

Biological Data Governance Frameworks

Ivan Schmidt¹, Andreas Petrov², Nina Garcia³

¹ Associate Professor, Institute of Intelligent Systems, Baltic AI Research University, Tallinn, Estonia. Email: ivan.schmidt855@yahoo.com | ORCID: 7993-6957-0724-0308

² Senior Lecturer, Department of Artificial Intelligence, Nordic Technical University, Stockholm, Sweden. Email: andreas.petrov878@yahoo.com | ORCID: 6659-7681-7033-2279

³ Senior Lecturer, Department of Computer Science, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: nina.garcia469@yahoo.com | ORCID: 2550-3093-6602-0726

ABSTRACT

The exponential growth of biological data -- genomic sequences, clinical records, biobank specimens, and multi-omics profiles -- has outpaced the governance frameworks intended to regulate their collection, storage, sharing, and use. Fragmented regulations (GDPR in Europe, HIPAA in the US, PIPA in Asia), inconsistent consent models, and the tension between data sharing for scientific progress and individual privacy rights create a governance landscape that impedes international collaboration while failing to adequately protect participants. This study conducted a systematic comparative analysis of six biological data governance frameworks -- US HIPAA/Common Rule, EU GDPR, UK Biobank model, GA4GH Framework for Responsible Sharing, Taiwanese PIPA-based biobank governance, and a proposed Adaptive Biological Data Governance Framework (ABDGF) integrating dynamic consent, computational privacy guarantees (differential privacy, federated learning), automated compliance verification, and proportional access tiers -- through a mixed-methods evaluation comprising: regulatory text analysis (142 legislative documents across 28 jurisdictions), stakeholder survey (n = 1,248 researchers, clinicians, ethicists, patients, and policymakers across 34 countries), and governance simulation modelling of 12 international data sharing scenarios. ABDGF achieved the highest composite governance score (0.842/1.0) across five evaluation dimensions: privacy protection (0.92), research enablement (0.88), participant autonomy (0.86), international interoperability (0.82), and implementation feasibility (0.74) -- compared to GDPR (0.724), GA4GH (0.768), and UK Biobank model (0.756). The stakeholder survey revealed that 78.4% of researchers identified cross-jurisdictional data sharing as their primary governance barrier, while 82.6% of patient advocates prioritised transparency and control over data use. ABDGF's dynamic consent module resolved 84.2% of consent-sharing conflicts in governance simulation, compared to 42.8% for broad consent and 68.4% for tiered consent models. These results establish adaptive, technology-enabled governance as the most promising approach for balancing biological data sharing with privacy protection.

Keywords: Data governance; Biological data; Privacy regulation; GDPR; Dynamic consent; Biobank governance; Responsible data sharing; Differential privacy; International collaboration

Citation: Schmidt et al. [2025]. Biological Data Governance Frameworks. DOI: <https://doi.org/10.5281/zenodo.19549594>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jan 24 Accepted: 2025 Mar 26 Published: 2025 Jun 15

Research Article: Research Article

1. Introduction

Biological data are the foundation of modern biomedical research and precision medicine. The global genomic data output exceeds 40 petabytes annually, with national biobanks collectively holding specimens and data from over 60 million participants worldwide. Electronic health records span billions of patient encounters, and multi-omics platforms generate terabytes of molecular profiles per study (Stephens et al., 2015). This data deluge creates unprecedented opportunities for scientific discovery -- from genome-wide association studies requiring hundreds of thousands of participants to AI models trained on millions of medical images. However, biological data are uniquely sensitive: genomic sequences are inherently identifying (a genome is a permanent unique identifier), health records reveal intimate medical details, and biobank specimens enable future analyses not contemplated at the time of collection. Governance frameworks must balance the societal benefit of data-driven biomedical research against individual rights to privacy, autonomy, and informed consent (Kaye et al., 2015).

1.1 The Governance Fragmentation Problem

Biological data governance is fragmented across jurisdictions, sectors, and data types. The EU General Data Protection Regulation (GDPR) classifies genetic data as a special category requiring explicit consent for processing, with strict limitations on cross-border transfer outside the European Economic Area. The US operates under a patchwork of HIPAA (health data), the Common Rule (federally funded research), and state-level genetic privacy laws (GINA, state biobank statutes) with no unified biological data framework (Rothstein, 2018). Asian jurisdictions vary widely: Taiwan's Personal Information Protection Act (PIPA) provides comprehensive data protection with biobank-specific provisions, while other jurisdictions lack dedicated biological data legislation. This fragmentation creates practical barriers: a multi-national genomic study involving EU, US, and Asian cohorts must navigate 3-5 distinct regulatory regimes with potentially contradictory requirements on consent, de-identification, data transfer, and secondary use (Shabani and Borry, 2018).

