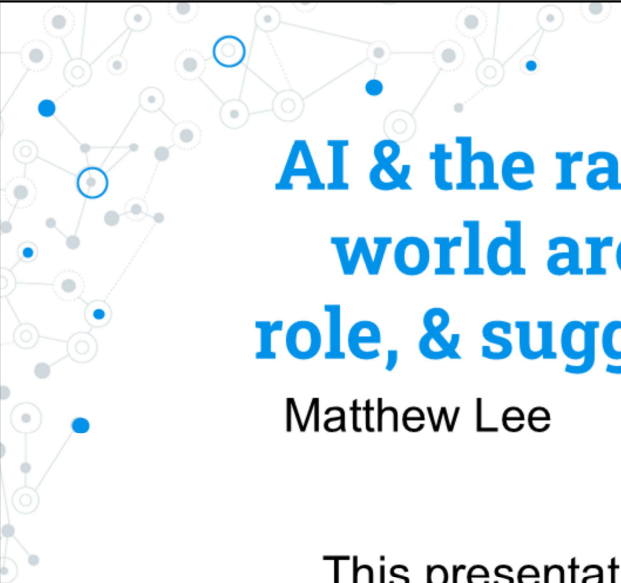


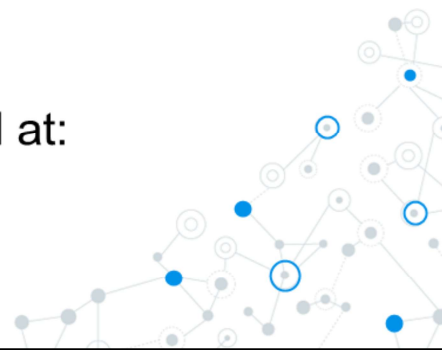


AI & the rapidly changing world around us, SIL's role, & suggestions for GDS

Matthew Lee

September 2, 2026

This presentation can be found at:
<https://cv.mattgyver.net/#GDS>



Matthew, Teresa and Juliette Lee

© SIL Cameroon 2010-2024

© Matthew:

- LangTech Consultant @ SIL Global
- Language Modeling Team
- Adjunct Faculty @DIU

© Teresa:

MK Education



Work in AI

- ◎ I was involved in the creation of Allosaurus, an AI IPA transcription engine.
- ◎ Since moving back to the US in 2024, most of my work has been in and around AI:
 - Advocating, learning, prototyping, problem solving
- ◎ My supervisor is working on changing my title to:
 - Language Technology Innovation Architect

Unstuck

- ◎ I love getting people and projects “unstuck”.
- ◎ I’ve always been a generalist.
- ◎ I bounce between many sciences and technologies, listen to the context of challenges and make uncommon connections.
- ◎ AI (mostly Claude) is just the latest tool in my toolbox.



Recent AI-related Projects



Software Support Agents (<https://build.chipp.ai/#/hq/sil-global-Ituse>)



Language Technology Cons...

I am here to help answer any of your questions about SIL softwar...



Adapt It Support

Adapt It is a free, open-source translation editor designed to hel...



Alpha Tiles Support

Alpha Tiles is an open-source Android app that creates literacy...



Audio Project Manager Sup...

Audio Project Manager is a web and desktop application designed...



BART Support

BART (Bible Adaptation and Revision Tool) is a tool that helps...



Bloom Support

Bloom is a tool for creating books and literacy materials in local...



Cog Support

Cog is a tool for comparative linguistic analysis of related...



Dekereke Support

Dekereke is a specialized linguistic software tool designed for...



Dictionary App Builder Sup...

Dictionary App Builder is a tool for creating mobile dictionary apps f...



DLS Courses

DLS Courses provide comprehensive training in...



FieldWorks Support

FieldWorks is an integrated application for documenting and...



FLExTools Support

FLExTools is an application for running custom Python scripts...



FLExTrans Support



Glyssen Support



HearThis Support

Practical AI for Kingdom Work

<https://podcast.langtech.cloud>



Claude Training (<https://claude.langtech.cloud>)

Claude Day

FR

Search...

