

Past the static deadlines: NeTEx on five European access points, checked against open schemas

EasyxLab · EasyByte Hub S. Coop. Mad. · Study S10

 Working draft Published 2026-10-03

Status: working draft, not peer-reviewed · EasyxLab, study S10 · snapshot of 2026-10-02

EasyxLab · a research lab by EasyByte

Abstract

The deadlines. Under Delegated Regulation (EU) 2017/1926 as amended by 2024/490 (MMTIS):

- every static-data deadline, where NeTEx applies, has passed (2019–2025);
- the public-transport real-time deadline on the comprehensive TEN-T passed on 1 December 2025;
- the 1 December 2026 date covers parking and vehicle sharing.

Since February 2026 the TEL TSI, Regulation (EU) 2026/253, has made CEN/TS 16614-4:2025 (EPIP) mandatory for rail passenger timetables, with NeTEx milestones in 2027 and 2028.

What we measured. On 2026-10-02 we censused 284 NeTEx feeds (639 files) on five national access points (France, Norway, the Netherlands, Belgium, Luxembourg) and validated 41 of 44 sampled datasets.

XSD results for the 34 timetable datasets:

- 8 are fully valid, all of them French;
- 6 are valid on the part checked, which ranges from 2.4 % to 66 % of their bytes;
- 16 are invalid and 4 were not checked.

Two schema details explain much of this:

- NeTEx 2.0.0 rejects 11 French timetables, 9 of which pass 1.3.2. The reason is that every passing time now needs an `id` and a stop-point reference.
- Ten of eleven Dutch datasets fail 2.0.0 only on references to objects outside the document.

EPIP. No dataset passes the archived 2021 EPIP XSD (0 of 36). That XSD predates the 2025 standard, so this says little about conformance to it.

Consistency checks.

- One feed is stale: Dutch AVV, valid until 2018.
- Six strings are double-encoded and one character was lost.
- Eleven of fourteen French datasets follow the profile's proposed identifier codification for at least 99 % of their ids.

- 248,286 internal references lack the `version` the French profile requires. 81 % of them come from one dataset.

Our own instrument. Our linter produced two false positives, now fixed.

We release `netex-lint` as a prototype.

1. The legal frame, read literally

Texts were fetched from the Publications Office Cellar and searched with `grep`. METHOD.md §2 has every quote. In short:

- **Static data (Art. 4(3) of 2017/1926 as replaced by 2024/490).** The deadlines run from 1 December 2019 (comprehensive TEN-T) to 1 December 2023 (rest of the network), 2024 (transport on demand) and "(f) ... point 1.4 ... by 1 December 2025". All have passed. NeTEx is one admissible format among several ("or any digital machine-readable format that can be proven fully compatible and interoperable"), represented through "minimum EU profiles or national profiles". 2024/490 never names EPIP.
- **Dynamic data (Art. 5(3)).**
 - **Annex 2.1** ("real-time status information, such as estimated departure and arrival times ... – for scheduled transport"), where SIRI applies, was due on the comprehensive TEN-T "by 1 December 2025".
 - **Annex 2.2**, due "by 1 December 2026", is "information service on parking tariffs" and "car-sharing ... bike-sharing ... scooter-sharing availability and location ... car parking spaces available".
 - Both points are due on the rest of the network by 1 December 2028.
- **Registration is lawful.** APIs must be "publicly accessible to data users, where relevant subject to registration" (Art. 4(4)).
- **Rail (TEL TSI, Implementing Regulation (EU) 2026/253).**
 - Art. 25 repeals the TAP TSI that MMTIS cites, and references to it "shall be construed as references to this Regulation".
 - Appendix C.4.A [P.4] lists "CEN/TS 16614-4:2025 ... Part 4: Passenger Information European Profile" for rail passenger timetable data (4.2.1), connection times (4.2.2) and "sharing with other modes of transport" (4.9).
 - Appendix G milestones: 4.2.1 passenger timetable data **14.12.2025**, connection times **12.12.2027**, tariff data **10.12.2028**.

So EPIP is law for rail, in an edition whose text is sold and whose only open schema (the 2021 Data4PT XSD) is archived. The question we can answer is not "will NAPs be ready?". It is whether the NeTEx already published validates and is internally consistent, using open schemas and open rules only.

