RISC-V Unified Database

Streamlining the Ecosystem with a Centralized Source of Truth

Afonso Oliveira

April 3rd, 2025

About

- Master's student in 

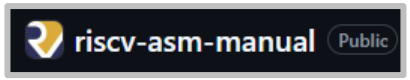
  - Hypervisors and MCS
- Software engineer at SYNOPSYS®
- Active contributor in UDB
  - Responsible for adding and maintaining instructions
  - Data validation against outside sources
- Sponsored by ESA Academy to be here today 

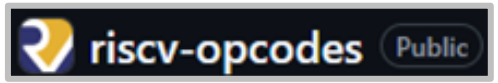
Challenges with the RISC-V Specification

Growth Challenges with the RISC-V Specification

Why this matters

RISC-V ecosystem relies on multiple disconnected specifications



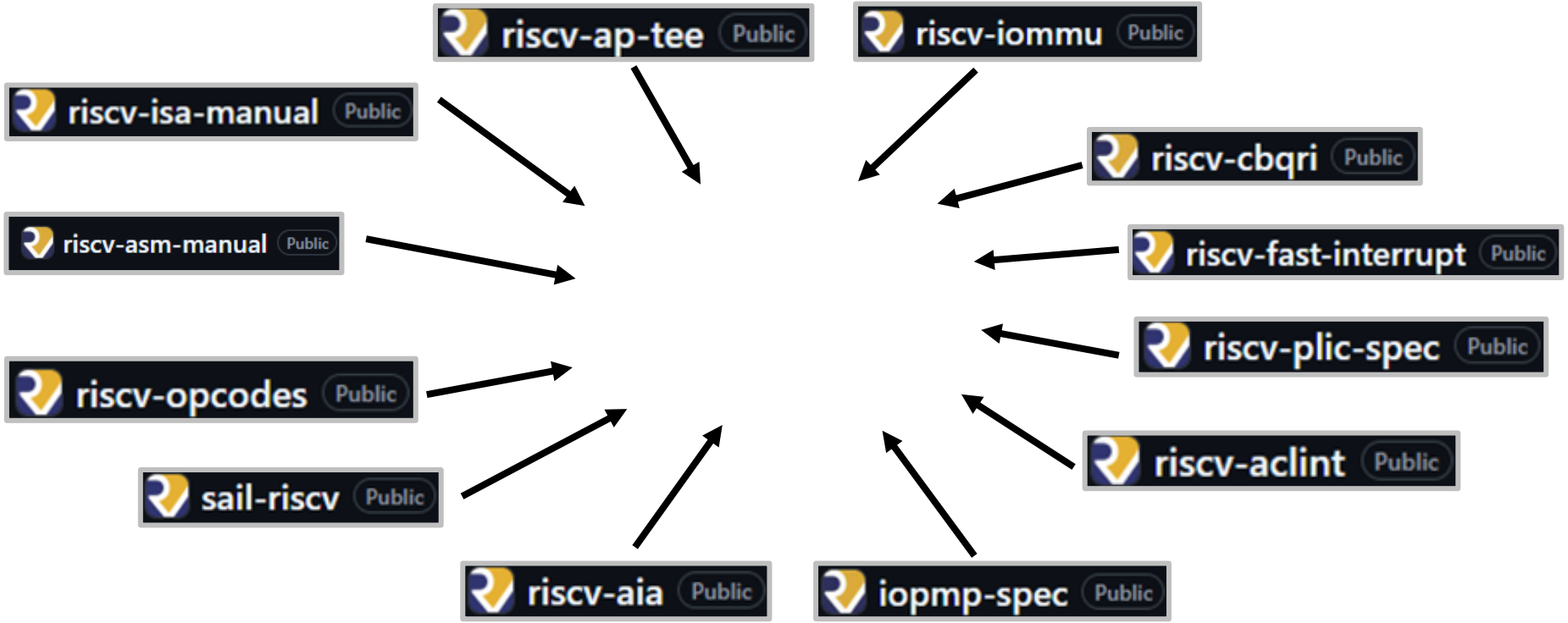


Why this matters