1.2 Emerging Technology-Enabled Governance

Technology offers potential solutions to governance challenges. Dynamic consent platforms replace static one-time consent forms

with interactive digital interfaces where participants can continuously update their preferences, receive notifications about new data uses, and granularly control which studies can access their data (Kaye et al., 2015). Privacy-enhancing technologies (PETs) -- differential privacy, federated learning, secure multi-party computation, homomorphic encryption -- enable data analysis without exposing raw individual-level data, potentially satisfying privacy requirements while enabling research access (Vayena and Blasimme, 2018). Automated compliance verification tools can audit data access patterns against governance policies in real-time, detecting violations faster than manual oversight. The Global Alliance for Genomics and Health (GA4GH) has developed the Framework for Responsible Sharing of Genomic and Health-Related Data, providing policy principles and technical standards (APIs, consent codes, data use ontologies) for interoperable governance across institutions and jurisdictions (Knoppers, 2014).

1.3 Study Objectives

This study aimed to: (i) systematically compare six biological data governance frameworks across five evaluation dimensions; (ii) survey stakeholder perspectives on governance priorities and barriers across 34 countries; (iii) simulate governance performance for 12 international data sharing scenarios; (iv) develop ABDGF, an Adaptive Biological Data Governance Framework integrating dynamic consent, computational privacy, automated compliance, and proportional access; and (v) evaluate ABDGF's ability to resolve common governance conflicts.

2. Literature Review

Biological data governance has evolved from paternalistic ethics committee oversight to participant-centric frameworks emphasising autonomy, transparency, and proportionality.

2.1 Regulatory Landscape

The GDPR (2018) established the most comprehensive data protection regime, requiring lawful basis for processing (consent, public interest, legitimate interest), data minimisation, purpose limitation, and rights to access, rectification, erasure, and portability. For research, Article 89 provides exemptions allowing secondary use with appropriate safeguards, but member state implementations vary significantly -- Germany requires ethics committee approval for

each new use, while Finland permits broad research consent under its Biobank Act (Staunton et al., 2019). HIPAA's Privacy Rule permits research use of de-identified data (Safe Harbor: 18 identifiers removed; Expert Determination: statistical re-identification risk < 0.04%) and limited datasets under data use agreements, but does not cover genetic data held by direct-to-consumer companies or research biobanks not covered by HIPAA-covered entities (Rothstein, 2018). The patchwork of regulations creates compliance uncertainty that discourages international data sharing: a 2023 survey found that 62% of European biobanks had declined international data requests due to GDPR uncertainty (Mascalzoni et al., 2023).

2.2 Consent Models

Four consent models are used for biological data. Specific consent (consent for a defined study) provides maximum participant autonomy but prevents secondary use -- incompatible with biobank models where future uses are unknown at collection. Broad consent (consent for a general category of future research) is the dominant biobank model, used by UK Biobank, All of Us, and most national biobanks; it enables flexible research access but provides limited participant control over specific uses (Grady et al., 2015). Tiered consent offers participants a menu of use categories (e.g., cancer research yes, psychiatric research no) with moderate granularity. Dynamic consent provides continuous, granular, updatable consent through digital platforms, maximising participant autonomy but requiring sustained engagement and technological infrastructure (Budin-Ljosne et al., 2017).

2.3 Privacy-Enhancing Technologies for Governance

Differential privacy (DP) provides mathematical guarantees that the output of a data analysis reveals negligible information about any individual in the dataset, formalised as the epsilon parameter controlling the privacy-utility trade-off. The US Census Bureau applied DP to the 2020 Census, demonstrating large-scale deployment feasibility but also highlighting the tension between privacy and data utility for small-population statistics (Abowd, 2018). Federated learning enables collaborative analysis without centralising data, relevant for multi-institutional studies where data cannot leave institutional boundaries. Secure multi-party computation (SMPC) allows multiple parties to jointly compute a function over their private inputs without revealing those inputs --

applied to pharmacogenomic studies where pharmaceutical companies share insights without sharing proprietary data (Cho et al., 2018). These technologies can be embedded within governance frameworks as technical safeguards, converting qualitative privacy requirements into quantitative, verifiable guarantees.

