Elements of Data Portability
Here’s What We Want

- Email.
The Worst Standard Ever

- It is intrinsically insecure
- It breaks all the time
- Nobody can fix it
- Every single person uses it constantly
Why?

- There is exactly one standard
- By the time anyone noticed that it sucked, it was broadly implemented and could not be changed
- The people who controlled the standards document were lazy and left it alone
- Anyone can write an email program and it works, period.
Like SIF in the Old Days

- Google data portability and you still get SIF
- Albeit in articles from a decade ago
- But, there were lots of them
- It actually was *the* standard
- Limit Google to the last few years emphatically does *not* find SIF (only GDPR)
Water under the bridge

- SIF was designed in the nineties
- It didn’t expect broadband to be ubiquitous
- It didn’t expect a supercomputer (by 90’s standards) on everyone’s desk
- It thought government cared about students
- The idea of API had not been invented
Not to put too fine a point...

- SIF was a brilliant step forward but was hobbled by technical choices inevitable at the time that were the seeds of downfall
A Rough Moment

• Awhile ago, someone tried to fix it and they made the classic, XKCD mistake…
HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION:
THERE ARE 14 COMPETING STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE’S
USE CASES. YEAH!

SITUATION:
THERE ARE 15 COMPETING STANDARDS.

[Soon:]
And so, Version 3

- Now, I want to say right up front that, as far as I can tell, Version 3 is better in every way.
- Unfortunately, it was like a prion disease that spread through the system causing irrevocable death and destruction.
- I cannot overstate the depth of my feeling about the horror that is SIF 3.
Maybe I go too far

- The whole thing wasn’t terrible
- The infrastructure spec was a good idea and necessary.
- The old SOAP architecture was obsolete and awful
- The new REST architecture is terrific, flexible, and 100% good.
- Most importantly, that part didn’t poison the system
The road to perdition

- We told the market, SIF 2 is obsolete and old and bad and only used by poopy-heads
- We said that cool, modern people use SIF 3.
- We said that software that uses SIF 2 had to be revised
- We ignored the fact that this was on its face, ridiculous
- It made everyone confused, unmotivated and unable to act on or support SIF
Last year about this time…

- After the meeting last year, a handful of us stayed on for a day of self-flagellation

- Not having the benefit of the summary I just presented, we wondered why we were falling on hard times.

- New adoption of SIF at about zero

- A4L membership declining

- No enthusiasm

- Doom doom doom doom
But we are sooooo smart

- We conceived a radical, amazing, hitherto unknown concept, so incredible that we almost could not believe it had taken us so long to figure it out…
We decided to talk to customers

(The Nobel Prize Committee should be notified)
The Healing Starts

- We put together a schedule of interviews with people who care about SIF
- Mainly, state level education technology managers
- We wanted to find out what works, how SIF is used, what’s wrong, what’s the future, what problems arise
It was GREAT

- Brief digression…
- Reporting mainly. Some outbound services
- Student information systems are run by evil corps
- Much success and goodness
- Kind of bleak on the future
The Motivating Insight

- Everyone loves SIF 3
- Nobody is going to use it
Well, not quite

- People were actively thinking about ways to use SIF 3
- But it was impossible to do
A Digression
A Language Problem

- Back in the day, if you had SIF, you had the data model and the transport. No other way.
- One SIF über alles
- Today transport is separate from data. That’s the SIF 3 innovation. That’s the world around us
- But, we only have one word: SIF
- Even though we have a lot of different things
Data Driven Design

• This is a project management strategy that is entirely unimportant here
• Except for one thing:
• One of the main principles is that a project should have a dictionary of agreed-upon terms
• Prevent confusion and make precise communication possible
This will make the rest of the conversation easier…

The components
  - SIF Data Model
  - SIF Infrastructure

eg, SIF Data Model 2, Sif Data Model 3, SIF Infrastructure 2, SIF Infrastructure 3
  - DM2, DM3, I2, I3
As I was saying...

