionable description of accessibility barriers. For example, the absence of predictive crowd-management tools could be understood as a source of unpredictable sound peaks constituting a barrier for autistic visitors with auditory hypersensitivity. This shift may help explain how an experience interpreted through Mezirow's notion of the disorienting dilemma could contribute to more reflexive ways of perceiving exclusion: the checklist may have provided cognitive and symbolic tools to name, classify, and contest previously less visible barriers.

The checklist appears to have functioned not only as a data-collection instrument but also as a pedagogical device supporting reflexive analysis of accessibility within heritage environments. This interpretation remains provisional, since the study did not include a comparison condition without the checklist. Future heritage education programmes may nevertheless benefit from exploring such observational tools as mediators between embodied experience and structural analysis. Third, the qualitative material is preliminarily consistent with the relevance of Mezirow's notion of the disorienting dilemma in heritage accessibility education (Theme 5) and suggests that BIP-format programmes with embedded site visits may offer pedagogical features not easily captured in lecture-only instruction. However, as no longitudinal follow-up was conducted, these findings reflect immediate situated change rather than confirmed transformative learning.

Fourth, the null result on L6 ($p = 0.963$) is theoretically suggestive: it may indicate that the personalisation–paternalism distinction requires sustained applied engagement rather than lecture exposure alone to generate attitudinal change. Indirect support for the curb-cut principle in AI heritage design [8,20] can be seen in POST respondents' tendency to frame inclusive AI as a universal quality enhancement rather than as a disability-specific accommodation.

Finally, the study tentatively contributes to the sociology of disability and technology by raising the hypothesis that AI, when framed through UDL–ICF principles, might in the future function as a de-stigmatising infrastructure—a possibility this pilot cannot demonstrate empirically but which warrants further investigation. Unlike assistive technologies that visibly mark users as different, more universally designed AI tools may embed flexibility within the baseline heritage experience, potentially reducing interactional friction and supporting neurodivergent modes of engagement. From this perspective, AI can be

approached not only as a possible surveillance risk, but also as a possible resource for greater autonomy and participation among neurodivergent visitors.

5.3. Practical Implications for Heritage Institutions and Policy

This orientation is also consistent with wider debates in heritage science concerning the evaluation of outcomes and social benefit. ICCROM has highlighted the need to strengthen the evidential basis through which heritage and conservation research demonstrates its impact, particularly in relation to public engagement, social inclusion, and decision-making processes. At the same time, these debates acknowledge the methodological challenges of capturing such impacts through standardised metrics. Within this perspective, the present study treats educational self-efficacy, interpretive reframing, and accessibility-oriented reflection as exploratory outcome domains that may contribute to a broader understanding of inclusive heritage practice.

The persistent low performance on Q3 (UDL v3.0 equity commitments) across both administrations has a practical implication: the equity-oriented sub-guidelines introduced in version 3.0—challenging exclusionary practices (Guideline 6.5), addressing biases in language and symbols (Guideline 2.4), cultivating belonging and community (Guideline 8.4)—are not yet widely known among heritage professionals or students in adjacent disciplines. These guidelines may therefore merit more explicit attention in professional development initiatives concerned with heritage AI deployment. The POST finding that sensory and environmental barriers were identified as most urgent (47% of respondents), combined with qualitative documentation of direct observation of inaccessibility at Pompeii and Herculaneum, lends preliminary support to the relevance of exploring AI-based non-invasive interventions in Italy's conservation-constrained heritage landscape, including personalised sensory mapping, predictive crowd management, NLP-driven content simplification, and pre-visit social story generation. The EU AI Act [18] introduces compliance obligations that may be relevant to cultural heritage institutions; the observed gain on L7 suggests that targeted BIP instruction on governance frameworks may contribute to participants' perceived preparedness to engage with these regulatory requirements.

For heritage institutions, these findings tentatively suggest three areas that may warrant priority attention: staff training in neurodiversity-informed interpretation, meaningful co-design with neurodivergent communities, and the cautious exploration of AI tools for non-invasive accessibility enhancement in conservation-constrained settings.