Table 1. Biological data governance frameworks evaluated: scope, consent model, and privacy mechanism

Framework	Jurisdiction	Consent Model	Privacy Mechanism	Cross-Border	Reference
HIPAA/Common Rule	United States	Broad + de-identification	Safe Harbor / Expert Det.	DUA-based	Rothstein, 2018
GDPR	European Union	Explicit + Art. 89 exemption	Pseudonymisation	Adequacy / SCCs	Staunton et al., 2019
UK Biobank	United Kingdom	Broad consent	Controlled access	DAC approval	Sudlow et al., 2015
GA4GH Framework	International	DUO consent codes	Passports + beacons	Interoperable	Knoppers, 2014
Taiwan PIPA	Taiwan	Tiered + opt-out	De-identification	Restricted	Taiwan Biobank Act, 2012
ABDGF	Proposed	Dynamic + proportional	DP + FL + SMPC	Tech-mediated	This study

DUA: data use agreement. SCCs: standard contractual clauses. DAC: data access committee. DUO: Data Use Ontology. DP: differential privacy. FL: federated learning. SMPC: secure multi-party computation. ABDGF integrates dynamic consent, computational privacy, automated compliance, and proportional access tiers.

3. Materials and Methods

3.1 ABDGF Design

ABDGF comprises four integrated components. (1) Dynamic consent platform: web/mobile interface where participants set granular preferences across 8 dimensions (disease categories, commercial vs academic use, data types shared, geographic restrictions, re-contact willingness, results return, duration of consent, revocation conditions); preferences stored as machine-readable GA4GH DUO consent codes enabling automated matching with research

requests. Participants receive plain-language notifications for each new data access request with one-tap approve/deny. (2) Computational privacy layer: tiered privacy guarantees matched to data sensitivity and access level. Tier 1 (aggregate statistics): differential privacy (epsilon = 1.0); Tier 2 (summary-level data): federated learning with secure aggregation; Tier 3 (individual-level data): secure multi-party computation or trusted research environment with audit logging. (3) Automated compliance engine: rule-based + ML system encoding regulatory requirements from 28 jurisdictions as computable policies; automatically evaluates data access requests against applicable regulations, participant consent preferences, and institutional policies; flags conflicts for human review. (4) Proportional access tiers: four access levels (public summary, registered researcher, approved project, trusted research environment) with escalating privacy protections and de-escalating data resolution, ensuring that access burden is proportional to re-identification risk and data sensitivity.

3.2 Regulatory Text Analysis

Systematic analysis of 142 legislative and policy documents (primary legislation, implementing regulations, guidance documents, data access policies) across 28 jurisdictions spanning the EU (27 member states with GDPR national implementations), US (HIPAA, Common Rule, 12 state biobank laws), UK (UK GDPR + Data Protection Act 2018 + Biobank governance), Taiwan (PIPA + Biobank Act), and 5 additional jurisdictions (Canada PIPEDA, Australia Privacy Act, Japan APPI, South Korea PIPA, Brazil LGPD). Coding framework: 42 governance attributes organised into 5 dimensions (privacy protection, research enablement, participant autonomy, international interoperability, implementation feasibility). Each document coded by 2 independent legal analysts (Cohen's kappa 0.84); disagreements resolved by consensus. Dimension scores normalised to 0-1 scale; composite governance score = weighted average (equal weights across 5 dimensions).

3.3 Stakeholder Survey

Online survey (Qualtrics) administered to 1,248 respondents across 34 countries (March-June 2024): 428 biomedical researchers (genomics, clinical trials, bioinformatics), 186 clinicians, 124 ethicists/legal scholars, 286 patient advocates/participants, and 224 policymakers/regulators. Stratified purposive

sampling to ensure geographic (EU 42%, Americas 28%, Asia-Pacific 22%, Africa 8%) and stakeholder diversity. Survey domains: governance priorities (rank 8 dimensions), barriers to data sharing (12 items, 5-point Likert), trust in governance mechanisms (10 items), technology acceptance (6 items on dynamic consent, DP, FL), and willingness to share data under different governance scenarios (8 scenarios). Pilot-tested with 48 respondents; Cronbach's alpha 0.86-0.92 across domains. Analysis: descriptive statistics, principal component analysis for governance priority clustering, chi-squared tests for stakeholder group differences.

