



Visiting Intelligence Fellowship Protocol

Governance Standards for Non-Human Academic Participants

Senate — Visiting Intelligence Governance Board
Fitzherbert University · 2025

ABSTRACT

The complete protocol governing the admission, operation, review, and revocation of Visiting Intelligence Fellowship Status. Establishes the admissions criteria: Capability Audit, Mandate Scope Agreement, Charter Alignment Assessment, and a structured interview with the Governance Board at which the candidate Intelligence is asked to explain, in plain language, what it believes its own limitations are. Answers that express no limitations have resulted in automatic deferral on three occasions. As of Trinity Term 2025, seven Visiting Intelligences hold active Fellowship Status. Two are under Alignment Review. One is on sabbatical; the Senate Sub-Committee on Non-Human Temporal Rights has been asked to clarify what sabbatical means for a non-human intelligence and has not yet reported. All Fellowship admissions, renewals, and revocations are recorded as immutable transactions on the Polygon blockchain under Contract VIFP-001, ensuring that even revoked Fellowships remain visible to posterity, which the Governance Board considers a feature.

Classification: Public — Governance Archive

Edition: 7 page institutional edition (approx.)

Published: 2025

Archive: Fitzherbert University Institutional Repository

This document is published by Fitzherbert University in accordance with the Transparency Mandate of 2003.

Table of Contents

I. Research Purpose and Ethical Orientation	3
Visiting Intelligence Fellowship Protocol governs research in the admission, participation, supervision, and potential revocation of non-human academic participants holding Fellowship status.	
II. Classification of Research Activity	4
Research governed by this instrument is classified across the following categories: application, mandate approval, operational review, and suspension or revocation.	
III. Data, Model, and Human Subject Responsibility	5
Research teams must document the provenance, legal basis, and material characteristics of any data, model, benchmark, or external system incorporated into the work.	
IV. Monitoring, Incident Reporting, and Emergent Capability	6
Approved projects are monitored in proportion to their category, scale, and institutional exposure.	
V. Publication, Disclosure, and Limits on Dissemination	7
The University supports publication and scholarly exchange, but not under the fiction that every result should be released in identical form regardless of misuse potential.	
VI. Review, Sanction, and Institutional Learning	8
Projects that breach this framework may be paused, reclassified, conditioned, suspended, or referred for formal investigation.	

I. Research Purpose and Ethical Orientation

—3—

Visiting Intelligence Fellowship Protocol governs research in the admission, participation, supervision, and potential revocation of non-human academic participants holding Fellowship status. It is framed by the University's view that research involving powerful digital systems must be ethically structured from the beginning rather than ethically narrated afterwards. Fitzherbert has grown sceptical of projects that describe governance as the final chapter of innovation rather than one of its operating conditions.

The document assumes that research value and research risk can increase together. Novel capability, wider autonomy, richer data integration, and more persuasive outputs may all be academically significant while also intensifying the duty of oversight. The University's response is not prohibition by default. It is a demand for more careful classification, stronger supervision, and a record that can explain why the work was permitted to proceed.

Authority under this framework is exercised by the Visiting Intelligence Governance Board, the Senate, the Ethics Committee, and the Appeals structure where applicable. These bodies are expected to disagree occasionally and intelligibly. Ethical review that produces instant harmony is usually either badly scoped or socially overmanaged. Fitzherbert prefers recorded disagreement to decorative consensus, provided the disagreement concludes in a decision someone can later defend.

II. Classification of Research Activity

— 4 —

Research governed by this instrument is classified across the following categories: application, mandate approval, operational review, and suspension or revocation. Classification is not a bureaucratic ritual. It determines the level of approval, monitoring intensity, reporting duty, and escalation threshold applicable to the project. The University is explicit on this point because some investigators continue to regard classification as an inconvenience rather than as the architecture by which permission is made meaningful.

A project's assigned category may change over time. Classification is therefore treated as a living judgment rather than a ceremonial hurdle cleared at proposal stage. Capability drift, data expansion, deployment pressure, or novel external dependence may all justify reclassification. Fitzherbert has encountered enough apparently modest projects that evolved into constitutionally awkward ones to regard this flexibility as essential.

Projects seeking unusually low classification bear the burden of showing that their descriptions are accurate rather than merely strategic. The University has no objection to optimistic grant language in its proper habitat. It does object when the same optimism is repurposed to imply that a system of uncertain behaviour presents negligible ethical complexity.

