## Konnaxion — Backend Debug Primer (Django 4 / PostgreSQL 15)

(updated 6 May 2025)

### 1. Stack & Core Components

* **Framework :** Django 4.x (monorepo konnaxion) with custom user model.
* **Database :** PostgreSQL 15 with pg\_trgm enabled; future migration to TimescaleDB for metrics.
* **Async Tasks :** Celery 5 + Redis broker (scheduled jobs in celerybeat table).
* **Auth :** Django builtin + django‑allauth (social providers disabled in prod).
* **Security :** Argon2 password hasher, django‑axes for brute‑force mitigation, CSP headers via SecureHeadersMiddleware.

### 2. Project Topology

konnaxion/

├─ manage.py

├─ konnaxion/ ←settings, urls, ASGI/Wsgi

│ └─ settings/

│ ├─ base.py ←common settings

│ ├─ local.py ←dev overrides

│ └─ prod.py ←prod overrides

├─ apps/

│ ├─ core\_ ←Auth, profiles, audit

│ ├─ ekoh\_ ←Reputation & voting

│ ├─ home\_ ←Ethikos debate

│ ├─ debate\_arena\_ ←Live debate sessions

│ ├─ stats\_ ←Metrics & logs

│ └─ …

└─ scripts/ ←maintenance & data‑migrate helpers

Naming convention: **snake\_case prefix** = domain; migration files follow 0001\_initial.py, 0002\_soft\_delete.py, etc.

### 3. Settings & Middleware Highlights

| Concern | Key setting / behaviour | Why it matters |
| --- | --- | --- |
| **Password hashing** | PASSWORD\_HASHERS = ['django.contrib.auth.hashers.Argon2PasswordHasher', …] | Argon2 binary must be present in Docker image; missing dependency blocks login silently. |
| **Middleware order** | SecurityMiddleware → CorsMiddleware → WhiteNoiseMiddleware | Reordering breaks CORS pre‑flight caching or CSP headers. |
| **Static files** | WhiteNoise serves pre‑compressed Brotli + Gzip; collectstatic output /staticfiles | Skipping collectstatic yields 404s in prod. |
| **CORS** | Allowed origins loaded from env var list; wildcard blocked in prod | Misconfigured list prevents frontend API calls. |
| **Database timezone** | Field timestamps stored **without TZ** except debate statistics (with TZ) | Mixing types causes equality bugs; convert to UTC at query boundaries. |
| **Soft delete** | Every model mixins SoftDeleteMixin (is\_deleted, deleted\_at) | ORM queries must filter is\_deleted=False; otherwise data appears duplicated. |

### 4. Database Schema: High‑Leverage Tables

| Table | Purpose | Critical Columns |
| --- | --- | --- |
| core\_customuser | User auth & profile | language\_preference, JSONB device\_details, is\_superuser |
| ekoh\_expertisetag | Taxonomy for reputation | GIN name gin\_trgm\_ops for fuzzy search |
| home\_debatetopic | Debate topics metadata | title unique via custom signal, scheduled\_for timestamptz |
| debate\_arena\_argument | Hierarchical arguments | parent\_id self‑FK enables nested replies |
| stats\_debatestatistic | Time‑series metrics | Planned hypertable once Timescale enabled |

### 5. Migrations & DDL Ownership

* **Role policy :** superuser ➝ konnax\_admin (schema owner), runtime ➝ konnax\_rw (app).
* All migrations run as **schema owner**; manual DDL must ALTER OWNER to avoid mixed ownership.
* Migration guards – wrap Timescale calls in if extension\_available('timescaledb').

### 6. Celery & Background Jobs

* CELERY\_TASK\_ALWAYS\_EAGER = False except in tests.
* Default queue **“default”** plus **“long‑running”**; tasks route via queue attribute.
* Healthcheck route /api/health pings Celery by dispatching a no‑op and waiting ≤ 3 s.
* Beat schedule YAML in celerybeat\_schedule table; editing schedule requires service restart.

### 7. Testing & Tooling

* **pytest‑django** with fixtures autouse database (transactional\_db).
* **Factory Boy** for model factories in tests/factories/.
* **coverage.py** minimum 85 % enforced in CI.
* **django‑debug‑toolbar** enabled only when DEBUG=True and INTERNAL\_IPS allows.

### 8. Known Pitfalls & Debug Checklist

1. **Argon2 libs** missing ⇒ silent auth failure; verify argon2‑cffi‑bind in requirements.
2. **Soft‑delete leakage**: always filter is\_deleted=False in custom SQL or .raw() queries.
3. **Timezone mixing**: cast timestamptz to UTC before comparison with naive timestamps.
4. **pg\_trgm on replica**: CREATE EXTENSION must exist in **every** DB including replicas.
5. **Long‑TX vacuum**: idle DB sessions during Celery tasks delay autovacuum; ensure tasks close connections.
6. **Migration conflicts**: squash after three cycles to keep django\_migrations table small.
7. **WhiteNoise caching**: change STATICFILES\_STORAGE invalidates etags; clear CDN cache.
8. **CORS list drift**: stale env var requires pod restart; check at runtime via /api/health?verbose.

### 9. Essential Management Commands

| Command | Purpose |
| --- | --- |
| python manage.py collectstatic --noinput | Build static assets for WhiteNoise. |
| python manage.py migrate --database=default | Apply schema migrations. |
| python manage.py rebuild\_trgm\_indexes | Recreate gin\_trgm\_ops after bulk‑loads. |
| python manage.py dumpdata --exclude auth.permission --exclude contenttypes > seed.json | Lightweight seed export. |
| celery -A konnaxion worker -Q default,long-running --concurrency=4 -l info | Start Celery workers. |
| celery -A konnaxion beat -l info | Scheduled tasks. |

### 10. Roadmap (Backend)

| Version | Planned Change | Impact |
| --- | --- | --- |
| **1.1** | Enable TimescaleDB, convert stats\_\* to hypertables | Improves metrics queries; requires new role privileges. |
| **1.2** | Introduce core\_auditlog table (CDC pattern) | Heavy insert volume; ensure monthly partitioning. |
| **1.3** | Shard debate\_arena\_argument by session\_id | Reduces lock contention during live debates. |

This primer focuses on ownership, soft‑delete safeguards, timezone consistency, extension dependencies, and background processing—recurring pain points unearthed during past incidents. Feeding these details to an AI assistant enables more precise diagnostics and migration advice.