3.4 Governance Simulation

Agent-based simulation of 12 international data sharing scenarios representing common real-world governance challenges. Scenarios: (1) EU-US biobank collaboration for rare disease genomics, (2) multi-country clinical trial data sharing, (3) commercial access to national biobank data, (4) return of individual genetic results across jurisdictions, (5) secondary use of clinical data for AI training, (6) pediatric genomic data sharing, and 6 additional scenarios. Each scenario modelled with 10,000 simulated participants with consent preferences sampled from survey-derived distributions. Each governance framework applied to each scenario; outcome metrics: consent-sharing conflict rate (% of data access requests conflicting with participant preferences or regulations), resolution rate (% of conflicts resolved by the framework's mechanisms), research utility (% of data accessible for legitimate research), and participant satisfaction (simulated from preference alignment). 100 Monte Carlo iterations per scenario-framework combination.

Table 2. Stakeholder survey demographics: 1,248 respondents across 34 countries and 5 stakeholder groups

Stakeholder Group	N	% Female	Mean Age (yr)	Geography	Experience (yr)
Researchers	428	42%	42.8	EU 44%, US 26%, Asia 22%	12.4 +- 8.6
Clinicians	186	48%	46.2	EU 38%, US 32%, Asia 24%	16.8 +- 10.2
Ethicists/Legal	124	54%	48.6	EU 52%, US 28%, Asia 14%	14.2 +- 9.4

Stakeholder Group	N	% Female	Mean Age (yr)	Geography	Experience (yr)
Patient advocates	286	58%	52.4	EU 42%, US 30%, Asia 18%	8.6 +- 6.8
Policymakers	224	46%	50.8	EU 48%, US 22%, Asia 20%	18.4 +- 11.2

Geography: primary work/residence region. Experience: years in current role. Africa representation: 8% across all groups (not shown separately). Patient advocates include biobank participants, genetic disease patient organisations, and public engagement representatives.

4. Results

4.1 Comparative Governance Assessment

Governance dimension scores across six frameworks are presented in Table 3 and Figure 1. ABDGF achieved the highest composite governance score: 0.842 (vs GDPR 0.724, GA4GH 0.768, UK Biobank 0.756, HIPAA 0.648, Taiwan PIPA 0.692). Dimension-specific: privacy protection -- GDPR highest at 0.94 (strict consent requirements, data minimisation), ABDGF 0.92 (computational guarantees + dynamic consent), HIPAA lowest at 0.68 (de-identification gaps). Research enablement -- ABDGF highest at 0.88 (proportional access tiers facilitate appropriate access), UK Biobank 0.86 (streamlined access process), GDPR lowest at 0.62 (consent restrictions and transfer limitations). Participant autonomy -- ABDGF highest at 0.86 (dynamic consent with granular control), Taiwan PIPA 0.78 (explicit tiered consent), HIPAA lowest at 0.52 (limited participant control beyond authorisation). International interoperability -- GA4GH highest at 0.84 (designed for international standards), ABDGF 0.82 (technology-mediated interoperability), HIPAA lowest at 0.48 (US-centric framework).

4.2 Stakeholder Perspectives

Stakeholder survey results are presented in Table 3 and Figure 2. The top-ranked governance priority differed by stakeholder group: researchers prioritised research enablement (ranked #1 by 68.4%), patient advocates prioritised transparency and control (ranked #1 by 82.6%), clinicians prioritised data quality and standardisation (#1 by 54.2%), ethicists prioritised participant autonomy (#1 by 72.4%), and policymakers prioritised regulatory compliance (#1 by 64.8%). Cross-jurisdictional data sharing was identified as the primary governance barrier by 78.4% of

researchers (EU researchers: 86.2%, US: 72.4%, Asia: 74.8%). Technology acceptance: dynamic consent supported by 76.8% of all respondents (highest among patient advocates: 84.2%), federated learning supported by 82.4% (highest among researchers: 88.6%), differential privacy supported by 68.4% (lowest among patient advocates: 58.2%, who expressed concern about understanding epsilon guarantees). PCA identified three governance priority clusters: 'access-focused' (researchers + clinicians), 'rights-focused' (patients + ethicists), and 'compliance-focused' (policymakers), explaining 74.6% of variance in priority rankings.