2. Prior work

Searched 2026-10-02/03 by API (`data/prior_work.csv`):

- arXiv: `all:NeTEx` gave 1 hit; `all:"national access point" AND all:transport` gave 3.

- Crossref: the top hits are catalogue records of the standard's parts.
- Semantic Scholar: no data without a key, so DOIs were looked up directly.
- GitHub code search `netex repo:etalab/transport-site validator` : 60 hits.
- NAPCORE deliverables page.

Closest results:

SOURCE	CLOSEST FINDING
Vidović, Mandžuka, Šoštarić (2019), ELMAR	Conceptual: "a brief analysis of regulations that define requirements ... on a conceptual level"; no measurement
arXiv:2011.06423 (2020)	Converts Italian and Spanish data to NeTeX; "These standards are complex and of limited practical adoption."
arXiv:2310.14054 (2023), NAPCORE	"progress has been made in standardizing NAP data, integration with operational ITS practices remain limited"
French NAP, <code>etalab/transport-site</code>	Validates every French NeTeX upload: "Validator for NeTeX files calling enRoute Chouette Valid API"; changelog 0.2.2 "Uses the profile <code>pan:french_profile:2</code> "
<code>Muspah/netex-belgium</code> (created 2026-09-05, last commit 2026-09-24, no licence)	"This repository validates Belgian NeTeX EPIP exports": whole-Belgium export at <code>data.gtfs.be</code> , with streaming validation
Entur <code>netex-validator-java</code> / <code>antu</code> , Fintraffic VACO	Nordic-profile validation (library and services)
NAPCORE	"National Body Compliance Assessment Guidelines" and "self-declaration forms": compliance is self-declared, not measured on published data

Novelty, narrowed. National validation exists:

- France validates every upload;
- the Nordic tools validate the Nordic profile;
- a public repository already validates Belgium's whole EPIP export.

We found no cross-country measurement that applies **one** yardstick (two public NeTeX releases, the archived EPIP XSD and open-text rules) to five NAPs and reports results by content type, nor a literal re-reading of the deadlines. `netex-lint` adds cross-file reference checks with cited open rules under Apache-2.0. Its profile-neutral rules are few, and its two bugs (§5.4) sat in exactly those.

3. Validator landscape

TOOL	SCOPE	LICENCE	LAST COMMIT
Greenlight, skinkie/DATA4PTTools	EPIP-oriented	MIT	2022-12-06 (Docker image 2023-04-27)
entur/netex-validator-java , entur/antu	Nordic	EUPL-1.2	2026-09-28, 2026-10-01
tmfg/digitraffic-tis-vaco	GTFS + Nordic NeTEx	EUPL-1.2	2026-09-25
enroute-mobi/chouette-core	French profile	AGPL-3.0	2026-08-07
theoremus-urban- solutions/netex-validator	claims "the EU NeTEx Profile"	MIT (README badge: EUPL-1.2)	2025-09-01
Muspah/netex-belgium	Belgian EPIP export	none	2026-09-24
enroute-mobi/netex-cli- validator	XSD only	Apache-2.0	2019-05-21

Any XSD validator (lxml, xmllint) is free and profile-neutral. Beyond XSD, the maintained tools are national. The NAPCORE catalogue lists no NeTEx validator and asks "Know a validator for ... NeTEx, SIRI ...?".

4. Method in brief

Details are in `METHOD.md`.

- Census.** Anonymous catalogue APIs of five NAPs, plus Mobilithek metadata. `data/nap_access.csv` records the other probes:
 - Spain (401) and Sweden (403 without key) need registration, which is lawful;
 - Finland, Austria and our first Swiss probe were not assessed (a session cookie, wrong endpoints);
 - a 1 kB `Range` request shows that the Swiss national timetable is anonymously downloadable, at 663,744,713 bytes, above our 300 MB limit.
- Sample.** 44 datasets with written rules:
 - FR: random within size terciles from 131 eligible resources;
 - NO: three regional authorities and three rail operators;
 - NL: the newest file per operator;
 - BE: every direct download;
 - LU: the newest snapshot.