- People really wanted SIF Infrastructure 3
- And had no intention of ever using SIF Data Model 3
- DM2 worked perfectly well
- They like things about DM3
  - But nowhere near enough to justify switching
- No new vendors even consider the idea of implementing I2
It’s obvious in retrospect

- SIF 2 and SIF 3 are not two aspects of SIF
- They are competing standards
- The existence of SIF3 says SIF2 is old, bad
- But SIF3 is impossible to implement
  - No way to get all vendors to revise their data structures
  - All districts to revise their vendors
- Incompatibility is a prion disease
So we killed it

- An insane person wrote an email entitled, “Let's Kill SIF 3: A Manifesto”

  I propose a different strategy. Take advantage of that enthusiasm by doing something big. Announce a roadmap for SIF 2 based on the idea of upgrading it with the best ideas from SIF 3 starting with the IEP data model. Make clear that the SIF 3 additions to the data model will be based on explicit community input. At the same time, announce that we are deprecating SIF 3.

- There was much arguing
- Eventually we convinced ourselves that we should have
- (drum roll)
The Goal

- Announce a bold, new direction
- A single, consolidated standard
- Strict compatibility with existing implementations
- All the goodness and none of the bad
- An unambiguous, comprehensible direction
- A story that could provide a rallying cry for the future
What we killed

- We have deprecated:
  - SIF Data Model 2
  - SIF Data Model 3
  - Sif Infrastructure 2
What We Kept

- SIF Data Model 4 (DM2 plus)
- SIF Infrastructure 3
SIF Data Model 4

- SIF Data Model 2 *plus*…
- IEP data model (everyone says they want it)
- Some xPress elements already in use at RIC One
- Some other DM3 things needed by Vermont
- Some details to improve transport over Infrastructure 3
- **Most important: Strictly compatible with DM2 implementations**
SIF Infrastructure 3

- It’s nearly perfect
- Everyone loves it
- Can be run in parallel with existing DM2
- Soon will have privacy specifications
- Everybody wants it because it is good and the old way is bad

(Can you imagine that there was a time when SOAP was not reviled?)
SIF Unity

• New standard comprising two elements
  o SIF Data Model 4
  o SIF Infrastructure 3

• Complete CEDS alignment

• Resources will be focused on Unity alone

• Nobody will be coerced to give up old Infrastructure 2
  o And we will help them keep it going if we can
I have to wear sunglasses…

- We have gotten rid of the cruft
- Defined a clear direction
- Heard the needs of our customers and community
- A clear recommendation for those considering educational data implementation
Feeling happy again...

- It is reasonable to ask vendors to support SIF Unity
- Infrastructure 3 is the work of days not months
- Essentially no work to update DM2 to DM4
- We can start to rebuild our community
While I was writing this…

- My state, MN, has chosen to use Ed-Fi
- He’s wondering if Ed-Fi will support certain customer needs for our data warehouse product
- I looked at their Discipline Entity and compared it to the SIF Discipline Object
- An order of magnitude more complete
- Really, no comparison
SIF is Better

- That Discipline Object has been under refinement for thirty years of practical use
- The Discipline Entity has been under development for less than a third of that time
- The SIF object was done by an independent community with a serious intent to seek an ideal of data portability
- The Ed-Fi entity was created by a corporation to sell dashboards
The other ‘standards’…

- Come from corporations that want your money
- They are not data transport standards
- At best, they are interface specifications sponsored by companies who want to use them to lock you in
- They have no intention of making it so that anyone can access the data they need to support students
- They are engineered products to provide the minimum wiggle room to allow them to enter the market
SIF is Different

- SIF provides a toolbox that is under constant improvement by independent education professionals
- It provides extensible data definitions that are consistent with federal CEDS standards
- It provides a transport definition that invites software to interface
- It is comprehensive, complete and proven
SIF Unity

IT IS THE FUTURE