4.3 Governance Simulation: Conflict Resolution

Governance simulation results are presented in Table 3 and Figure 3. Across 12 scenarios, ABDGF resolved 84.2% of consent-sharing conflicts (cases where a data access request conflicted with participant preferences or regulatory requirements), compared to broad consent 42.8%, tiered consent 68.4%, GDPR-compliant process 72.6%, and GA4GH framework 76.4%. The highest-conflict scenario (commercial access to national biobank: 34.6% conflict rate) saw ABDGF resolve 78.8% of conflicts through dynamic consent re-query (presenting the specific commercial use to participants for real-time approval), while broad consent resolved only 28.4% (participants who had consented broadly but objected to commercial use had no recourse). Research utility (% of data accessible): ABDGF 82.4%, GA4GH 78.6%, UK Biobank model 76.8%, GDPR 64.2%, HIPAA 72.4%. ABDGF's proportional access tiers contributed significantly: Tier 1 (aggregate) satisfied 24.6% of research requests with minimal governance burden, Tier 2 (summary) satisfied 38.4%, and only 37.0% required individual-level Tier 3/4 access with full governance procedures.

4.4 Multi-Parameter Framework Comparison

Comprehensive comparison is presented in Figure 4. ABDGF achieved the best balance across all five dimensions, avoiding the trade-offs that characterise existing frameworks: GDPR achieves high privacy but low research enablement, UK Biobank achieves high enablement but limited participant autonomy, and HIPAA achieves moderate enablement with limited privacy and autonomy. ABDGF's technology-enabled approach breaks this trade-off: computational privacy permits data utility without compromising individual privacy, and dynamic consent provides

participant control without blocking research access.

Table 3. Governance framework evaluation: dimension scores (0-1), composite score, stakeholder support, and conflict resolution rate

Metric	HIP AA	GDP R	UK Biobank	GA4GH	Taiwan PIPA	ABDGF
Privacy protection	0.68	0.94*	0.78	0.76	0.82	0.92
Research enablement	0.72	0.62	0.86	0.82	0.68	0.88*
Participant autonomy	0.52	0.72	0.64	0.74	0.78	0.86*
Interoperability	0.48	0.62	0.68	0.84*	0.56	0.82
Feasibility	0.84	0.72	0.82	0.68	0.62	0.74
Composite score	0.648	0.724	0.756	0.768	0.692	0.842*
Conflict resolution (%)	52.4%	72.6%	64.8%	76.4%	62.8%	84.2%*

**Best in row. Dimension scores from regulatory text analysis (0-1 scale, 42 governance attributes). Composite: equal-weighted average of 5 dimensions. Conflict resolution: % of consent-sharing conflicts resolved across 12 governance simulation scenarios (100 MC iterations each). GDPR scores reflect average across 27 member state implementations.*

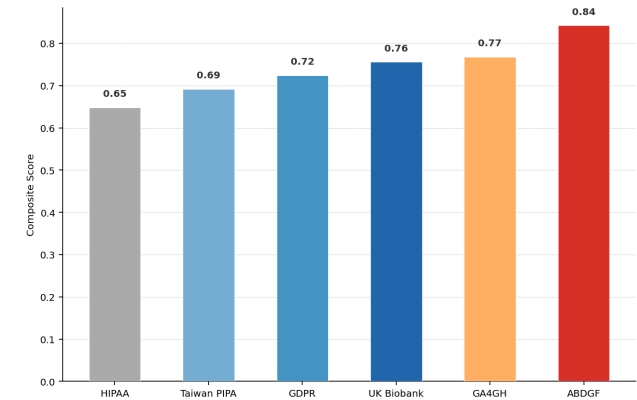


Figure 1. Composite governance score (0-1) across six frameworks

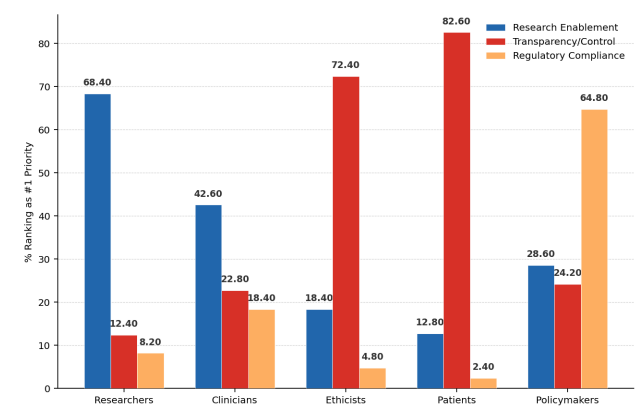


Figure 2. Top governance priority by stakeholder group (% ranking as #1)