3. **XSD.** NeTEx 1.3.2 and 2.0.0 (pinned by SHA) and the archived EPIP XSD, per document, within per-dataset byte budgets.

4. **Consistency.** One streaming pass per dataset, with every rule cited in `prototype/RULES.md`.

Results are reported by **content class**: timetable, stops, parking, scooter-sharing, codespace list.

5. Results

5.1 What the NAPs publish

NAP	FEEDS (FILES)	MB	MEDIAN MB	UPDATED ≤ 30 D	MACHINE- READABLE LICENCE
FR transport.data.gouv.fr	178	775 (144 sized)	0.71	115	169
NO Entur (EEA)	63	539, incl. a 290 MB national aggregate of the others	0.11	61	0
NL NDOV Locket	17 (62)	38	0.79	14	0
BE transportdata.be	7	270	15.4	4	1
LU data.public.lu	1 (311 weekly snapshots)	27	26.7	1	1 (CC0)
DE Mobilithek (metadata)	18 offers	—	—	9 created	9

The German offers matching "NeTEx" are mixed:

- 6 scheduled public transport;
- 10 car/bike sharing;
- 2 railway, one of which is a SIRI facility feed.

The data.europa.eu `netex` format facet finds French datasets only, because format metadata is not harmonised.

5.2 XSD validation, by content

Best verdict of the two NeTEx releases, for the 41 completed datasets:

CONTENT	N	VALID	VALID ON CHECKED PART	INVALID	NOT CHECKED
Timetable	34	8	6	16	4
Stops only	3	3	0	0	0
Parking	2	2	0	0	0
Scooter operator	1	1	0	0	0
Codespace list	1	0	0	1	0
All	41	14	6	17	4

What "valid" covers. Validity is not a single property here:

- **Fully valid timetables.** All 8 are French, and 7 of them are valid only against 1.3.2.
- **"Valid on checked part".** The checked share is small for some feeds: Belgian TEC 2.4 % and STIB 10.3 % of their bytes. Luxembourg 45.5 %, Norwegian INN 61.4 % and two French feeds 61–66 %.
- **Not checked.** The 631 MB SNCF national file exceeds every cap. Dutch ARR, GVB and Keolis are single documents of 185–358 MB that the memory guard skipped on this 8 GB machine. They were "not checked" first because we lowered the per-document cap mid-run (400 → 150 MB); re-checking them under the original cap was not possible within the free memory.

The failures follow profiles:

- **NeTEx 2.0.0 versus French timetables.** 2.0.0 makes `TimetabledPassingTime/@id` mandatory and `PointInJourneyPatternRef` no longer optional. French passing times carry neither, so 11 French timetables fail 2.0.0; 9 of them pass 1.3.2.
- ****Dutch BISON (ntx:1.1).****
 - **1.3.2** rejects `privateCodes` (11 datasets).
 - **2.0.0** accepts the structure: 10 of the 11 Dutch datasets invalid under 2.0.0 fail **only** on unresolved keyrefs. These point to objects outside the document (the national stop registry CHB, profile `TypeOfFrame s`, codespaces), which per-document XSD validation cannot tell apart from broken references. PNB also has 35 in-file duplicate keys.
- ****Nordic (1.15 / 1.16).**** 1.3.2 rejects `ServiceJourneyRef` and 2.0.0 rejects `DatedServiceJourneyRef` (3 datasets each). The data fit neither public release.
- **EPIP (archived 2021 XSD).** 0 of 36 checked datasets pass. The rejection is usually the first divergent element of each document: `GeneralFrame` (18 datasets; the French profile is built on it), `validityConditions` (12) or `privateCodes` (9). This is a profile-design mismatch with an old schema. Luxembourg (frames typed `epip:EU_PI_*`) and TEC ("EPIP Profile", 2.4 % checked) fail it too.

Error counts are lower bounds per category. Only 20,000 errors per document were categorised: 928,920 of 3,357,691 under 2.0.0. We therefore report datasets.