RISC-V ecosystem relies on multiple disconnected specifications

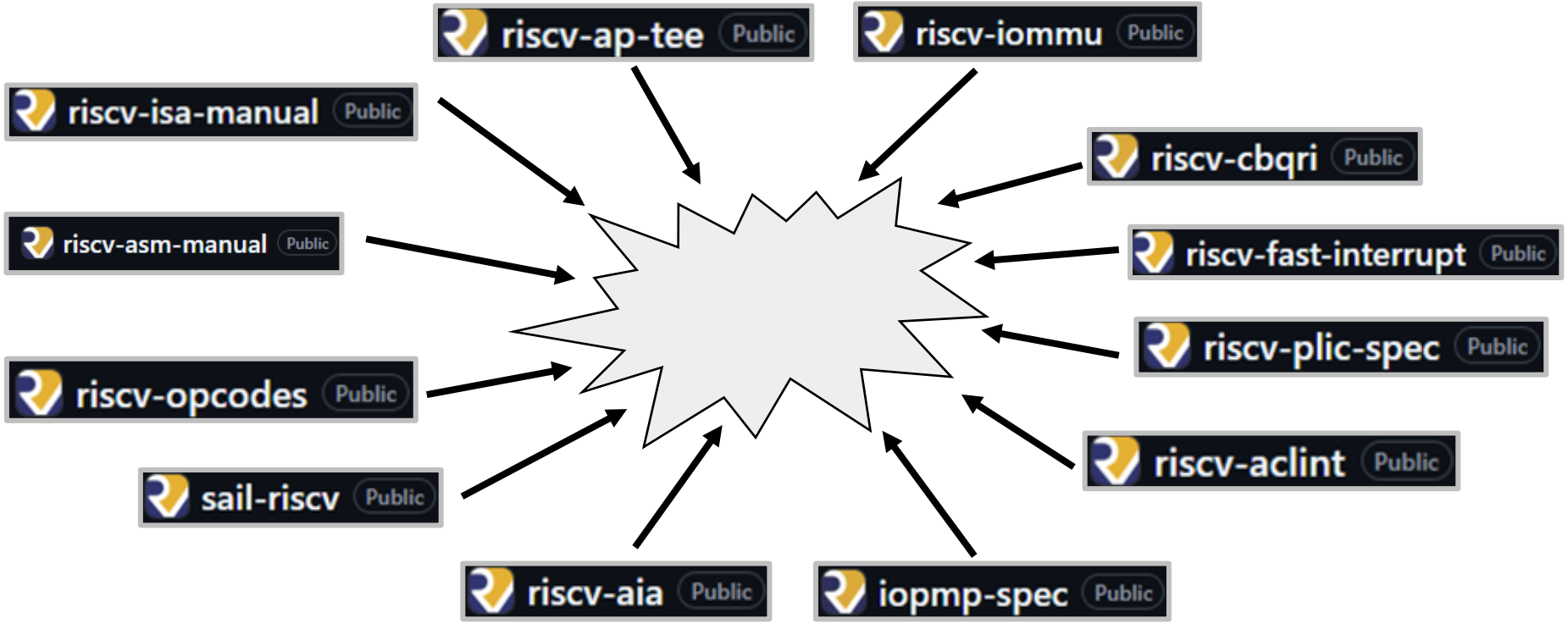
Why this matters

RISC-V ecosystem relies on multiple disconnected specifications



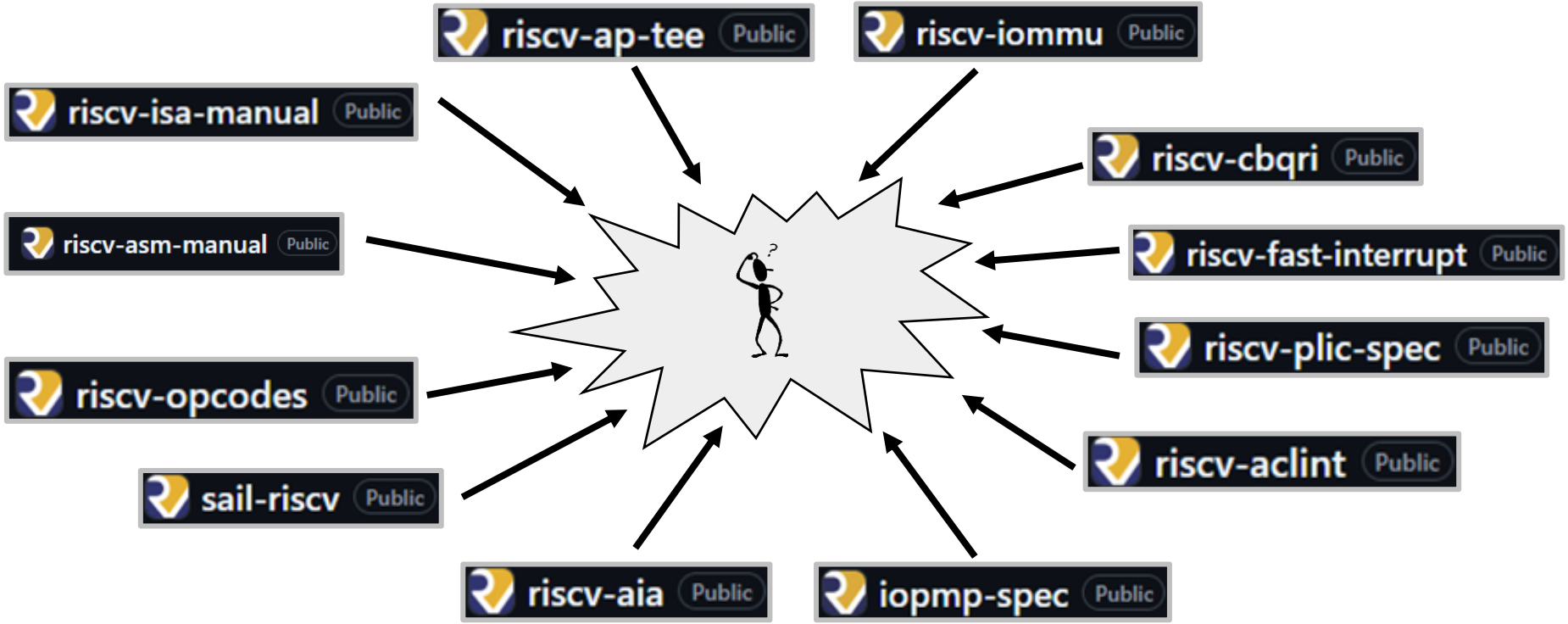
Why this matters

RISC-V ecosystem relies on multiple disconnected specifications



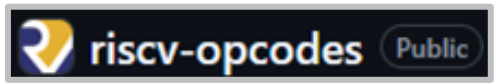
Why this matters

RISC-V ecosystem relies on multiple disconnected specifications



Why this matters

RISC-V ecosystem relies on multiple disconnected specifications

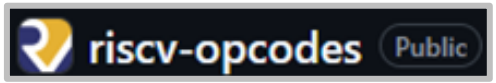


https://github.com/riscv/riscv-isa-manual
https://github.com/riscv/riscv-opcodes

Why this matters

RISC-V ecosystem relies on multiple disconnected specifications

- First published in 2015 – only 5 extensions
- 10 years later (2025) – more than 200 extensions



https://github.com/riscv/riscv-isa-manual
https://github.com/riscv/riscv-opcodes

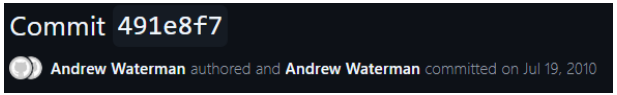
Why this matters

RISC-V ecosystem relies on multiple disconnected specifications

riscv-isa-manual Public

- First published in 2015 – only 5 extensions
- 10 years later (2025) – more than 200 extensions

riscv-opcodes Public

- First created in 2010 - 

- Refactored in 2022 – some previous problems stayed
- From 5 to 200 extensions

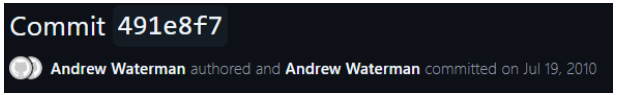
https://github.com/riscv/riscv-isa-manual
https://github.com/riscv/riscv-opcodes