6. Limitations and Directions for Future Research

The study has five limitations. First, the sample size ($n = 14$ matched pairs) severely limits statistical power; all quantitative findings should be treated as exploratory, and effect sizes are more informative than p -values given this constraint. Given the small matched sample, the present study was not designed to detect small effects reliably and should be interpreted as exploratory. The quantitative component was sufficient to identify only relatively large directional shifts, while smaller or more nuanced changes may have remained undetected. Future research should therefore employ a priori sample-size planning based on minimally meaningful effect sizes, together with larger and more diverse samples and, where possible, multi-site designs. Such approaches would enable a clearer distinction between directional patterns observed in pilot contexts and more robust and generalisable effects. Second, the absence of a control group prevents causal attribution of attitudinal changes to programme participation. Third, social desirability bias may inflate positive Likert responses; anonymous administration partially mitigates this. Fourth, the dual role of lecturer and sole analyst may have had some influence on certain aspects of the interpretation, although steps such as peer review of the instrument and the use of an analytical audit

trail were taken to help limit potential bias. Fifth, neurodivergent individuals were not involved as co-researchers in instrument design, data collection, or analysis. This remains a limitation, particularly in light of the study's emphasis on co-design. Future phases of this research should therefore move beyond consultation toward the direct involvement of neurodivergent participants as research and design partners.

This means not only including neurodivergent individuals as participants or consultants but also ensuring they hold meaningful authority over research questions, instrument design, data interpretation, and dissemination—consistent with the Nothing About Us Without Us principle that the programme itself teaches.

Future research should address these limitations through (1) replication across multiple BIP cohorts with longitudinal follow-up at six months; (2) a matched comparison group to separate programme effects from maturation and other confounding variables; (3) systematic involvement of neurodivergent co-researchers as full research partners; (4) empirical evaluation of specific heritage AI tools against the UDL and six-component governance framework criteria; and (5) action research at Pompeii, Herculaneum, or the MANN to co-design and pilot a UDL-informed AI accessibility intervention with neurodivergent community members as equal design partners.

Taken together, the findings do not provide confirmatory evidence of programme effectiveness, but they do suggest that the UDL–ICF framework offers a promising conceptual and pedagogical basis for rethinking inclusive AI in cultural heritage. In this sense, the present study should be read primarily as a pilot contribution to framework transfer, methodological refinement, and future co-designed research rather than as a basis for strong causal or policy claims.

7. Conclusions

This pilot study offers three preliminary contributions to the emerging discussion on inclusive AI in cultural heritage, together with one critical sociological qualification.

First, it proposes a dual ICF–UDL analytical and design framework adapted from formal education to cultural heritage and offers preliminary indications that this framework can be meaningfully introduced in a BIP context, as suggested by gains in participants' self-efficacy in applying both models. These findings remain preliminary, and replication is required before stronger claims can be made.

Second, it suggests that direct experiential engagement with the inaccessibility of UNESCO-protected Italian heritage sites—Pompeii and Herculaneum—may have contributed to forms of conceptual reframing that participants themselves described in ways initially consistent with Mezirow's framework for situated perspective shifts. More broadly, the findings point to the possible pedagogical value of site-based experiential components in heritage accessibility education.

Third, it proposes a six-component governance framework encompassing universal design principles, community co-design authority, technical accountability and algorithmic audit, data ethics and informed consent, longitudinal impact evaluation, and professional development, grounded in the EU AI Act [18], the UN CRPD [19], the GDPR, and the empirical co-design literature.

Fourth, and most critically, the study raises a question left open by the first three contributions: is inclusive AI design sufficient to transform cultural institutions as sites of social reproduction, or does it leave the structure of the field largely intact?

From a Bourdieusian perspective [25], cultural heritage institutions can be understood as structured fields in which symbolic authority, professional roles, and institutional norms shape who is recognised as a legitimate visitor and what forms of engagement are treated as appropriate. Within this field, AI tools designed through UDL–ICF principles may reduce

barriers, but they do not automatically alter underlying power asymmetries, including who defines heritage value, whose needs are prioritised in design trade-offs, and which forms of expertise are institutionally rewarded.

Inclusive AI design is therefore necessary but not sufficient. Without parallel intervention in institutional structures—for example, through neurodivergent representation in governance, broader recognition of accessibility expertise, and a shift from reactive accommodation to proactive universal design—AI may risk becoming a technological layer placed over otherwise largely unchanged hierarchies. In this sense, AI can support transformation but cannot substitute for it. Without such field-level reflexivity, AI interventions may easily be absorbed into existing practices and used to adapt visitors to institutions rather than institutions to visitors.