OVERVIEW

Getting Started

Installation (Claude Chat and Cowork)

Installation (VS Code, Claude Code, and Git)

The Anthropic Tools

Deterministic Results

CLAUDE CHAT

Claude Chat

CLAUDE COWORK

Workshop Docs

Select a topic from the sidebar, or jump to a section below:

Overview

[Getting Started](#)

[Installation \(Claude Chat and Cowork\)](#)

[Installation \(VS Code, Claude Code, and Git\)](#)

[The Anthropic Tools](#)

[Deterministic Results](#)

Claude Chat

[Claude Chat](#)

Claude Cowork

[Claude Cowork](#)

Language Technology Competency Analytics

Navigation

Select Analysis

- ☐ Overview
- ☐ Staff Certification Report
- ☒ Category Blocker Analysis
- ☐ Weak Spot Analysis
- ☐ User Analysis
- ☐ Misaligned Users
- ☐ Certification Alignment
- ☐ Last Mile
- ☐ Missing Translations
- ☐ Drill Down

Dataset Summary

- Users: 141
- Roles: 5
- Categories: 14
- Competencies: 52

Deploy

Category	Competencies	Users Blocked	Average Gap	Impact
Mentoring		3	94 1.9	177
Adult Education		2	93 1.9	176
Technology for Consulting		2	75 1.9	144
Working in a Multicultural Environment		2	70 1.9	135
Scholarship and Documentation		3	105 1.1	112
Program Design and Engagement		3	97 1	100
Consulting Process Skills		3	85 1	88
Reading to other organisations		2	84 1	85
Interpersonal Skills		3	71 1.1	75
Language and Culture		2	73 1	73
Education		1	1 1	1



Finance Documentation Analytics

LANGUAGES

- ☒ EN
- ☐ ES
- ☐ FR
- ☐ KO
- ☐ PT

Filter documents...

All Documents 1607

Finance Guidance - resource index 4

Accounting for Contributions at Year-End 70

Accounting for Internal Sales & Services in NetSuite 14

Approval Checklist - NetSuite JEs 7

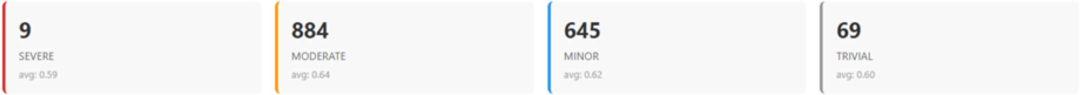
Approved SIL Payment and Reimbursement Methods Update 20

Bank Accounts Form FY25 2

Cash & Bank Reconciliations in NetSuite 10

Policy Contradiction Report

Interactive findings viewer



Dashboard Severe (9) Moderate (884) Minor (645) Trivial (69)

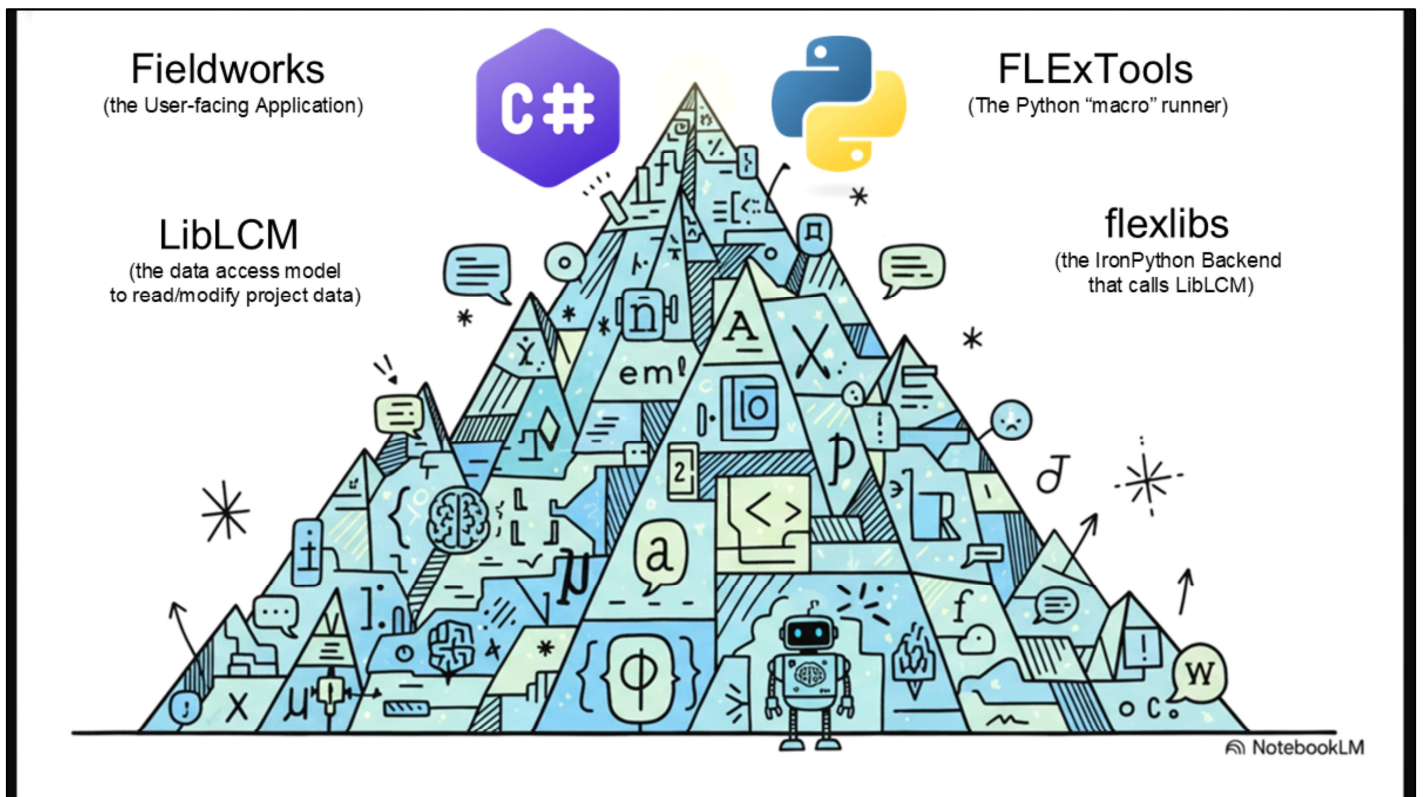
SEARCH: Document name, verdict... MIN CONFIDENCE: 0.00 FINDING TYPE: All Types