Why this matters

RISC-V ecosystem relies on multiple disconnected specifications

- First publish… 5 extensions
- 10… n 200 extensions

Not a RISC-V problem –
Infrastructure was outgrown



in 2010 -

…actored in 2022 – some previous problems stayed

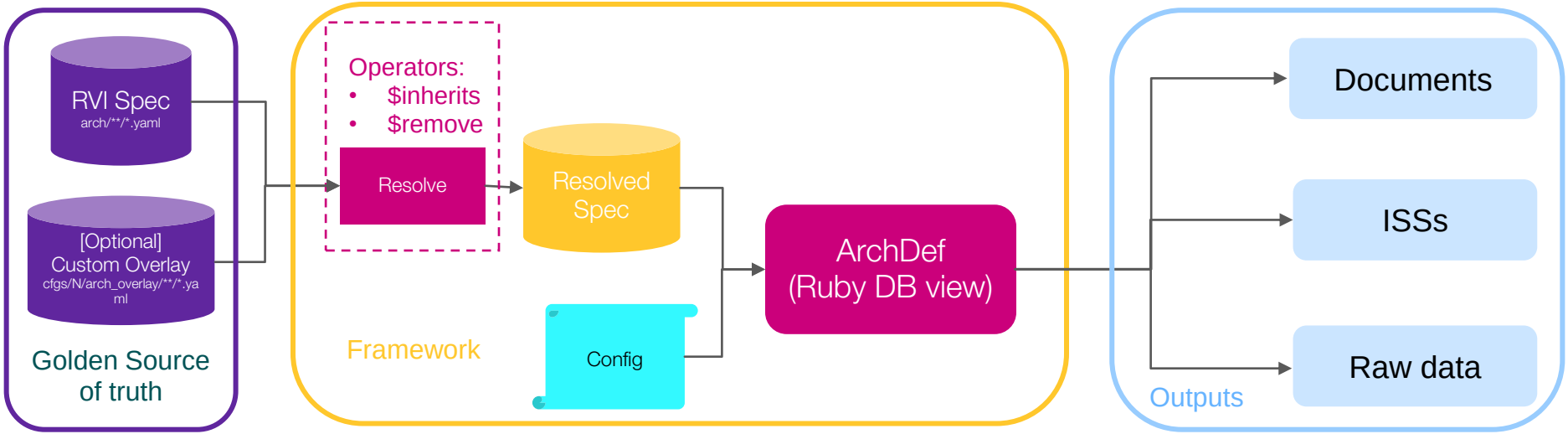
- From 5 to 200 extensions

https://github.com/riscv/riscv-isa-manual
https://github.com/riscv/riscv-opcodes


RISC-V Unified Database (UDB)
RISC-V®

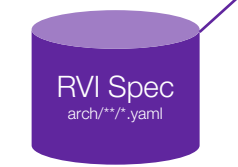
What is the UDB?

1. A centralized, machine readable, source of truth
2. Tools to generate several outputs



Adapted from Derek Hower's ARC Presentation

Golden Source of Truth



Golden Source of truth

	Instructions	CSRs	Profiles	Architectural Parameters
Definition	Operations executed by the processor	Special-purpose state management registers	Subsets of the ISA for specific use cases	Customizable implementation attributes
Purpose	Arithmetic, logic, memory, control	System config, monitoring, exceptions	Compatibility across implementations	Tailored performance and design
Examples	ADD, SUB, LW	mstatus, mtvec, mip	RVI20U32, RVA23U64	XLEN, VXLEN, TRAPS

Instructions

- Name
- Long name
- Description
- Defined by
  - Extensions
- Encoding
  - Match
  - variables
- Access Mode
- Formal Specification (Sail and IDL)

https://github.com/riscv-software-src/riscv-unified-db/blob/main/arch/inst/I/add.yaml

```yaml
kind: instruction
name: add
long_name: Integer add
description: |
  Add the value in rs1 to rs2, and store the
  result in rd. Any overflow is thrown away.
definedBy: I
assembly: xd, xs1, xs2
encoding:
  match: 0000000----------000-----0110011
  variables:
    - name: rs2
      location: 24-20
    - name: rs1
      location: 19-15
    - name: rd
      location: 11-7
access:
  s: always
  u: always
  vs: always
  vu: always
data_independent_timing: true
operation(): X[rd] = X[rs1] + X[rs2];

sail(): |
  {
    let rs1_val = X(rs1);
    let rs2_val = X(rs2);
    let result : xlenbits = match op {
      RISCV_ADD  => rs1_val + rs2_val,
    };
    X(rd) = result;
    RETIRE_SUCCESS
  }
```