The study's answer to this pilot question is therefore cautious: inclusive AI design does not automatically transform cultural institutions as sites of social reproduction. It may create conditions of possibility—by making exclusion more visible and by offering actionable alternatives—but such possibilities are more likely to translate into lasting change when accompanied by structural reforms in governance, accountability, and the institutional recognition of accessibility expertise. In this sense, AI may function as an enabling instrument, but not as a substitute for institutional transformation.

This study cannot demonstrate that AI fulfils this role, but it suggests the conditions under which such a role might be possible: AI may function as a more structural rather than merely supplementary instrument of heritage inclusion under at least three conditions: it is designed through UDL and ICF principles from the outset; it is developed in genuine co-production with neurodivergent communities holding meaningful decision-making authority; and it is embedded within a governance framework that ensures algorithmic accountability, ethical data stewardship, and longitudinal impact evaluation. A fourth condition should also be considered: the intervention should be accompanied by critical attention to the field's power asymmetries and exclusionary assumptions. Where these conditions are present, AI-enhanced heritage accessibility may operate not only as a compensatory accommodation for a minority but also as a broader improvement in heritage quality and participation, consistent with the curb-cut principle [27]. Where they are absent, AI may reproduce existing institutional limitations in a more technologically sophisticated form.

Future research should extend the present pilot by recruiting interdisciplinary cohorts—including museum curators, exhibition designers, conservation specialists, and AI engineers—to assess how the UDL–ICF framework is understood and operationalised across different professional roles and institutional constraints.

Author Contributions: Conceptualization, M.C.C. and A.S.; Methodology, M.C.C. and A.S.; Writing—original draft, M.C.C.; Writing—review and editing, M.C.C. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki (revised 2013). Formal IRB approval was not required for this non-clinical, anonymous educational survey of competent adult volunteers under Italian national legislation (D.Lgs. 196/2003 as amended by D.Lgs. 101/2018); the study design was reviewed and approved by the Erasmus+ BIP academic coordinating committee.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Participation was voluntary, anonymous, and without evaluative consequences; completion of the questionnaire constituted informed consent to the use of responses in aggregated, anonymised research outputs. No participant withdrew.

Data Availability Statement: Original data are available on request from the corresponding authors due to participant anonymity considerations.

Acknowledgments: The authors gratefully acknowledge the participants in the Erasmus+ BIP “Ethics in Social Science and Humanities” (Naples, Italy, 2026) and the partner institutions—University of Wrocław and IPB Bragança—for their collaboration, and the staff of the Archaeological Parks of Pompeii and Herculaneum and the MANN for facilitating the research visits.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificial Intelligence
ASC	Autism Spectrum Condition
AuND	Autistic and Neurodivergent
BIP	Blended Intensive Programme
CAST	Center for Applied Special Technology
CRPD	Convention on the Rights of Persons with Disabilities [19]
EU AI Act	Regulation (EU) 2024/1689 on Artificial Intelligence [18]
GDPR	General Data Protection Regulation (EU) 2016/679
ICF	International Classification of Functioning, Disability and Health (WHO, 2001)
IRB	Institutional Review Board
ISTAT	Istituto Nazionale di Statistica
MANN	Museo Archeologico Nazionale di Napoli
NLP	Natural Language Processing
UDL	Universal Design for Learning (CAST, v3.0, 2024)
WHO	World Health Organization