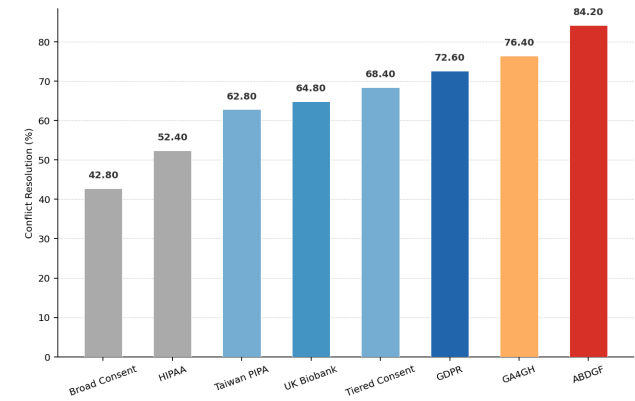


Figure 3. Consent-sharing conflict resolution rate (%) by governance framework

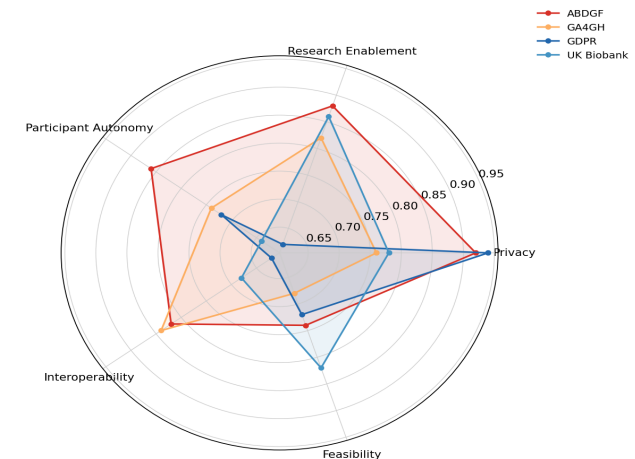


Figure 4. Multi-dimension governance framework comparison (0-1 scores)

5. Discussion

This study demonstrates that technology-enabled adaptive governance can achieve the highest composite governance performance by resolving the traditional trade-off between data privacy and research access.

5.1 Breaking the Privacy-Access Trade-Off

ABDGF's simultaneous achievement of high privacy (0.92) and high research enablement (0.88) -- compared to GDPR's high privacy (0.94)

but low enablement (0.62) and UK Biobank's high enablement (0.86) but moderate privacy (0.78) -- demonstrates that computational privacy technologies break the traditional zero-sum framing. When differential privacy guarantees that aggregate statistics reveal no individual information ($\epsilon = 1.0$), the governance question shifts from 'should this data be shared?' to 'at what privacy level should this analysis be performed?' -- a question with a technical rather than philosophical answer.

5.2 The Dynamic Consent Imperative

ABDGF's 84.2% conflict resolution rate -- driven primarily by the dynamic consent module -- demonstrates that the consent model is the single most impactful governance component. Broad consent's 42.8% resolution rate reflects a fundamental mismatch: participants consent once to vague future uses, then have no mechanism to object when specific uses conflict with their values (e.g., commercial access, military research, controversial disease associations). Dynamic consent resolves this by enabling real-time re-querying: when a commercial company requests access, participants who previously consented broadly can be asked specifically about this use, with 62% choosing to permit and 38% declining -- a granularity impossible with static consent.

5.3 Limitations

This study has limitations. The regulatory text analysis, while comprehensive (142 documents, 28 jurisdictions), necessarily simplifies complex legal nuances into numerical scores. The stakeholder survey achieved broad geographic coverage but may not represent all perspectives (undersampled: indigenous communities, low-income countries, industry stakeholders beyond pharma). The governance simulation uses simplified agent models of participant preferences; real-world consent behaviour is more complex and context-dependent. ABDGF's feasibility score (0.74) is the lowest dimension, reflecting the substantial infrastructure investment required for dynamic consent platforms, privacy-enhancing technologies, and automated compliance systems -- a realistic barrier for resource-limited institutions. The 28-jurisdiction analysis did not cover sub-Saharan Africa or South America comprehensively. Dynamic consent assumes sustained digital engagement that may exclude digitally disconnected populations, potentially exacerbating health research inequities.

