COURAGE GARIKAI

FACTS .JSON AS INFRASTRUCTURE

v1.0 / CC BY / HARARE 2026

For the volunteers and maintainers who keep the boring data accurate.

For every developer who has ever fixed a country list at 2 AM. For every civic technologist who has scraped a government website because the data should have been public.

Facts are infrastructure.

Facts.json as Infrastructure

Treating Public Data as Code

Courage Garikai

ColdCode Si Labs · Harare

About this book

There is a category of software infrastructure that almost no one talks about: the small, structured, factual data files that make the rest of the internet work.

I am talking about files like `countries.json` (a list of every country, with ISO codes, currencies, capitals, phone codes). Files like `currencies.json` (every world currency, with exchange rate metadata, subunit information, symbols). Files like `languages.json` (ISO

639 language codes and names). Files like `timezones.json`, `postal-code-regexes.json`, `holidays.json`, `tax-rates.json`, `airports.json`, `medical-codes.json`, `zip-codes.json`, `unit-conversions.json`, `ng-counties.json`, `za-provinces.json`, `zw-districts.json`.

These files are not glamorous. They are not "artificial intelligence." They are not distributed systems, blockchains, or real-time data pipelines. They are static JSON (or CSV, or YAML, or XML) files — structured collections of facts about the world, stored in a format computers can read, version-controlled, maintained, and published for other software to depend on.

And yet almost every non-trivial software application depends on at least one of these files. If you have ever filled out a country dropdown on a web form, you have used `countries.json`. If you have ever been shown a price in your local currency, that is `currencies.json` at work. If you have ever seen a public holiday marked on a calendar, that is `holidays.json`. If you have ever been asked for your postal code and the form validated it, that is `postal-code-regexes.json`. These files are everywhere; they are invisible; they

are critical infrastructure; and they are surprisingly poorly made.

This book is about why these facts-as-data-files matter, why they are a form of infrastructure as critical as DNS or TLS or open-source libraries, why most of them are badly maintained, out of date, incomplete, wrong, unversioned, or paywalled, and how we can build them properly — as open data, version-controlled, well-maintained, tested, validated, and treated with the same engineering rigour we apply to any other critical software dependency.

I have spent years building and maintaining this kind of data infrastructure for African markets — where the problem is most acute because most public data files are built from a Western perspective and either omit African countries, get them wrong, or treat them as afterthoughts. I maintain `zw-districts.json` (a structured list of Zimbabwe's 10 provinces and 64 districts with administrative codes, coordinates, and population estimates), `zw-schools.json` (a list of all registered primary and secondary schools in Zimbabwe), `african-currencies.json` (currency metadata for all 54 African countries with current exchange-

rate regimes and mobile-money integrations), `african-mobile-operators.json` (MNO metadata, USSD codes, M-Pesa-style integrations), and several others. These files are published as open data under CC BY licenses, versioned on GitHub, validated with CI, and depended on by fintechs, edtechs, health-techs, and government systems across southern Africa.

In this book I argue: - Structured factual data files are infrastructure. They should be engineered, maintained, and respected as such. - Most reference data files are terrible — incomplete, outdated, unmaintained, wrong, paywalled, or lacking provenance. - Building good ones is a solvable engineering problem: schemas, validation, testing, versioning, attribution, and governance. - For Africa and the global periphery, building and maintaining these reference datasets is a prerequisite for functioning software ecosystems. You cannot build reliable fintech if your currency data is wrong. You cannot build reliable health systems if your medical facility list is three years out of date. You cannot build edtech if your school registry is incomplete. - These files should be open, public, freely reusable, and maintained as public goods. Locking reference

facts about the world behind paywalls is rent-seeking on public knowledge.

This book covers: - Part I: Why facts.json is infrastructure — the country dropdown problem, reference data as critical infrastructure, the structural problems with how public data is maintained today. - Part II: How to build good facts.json — schemas, sourcing, provenance, validation, versioning, distribution. - Part III: Specific datasets — countries, currencies, subnational divisions, the African data gap. - Part IV: Governance and community — maintenance, sustainability models, open licensing. - Part V: Using facts.json in production — how to depend on reference data responsibly, how to publish for consumers, the open data flywheel. - Part VI: Call to build — what is missing today, facts as a commons, why African developers must build our own data infrastructure. - Part VII: Case studies — zw-districts.json, the IANA Time Zone Database, holiday data, school registries. - Part VIII: Technical implementation — writing schemas, CI pipelines, repository structure, publishing to package managers, APIs.

This is a technical book, but it is short (~80 pages) and practical. It is written for developers who build and use these kinds of files, for data engineers, for civic technologists, for open-data advocates, for governments thinking about publishing public data, and for anyone who believes that basic factual data about the world should be accurate, accessible, and free.

License

This book is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0). You are free to share (copy and redistribute in any medium or format) and adapt (remix, transform, and build upon the material) for any purpose, even commercially, as long as you give appropriate credit, provide a link to the license, and indicate if changes were made.

The code examples and data structures described in this book are released under CC0 (public domain dedication) — use them freely without restriction or attribution, though attribution is appreciated.

Free sample

This is the opening of "Facts.json as Infrastructure" by Courage Garikai.
The full edition — PDF and EPUB — is at library.coldcode-si-labs.workers.dev

(c) 2026 Courage Garikai. Licensed under CC BY 4.0.