References

1. Singer, J. *NeuroDiversity: The Birth of an Idea*; Judy Singer: Melbourne, Australia, 2017.
2. Granland, M.E.; Sadia, T.; Cooper, E. Beyond the minimum: Designing for neurodiversity in museums and galleries. *Archnet-IJAR* **2026**, *20*, 522–544. [CrossRef]
3. Bourdieu, P.; Passeron, J.-C. *Reproduction in Education, Society and Culture*; Sage: London, UK, 1977.
4. Fletcher, A.L. Neurodivergence in Museums: Making the Cultural Heritage Sector Accessible Using Participatory Practices. Ph.D. Thesis, University of Glasgow, Glasgow, UK, 2024. Available online: <https://theses.gla.ac.uk/85238/> (accessed on 1 March 2026).
5. World Health Organization. *International Classification of Functioning, Disability and Health (ICF)*; WHO: Geneva, Switzerland, 2001. Available online: <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health> (accessed on 1 March 2026).
6. Hutson, J.; Hutson, P. *Inclusive Smart Museums: Engaging Neurodiverse Audiences and Enhancing Cultural Heritage*; Palgrave Macmillan: Cham, Switzerland, 2024. [CrossRef]
7. Siri, A. Emerging Trends and Future Directions in Artificial Intelligence for Museums: A Comprehensive Bibliometric Analysis Based on Scopus (1983–2024). *Geopolit. Soc. Secur. Freedom J.* **2024**, *7*, 20–38. [CrossRef]
8. Aruanno, B.; Carruba, M.C.; Mondellini, M.; Santos-Paz, J.A.; Ferrise, F.; Karaki, J.; Covarrubias, M. Enhancing inclusive education for young students with special needs through mixed reality. *Comput.-Aided Des. Appl.* **2024**, *21*, 522–535. [CrossRef]
9. Karaçay, Z.; Doğan, Y.; Doğan, N. Artificial intelligence application for museum to experiential transformation of cultural heritage and learning. *Smart Learn. Environ.* **2025**, *12*, 21. [CrossRef]
10. CAST. *Universal Design for Learning Guidelines Version 3.0*; CAST: Lynnfield, MA, USA, 2024. Available online: <https://udlguidelines.cast.org> (accessed on 1 March 2026).
11. CAST. Artificial Intelligence and UDL. 2024. Available online: <https://www.cast.org/what-we-do/artificial-intelligence/> (accessed on 1 March 2026).

12. Carruba, M.C.; Barca, A. App come ambienti di apprendimento inclusivi: Buone prassi per un approccio pedagogico. *Personae* **2023**, *2*, 24–48.
13. Molek, J.; Vela, R.C.; Olcese, G.; Siri, A. Accessibility as a Shared Cultural Responsibility: The *Entre Lucas* Project at the Pablo Gargallo Museum. *Heritage* **2025**, *8*, 475. [[CrossRef](#)]
14. Pellicano, E.; den Houting, J. Annual research review: Shifting from ‘normal science’ to neurodiversity in autism science. *J. Child Psychol. Psychiatry* **2022**, *63*, 381–396. [[CrossRef](#)] [[PubMed](#)]
15. Goffman, E. *Stigma: Notes on the Management of Spoiled Identity*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1963.
16. Mezirow, J. *Transformative Dimensions of Adult Learning*; Jossey-Bass: San Francisco, CA, USA, 1991.
17. European Parliament; Council of the European Union. Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the Accessibility Requirements for Products and Services (European Accessibility Act). *Off. J. Eur. Union* **2019**, *L 151*, 70–115.
18. European Parliament; Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). *Off. J. Eur. Union* **2024**, *L 1689*.
19. United Nations. *Convention on the Rights of Persons with Disabilities (CRPD)*; United Nations: New York, NY, USA, 2006. Available online: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html> (accessed on 20 April 2026).
20. UNESCO World Heritage Centre. World Heritage List. n.d. Available online: <https://whc.unesco.org/en/list/> (accessed on 1 March 2026).
21. BSI. *PAS 6463: Design for the Mind—Neurodiversity and the Built Environment*; British Standards Institution: London, UK, 2022.
22. Creswell, J.W.; Plano Clark, V.L. *Designing and Conducting Mixed Methods Research*, 3rd ed.; SAGE: Thousand Oaks, CA, USA, 2018.
23. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [[CrossRef](#)]
24. Braun, V.; Clarke, V. *Thematic Analysis: A Practical Guide*; SAGE: London, UK, 2022.
25. Bourdieu, P. *The Logic of Practice*; Stanford University Press: Stanford, CA, USA, 1990.
26. Goodwin, C. Professional vision. *Am. Anthropol.* **1994**, *96*, 606–633. [[CrossRef](#)]
27. Blackwell, A.G. The Curb-Cut Effect. *Stanf. Soc. Innov. Rev.* **2016**, *15*, 28–33. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.