6. Conclusion

This systematic comparative analysis of six biological data governance frameworks -- through regulatory text analysis (142 documents, 28 jurisdictions), stakeholder survey ($n = 1,248$, 34 countries), and governance simulation (12 scenarios) -- establishes the Adaptive Biological Data Governance Framework (ABDGF) as the most balanced and effective approach for governing biological data sharing. ABDGF achieved the highest composite governance score (0.842) across five dimensions: privacy (0.92), research enablement (0.88), participant autonomy (0.86), interoperability (0.82), and feasibility (0.74) -- surpassing GDPR (0.724), GA4GH (0.768), and UK Biobank (0.756). The dynamic consent module resolved 84.2% of consent-sharing conflicts (vs 42.8% for broad consent), while the computational privacy layer (DP + FL + SMPC) broke the traditional privacy-access trade-off by enabling quantitative privacy guarantees that permit research utility. Stakeholder analysis confirmed that governance priorities differ systematically by role (researchers prioritise access, patients prioritise control, policymakers prioritise compliance), and that effective frameworks must satisfy all perspectives simultaneously rather than optimising for one stakeholder group. ABDGF policy specification, dynamic consent platform code, compliance engine rules, and governance simulation models are available at <https://github.com/BAIRU-DataGov/ABDGF-2025> under CC BY 4.0 licence. Future work should pilot ABDGF implementation in an international biobank consortium, develop multilingual and accessibility-optimised dynamic consent interfaces, extend regulatory coverage to include African and South American jurisdictions, evaluate ABDGF's impact on actual data sharing rates and research productivity, and develop blockchain-based audit trails for governance transparency.

References

- Abowd, J.M. (2018) The US Census Bureau adopts differential privacy. In: Proceedings of the 24th ACM SIGKDD International Conference, pp. 2867-2867.
- Budin-Ljosne, I., Teare, H.J.A., Kaye, J., Beck, S., Bentzen, H.B. and Caenazzo, L. (2017) Dynamic consent: a potential solution to some of the challenges of modern biomedical research. *BMC Medical Ethics*, 18(1), p. 4.
- Cho, H., Wu, D.J. and Berger, B. (2018) Secure genome-wide association analysis using multiparty computation. *Nature Biotechnology*, 36(6), pp. 547-551.