- **Stale.** One feed: Dutch AVV, whose newest file dates from 2018 and whose frame validity ended on 2018-12-08. Of the 41 datasets, 12 yield a calendar end date; none ends before the snapshot.
- **Encoding.**
 - Six double-encoded strings in two French datasets, e.g. «Provence-Alpes-Côte d'Azur» in the SNCF file.
 - One lost character (U+FFFD) in Dutch EPIAP ("Port Zeevlakte").
 - No document declares a non-UTF-8 encoding. One (Voi) has no XML declaration, which means UTF-8.
- **Duplicates.** The same id and version is defined twice in one document in 3 datasets (429 times).
- **References.** No versioned reference fails to resolve at dataset level. This holds partly by construction: `version="any"` and `TypeOf...Ref` are excluded. At least 15 such `TypeOf...` references with an explicit version resolve nowhere; they are counted only in the 20 re-scanned records. Unversioned references that resolve nowhere were not checked against the national registries.
- **** version on PublicationDelivery **** STIB and TEC (940 documents), Indigo and AVV omit it. The attribute is optional in the XSD: this is a fact, not an error.
- **French profile, applied to the 14 French datasets detected as such.** SNCF declares only `1.09` and was not tested. The profile *proposes* `[CODESPACE]:[type d'objet]:[identifiantTechnique]:[LOC ou Nom attributaire]` and adds that "d'autres structures peuvent être utilisée ... pour peu que l'unicité soit conservée au niveau national".
 - **The proposal is followed almost everywhere.** In 11 of 14 datasets, only 1–57 ids out of thousands depart from it, mostly frame and header objects. Two stop-only files (99.9 %) and one timetable (30 %) use another structure, which the text allows.
 - **The firmer rule is not.** "Version de l'objet référencé ... Doit systématiquement être instancié pour un objet présent dans le jeu de donnée". It is broken by 248,286 internal references in all 14 datasets: 200,428 (81 %) in one dataset (ZOU!, 100 % of its internal references), and a median of 2.9 % in the other 13.
- **Nordic.** The four Nordic datasets meet the quoted Nordic id and versioning rules.

5.4 Our instrument's errors

The independent review found two false positives in our linter:

1. **Open-ended validity was read as expired.** It flagged the Möbius stop file (5,554 of 5,555 `ValidBetween` without an end date). The NAP's own metadata says `no_validity_dates: true`. Rule v4 now uses frame-level validity, and open-ended validity never expires.
2. **A missing XML declaration was counted as "non-UTF-8"** (Voi). It now counts as UTF-8.

Both are covered by new tests (16 in total). The review also found that we had quoted the French identifier rule selectively; §5.3 now quotes it in full.

6. Discussion

On these five NAPs the static NeTeX is mostly current and internally consistent. It is not interchangeable at schema level:

- each profile validates against a different NeTeX snapshot;
- the newest public release (2.0.0) rejects the largest corpus we tested;
- the only open EPIP schema predates the EPIP now mandated for rail.

A consumer combining NAPs across borders therefore validates nationally. Germany (DELFI), Spain, Sweden, Italy and Switzerland are absent from this measurement, and they are among the largest NeTeX producers.

The 2026 MMTIS date is for parking and sharing data. The static counterpart of that data appears here as NeTeX parking and operator files (Indigo, Interparking, Voi): small, and XSD-valid. For rail, the TEL TSI milestones (12.12.2027, 10.12.2028) concern EPIP and tariffs.

****What `netex-lint` would need to be useful:****

- SIRI checks, including resolving SIRI references against the NeTeX ids;
- an open schema or open rules for CEN/TS 16614-4:2025, so that EPIP conformance can be checked against the TEL TSI milestones;
- a disk-backed index, so that memory no longer grows with the size of the dataset.

Until then it stays a prototype.

7. Competing interests

EasyByte, which runs EasyxLab, builds software and could build a NeTeX/SIRI validation tool or service. No such product exists today. `netex-lint` is an unpackaged Apache-2.0 prototype. No funding was received, and no NAP, operator or vendor was contacted.