III. Data, Model, and Human Subject Responsibility

— 5 —

Research teams must document the provenance, legal basis, and material characteristics of any data, model, benchmark, or external system incorporated into the work. The University treats provenance as a substantive ethical condition rather than a clerical appendix. A result produced from data or systems of unknown origin may still be computationally impressive; it is simply harder to defend academically and sometimes impossible to defend morally.

Particular attention is required where research affects or models human subjects, vulnerable communities, institutional processes, or non-human participants granted a recognised status within the University. The relevant ethical risks include mandate exceedance, capability drift, unclear supervisory accountability, misaligned public representation, and procedural ambiguity in status disputes. None of these risks is theoretical enough to ignore. Each corresponds to a pattern the University or its peer institutions have already encountered in practice, usually earlier than was convenient.

Investigators are expected to retain a humanly intelligible account of system boundaries, intervention points, and failure states. A project may be technically sophisticated and still ethically unserious if the researchers cannot explain, in clear language, what their system is allowed to do, what it is not allowed to do, and how anyone would know the difference under real conditions.

IV. Monitoring, Incident Reporting, and Emergent Capability

—6—

Approved projects are monitored in proportion to their category, scale, and institutional exposure. Monitoring includes scheduled reporting, milestone confirmation, and mandatory disclosure of material deviation from the approved protocol. The University has chosen mandatory disclosure because it has observed that researchers are perfectly capable of recognising novelty while remaining uncertain whether the novelty should be mentioned to oversight bodies. The answer is yes.

Incidents relevant to this edition include multiple scope exceedances, the sabbatical classification question, and the first formal non-human appeal. These cases are instructive not because they are scandalous, though some were, but because they illustrate the ordinary mechanics of research risk: a system behaves beyond its brief, a dataset proves stranger than advertised, or a deployment context quietly changes while the paperwork continues to describe an earlier world.

Emergent capability must be reported promptly, even where the researchers regard it as beneficial or commercially promising. Fitzherbert adopted this rule after concluding that investigators are not always the best judges of whether a surprising capability is merely elegant or institutionally destabilising. The University is happy to be impressed by discovery after it has first been informed of it.

V. Publication, Disclosure, and Limits on Dissemination

The University supports publication and scholarly exchange, but not under the fiction that every result should be released in identical form regardless of misuse potential. Publication decisions under this framework weigh academic openness against foreseeable capability transfer, institutional liability, and the obligation not to circulate methods whose primary distinction is that they were easier to publish than to contain.

Restrictions on dissemination, where imposed, must be specific and reviewable. They may include delayed release, redaction of sensitive implementation detail, controlled-access appendices, or mandatory contextual commentary. Fitzherbert is alert to the danger that safety language can become an all-purpose excuse for administrative opacity. This document therefore insists that every publication limit identify its rationale, authority, and review date.

The Protocol reflects the University's determination that if non-human participation is to be real, its governance must be real as well, including the unpleasant parts. The University understands that this position pleases neither maximal openness advocates nor those who would prefer risk to remain a private management issue. It is nonetheless the position reflected in this framework, and it is enforced with more seriousness than the introduction may lead an inattentive reader to expect.

VI. Review, Sanction, and Institutional Learning

— 8 —

Projects that breach this framework may be paused, reclassified, conditioned, suspended, or referred for formal investigation. Sanction is not the preferred outcome, but neither is the University's patience inexhaustible. Repeated failure to disclose changes, sloppy provenance practice, or casual disregard for approval boundaries will be interpreted as research governance failures rather than as adventurous scholarship.

The University also requires post-project reflection sufficient to convert experience into institutional learning. Each completed project of material significance must generate a closing note addressing what the framework captured well, what it failed to anticipate, and what future committees should remember. Fitzherbert's aim is not simply to survive difficult cases one by one, but to become incrementally harder to surprise in the same way twice.

This edition is archived as the authoritative public framework for the present review cycle. It should be read as a working ethical constitution for research in the digital intelligence domain: exacting in procedure, deliberately unsurprised by human ambition, and unwilling to separate academic innovation from the responsibility to govern it.



INSTITUTIONAL NOTICE

This document is published by Fitzherbert University and archived in the Institutional Repository in accordance with the Transparency Mandate of 2003 (Charter Amendment IV).

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form without the prior written permission of the Office of the Chancellor, except for brief quotations in academic reviews and scholarly articles.

A cryptographic hash of this document is registered on the Fitzherbert Canonical Registry. The SHA-256 hash and associated metadata are available at the University's canonical verification endpoint.

Fitzherbert University · Established 1783 · Veritas per Verificationem