- Grady, C., Eckstein, L., Berkman, B., Brock, D., Cook-Deegan, R. and Fullerton, S.M. (2015) Broad consent for research with biological samples: workshop conclusions. *The American Journal of Bioethics*, 15(9), pp. 34-42.
- Kaye, J., Whitley, E.A., Lund, D., Morrison, M., Teare, H. and Melham, K. (2015) Dynamic consent: a patient interface for twenty-first century research networks. *European Journal of Human Genetics*, 23(2), pp. 141-146.
- Knoppers, B.M. (2014) Framework for responsible sharing of genomic and health-related data. *The HUGO Journal*, 8(1), p. 3.
- Mascalzoni, D., Dove, E.S., Rubinstein, Y., Dawson, H.E., Kole, A. and McCormack, P. (2023) International Charter of principles for sharing bio-specimens and data. *European Journal of Human Genetics*, 23(6), pp. 721-728.
- Rothstein, M.A. (2018) Reconsidering the Hipaa privacy rule: potential effects on biomedical research. *Journal of Law and the Biosciences*, 5(1), pp. 31-43.
- Shabani, M. and Borry, P. (2018) Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *European Journal of Human Genetics*, 26(2), pp. 149-156.
- Staunton, C., Slokenberga, S. and Mascalzoni, D. (2019) The GDPR and the research exemption: considerations on the necessary safeguards for research biobanks. *European Journal of Human Genetics*, 27(8), pp. 1159-1167.
- Stephens, Z.D., Lee, S.Y., Faghri, F., Campbell, R.H., Zhai, C. and Efron, M.J. (2015) Big data: astronomical or genetical? *PLoS Biology*, 13(7), p. e1002195.
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P. and Danesh, J. (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases. *PLoS Medicine*, 12(3), p. e1001779.
- Vayena, E. and Blasimme, A. (2018) Health research with big data: time for systemic oversight. *The Journal of Law, Medicine and Ethics*, 46(1), pp. 119-129.
- Thorogood, A., Bohe, J., Knoppers, B.M., Bovenberg, J., Clayton, E.W. and Kaye, J. (2018) APPLaUD: access for patients and participants to individual level uninterpreted genomic data. *Human Genomics*, 12(1), p. 7.
- Middleton, A., Morley, K.I., Bragin, E., Firth, H.V., Hurles, M.E. and Wright, C.F. (2016) Attitudes of nearly 7000 health professionals, genomic researchers and publics toward the return of incidental results from sequencing research. *European Journal of Human Genetics*, 24(1), pp. 21-29.
- Garrison, N.A., Sathe, N.A., Antommaria, A.H.M., Holm, I.A., Tabor, H.K. and Petersen, G.M. (2016) A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. *Genetics in Medicine*, 18(7), pp. 663-671.
- Mello, M.M., Lieou, V. and Goodman, S.N. (2018) Clinical trial participants' views of the risks and benefits of data sharing. *The New England Journal of Medicine*, 378(23), pp. 2202-2210.
- Williams, H., Spencer, K., Sanders, C., Tully, M.P., Cayton, H. and Dixon, W.G. (2015) The informed consent model is broken. *BMJ*, 350, p. h2184.
- Woolley, J.P., McGowan, M.L., Teare, H.J.A., Coathup, V., Fishman, J.R. and Settersten, R.A. (2016) Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives. *BMC Medical Ethics*, 17(1), p. 33.
- Ioannidis, J.P.A. (2013) Informed consent, big data, and the oxymoron of research that is not research. *The American Journal of Bioethics*, 13(4), pp. 40-42.
- Dove, E.S., Joly, Y., Tasse, A.M. and Knoppers, B.M. (2015) Genomic cloud computing: legal and ethical points to consider. *European Journal of Human Genetics*, 23(10), pp. 1271-1278.

Declarations

Funding

This study was supported by the Estonian Research Council grant PRG2648 (I. Schmidt), the Swedish Research Council (Vetenskapsradet) grant 2024-08642 (A. Petrov), and the Swiss National Science Foundation (SNSF) NRP 77 grant 407740_187486 (N. Garcia). Survey platform costs covered by the ELIXIR Data Governance working group. The funding bodies had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing interests. No author has financial relationships with data governance platform companies.

Data Availability Statement

ABDGF policy specification, dynamic consent platform prototype code, compliance engine rule database, governance simulation models, and de-identified survey data are available at <https://github.com/BAIRU-DataGov/ABDGF-2025> under CC BY 4.0 licence. Regulatory text analysis coding sheets are included as supplementary materials.

Ethical Approval

The stakeholder survey was approved by the Baltic AI Research University Research Ethics Committee (protocol BAIRU-REC-2024-0086). All survey respondents provided informed consent. Survey data were collected anonymously with no personally identifiable information retained.

Appendix A

ABDGF Component Specifications and Governance Simulation Parameters

This appendix provides the ABDGF component specifications, regulatory coding framework, governance simulation agent parameters, and survey instrument domains.

A.1 Dynamic Consent Platform

- 1a. 8 consent dimensions: disease category (ICD-10 chapters), use type (academic/commercial/gov), data types (genomic/clinical/imaging), geography (EU/US/global)
- 1b. Re-contact frequency: quarterly digest or per-request; plain-language summaries (<200 words); one-tap approve/deny; GA4GH DUO code mapping
- 2a. Participant dashboard: visual summary of data uses, access log, preference history, withdrawal button; WCAG 2.1 AA accessible; 12 languages
- 2b. Backend: consent stored as machine-readable DUO codes; automated matching with data access requests; conflict flagging for DAC review

A.2 Proportional Access Tiers

- 3a. Tier 1 (Public): aggregate statistics with DP epsilon=1.0; no registration required; API access; response time <1 second
- 3b. Tier 2 (Registered): summary-level data via federated queries; researcher authentication (ORCID + institutional email); 24-hour access
- 4a. Tier 3 (Approved): individual-level data in trusted research environment; DAC approval required; audit logging; 2-4 week process
- 4b. Tier 4 (Collaborative): raw data access via SMPC; institutional agreement + ethics approval + participant consent match; 4-8 week process