Current outputs

Documents

ISSs

Raw Data

Outputs

Current outputs

Documents

ISSs

Raw Data

Outputs

Manuals

- ISA Manual
- Certification Documents
- Profile Manuals
- Extensions Manuals

Detailed Appendices

- Instructions
- Extensions
- CSRs
- Architectural Parameters

Instruction Set Simulator

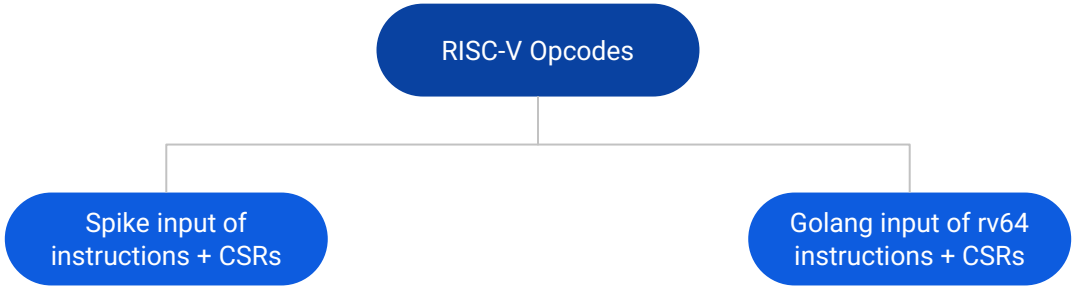
Configurable on UDB parameters

RISC-V Opcodes outputs

- Json
- Encodings.h
- Inst.go

RISC-V Opcodes

```
add  rd rs1 rs2 31..25=0 14..12=0 6..2=0x0C 1..0=3
```



RISC-V Opcodes

```
add  rd rs1 rs2 31..25=0 14..12=0 6..2=0x0C 1..0=3
```

RISC-V Opcodes

```
add  rd rs1 rs2 31..25=0 14..12=0 6..2=0x0C 1..0=3
```

Porting guide to all downstream users is coming soon


UDB and the Space industry
RISC-V®

Current use cases

- Used by the Certification Steering Committee SIG to create certification documents
- Qualcomm created Xqci: an extension that is only available through the UDB
- Synopsys is using the UDB to generate ARC-V processors documentation
- Cross-verification tools - checks against LLVM, Binutils and more

Use cases – for Space

Building any documents **you** may want to

Provide content for new tools

Community Endorsement

UDB is being collaboratively developed

- Several engineers from different companies are actively developing and meeting weekly
  - Antmicro, Akeana, Qualcomm, Rivos, Sifive, Synopsys, Ventana
- We have 7 RVI mentees working on UDB right now! – under official LFX/RVI mentorship platform
- TSC gave initial endorsement to UDB and approved a UDB SIG
  - UDB is engaging with other Working Groups – such as the Doc SIG, TSC , CSC

UDB is being collaboratively developed

- Several engineers from different companies are actively developing and meeting weekly
  - Antmicro, Akeana, Qualcomm, Rivos, Sifive, Synopsys, Ventana
- We have 7 RVI mentees working on UDB right now! – under official LFX/RVI mentorship platform
- TSC gave initial endorsement to UDB and approved a UDB SIG
  - UDB is engaging with other Working Groups – such as the Doc SIG, TSC , CSC

Join us!

- Check our github repository https://github.com/riscv-software-src/riscv-unified-db/
  - Several issues and discussions on going!
  - Ask questions!
- On our weekly meetings – now under LFX/RVI calendar – click here
- Contact me Afonso.Oliveira@synopsys.com


Thank you
RISC-V®


UDB Meetings
RISC-V®

Backup slides

What is the UDB?

Centralization of Disconnected Resources

- Combines scattered resources like ISA manuals, opcode definitions, and Sail formal specifications.
- A single, reliable source of truth for RISC-V developers and vendors.

Machine-Readable, YAML-Based Format

- Standard YAML, **versioned** schema.
- Allows for a lot of use case (ISSs, QEMU-configs…) due to being easy to parse

Consistency and Validation

- Verifiable format
- Cross-Validated against resources like Binutils, LLVM and riscv-opcodes

Artifact Generation

- Generates specific artifacts like PDF/HTML (no more copy paste from the ISA manual)