8. Automation and review

- **Who did what.** An LLM-based agent supervised by EasyByte:
 - fetched the legal and profile texts and **selected the quotes**;
 - wrote the scripts and the prototype, and ran them;
 - drafted this text.
- **What is mechanical.** The tables come from `scripts/04_aggregate.py`. `scripts/05_check_headlines.py` recomputes every headline number in this paper and the README from `data/` and asserts it. The prose and the choice of figures are the agent's.
- **Independent review.** A second, independent agent reviewed the study adversarially. It re-validated three datasets with identical XSD verdicts, reproduced the aggregation byte for byte, and found the

errors corrected here: the deadline premise, the TEL TSI omission, the two lint bugs, the selective French quote and the access claims.

9. Limitations

- **Coverage.**
 - Five NAPs.
 - Ruter, Skyss and SNCB (BE) did not complete: the scan keeps every id and reference in memory (about 6.9 M references for TEC), up to five workers ran in parallel, and the 8 GB machine swapped.
 - Four large single documents were never XSD-checked.
 - **Mixed parameters.** The 16 earliest records used 300/400 MB budgets and the others 150/150 MB (stored per record). A full re-run with today's defaults would not reproduce the early records.
 - **Releases, not intent.** We validate against public releases, not against the snapshot each publisher targets.
 - **Rule versions.** Twenty records were re-scanned after the review (18 under rule v3, 2 under v4); 21 were upgraded from their stored summaries. The frame-level validity rule (v4) was re-run only on the two datasets that earlier rules flagged; elsewhere, object-level validity extends past the snapshot.
 - **Calendars.** `DayTypeAssignment` patterns are not expanded.
-

10. Data and reproducibility

`data/` holds derived files only:

- census and access (`catalogue_netex.csv` , `catalogue_summary.csv` , `nap_access.csv`);
- the sample;
- one JSON record per dataset;
- the aggregates and `summary.json` ;
- prior-work and landscape CSVs.

Raw catalogue responses contain third-party contact details and are not published. No NeTEx dataset is redistributed. Re-run: `scripts/run.sh` (or `SKIP_DOWNLOAD=1` to re-aggregate and re-check the headlines). Code: Apache-2.0. Data and text: CC BY 4.0, which does not cover third-party metadata quoted in `catalogue_netex.csv` .

Cite as: EasyxLab (2026). *Past the static deadlines: NeTEx on five European access points, checked against open schemas*. Study S10. EasyByte Hub S. Coop. Mad. <https://github.com/easybytehub/easyxlab>

References

- Delegated Regulation (EU) 2024/490 amending (EU) 2017/1926, CELEX 32024R0490; Implementing Regulation (EU) 2026/253 (TEL TSI), CELEX 32026R0253. Publications Office Cellar, retrieved 2026-10-02/03.

- NeTEx XSD, github.com/TransmodelEcosystem/NeTEx, v1.3.2 ([4f42794](#)) and v2.0.0 ([a94e5e1](#)), GPL-3.0. NeTEx-Profile-EPIP XSD, commit [e5eaf83](#) (archived).
- Entur, *Nordic NeTEx Profile* (Confluence, versions of 2025-11-05 and 2025-11-01). *Profil NeTEx France — Éléments communs*, normes.transport.data.gouv.fr. Both retrieved 2026-10-02.
- Vidović K., Mandžuka S., Šoštarić M. (2019), ELMAR, doi:10.1109/elmar.2019.8918828. Mandžuka S. et al. (2020), TELFOR, doi:10.1109/telfor51502.2020.9306634. arXiv:2011.06423, 2310.14054, 2010.12036.
- etalab/transport-site ([apps/transport/lib/validators/netex/](#)); Muspah/netex-belgium; NAPCORE tool catalogue and deliverables; repositories in [data/validator_landscape.csv](#) .

Cite this study

CITATION

EasyxLab (2026). Past the static deadlines: NeTEx on five European access points, checked against open schemas. Study S10. EasyByte Hub S. Coop. Mad.
<https://github.com/easybytehub/easyxlab/tree/main/studies/s10-netex-naps>

BIBTEX

```
@techreport{easyxlab_s10,
  title      = {Past the static deadlines: NeTEx on five European access points,
checked against open schemas},
  author     = {{EasyxLab}},
  institution = {EasyByte Hub S. Coop. Mad.},
  number     = {S10},
  year      = {2026},
  url       = {https://github.com/easybytehub/easyxlab/tree/main/studies/s10-
netex-naps}
}
```