Taxes & Tax Credits in Chrome River ↔ PNG GST Info for NetSuite

Document A states that standalone taxes are not recorded with this policy, while Document B doesn't explicitly mention these but is specific to PNG GST which might imply otherwise. **SEVERE** 0.45

Type: direct contradiction

Sections: Overview ↔ Overview



- This project set out with this monster of a goal. Could they make an AI a truly useful coding partner for this super niche, incredibly complex piece of software for linguists? I mean, this isn't just about knowing a standard programming library.
- This is like the final exam for an AI's ability to understand context.

FLExTools MCP

- What are the grammatical categories (POS) in this project and how many entries does each have?

• The Esperanto project has 13 grammatical categories with 1,442 total entries:

Category	Entries
Locution	439
Noun	408
Adverb	336
Verb	217
Root	25
Adjective	9
Preposition	3
Interjection	2
Definite article	1
Conjunction	1
Junk Verb	1
Pro-form	0
Particle	0

Notable observations:

- Locution (phrases/expressions) is the largest category at 439 entries
- Noun, Adverb, and Verb make up the bulk of the remaining entries
- Pro-form and Particle exist as categories but have no entries assigned
- Junk Verb appears to be a cleanup/staging category with 1 entry

Would you like to explore any of these categories further?

Scripture Use Survey Analytics

Claude Code Chat Log

iMessage-style Chat Log

U User Messages: Your direct requests and questions **W** Tool Executions: Commands with inputs and outputs **R** Chat Responses: Analysis and explanations

RESPONSE

Done! I've saved a detailed analysis to `/reports/Data_Findings.md`. Here are the **key discoveries**:

Major Findings

- 1. Critical Gap in Scripture Access** - 86% own Swahili Bibles, but only 53% own mother tongue Bibles - Audio recordings in mother tongue are virtually non-existent (3% listen daily vs. 22% for Swahili)
 - 2. Massive Underutilization of Mother Tongue Scripture** - Only 9% read Scripture daily in mother tongue vs. 40% in Swahili - 66% never read mother tongue Scripture at all - Yet 90% are literate
 - 3. Strong Perceived Impact** - 71% believe mother tongue would improve understanding - 84% of church leaders say it would increase conversions/evangelism - 66% report learning something new when MT Scripture is available - 76-94% want MT materials created
 - 4. Paradoxical Church Language Use** - Mother tongue is used in 80% of announcements, 67% of evangelism, 55% of sermons - BUT only 3% use it for formal Bible reading - Only 14% use it in Sunday School - Suggests cultural belief that Scripture requires "official" language (Swahili)
 - 5. Strong Market Demand** - Requests for: complete Bibles, audio recordings, songbooks, dictionaries, study guides - Church leaders unanimously want translation materials - Gap between what exists (53% have MT Bibles) and what people want (94% want MT materials)
- The survey was clearly designed to **justify and guide mother tongue Bible translation work** by documenting the gap between availability and need.

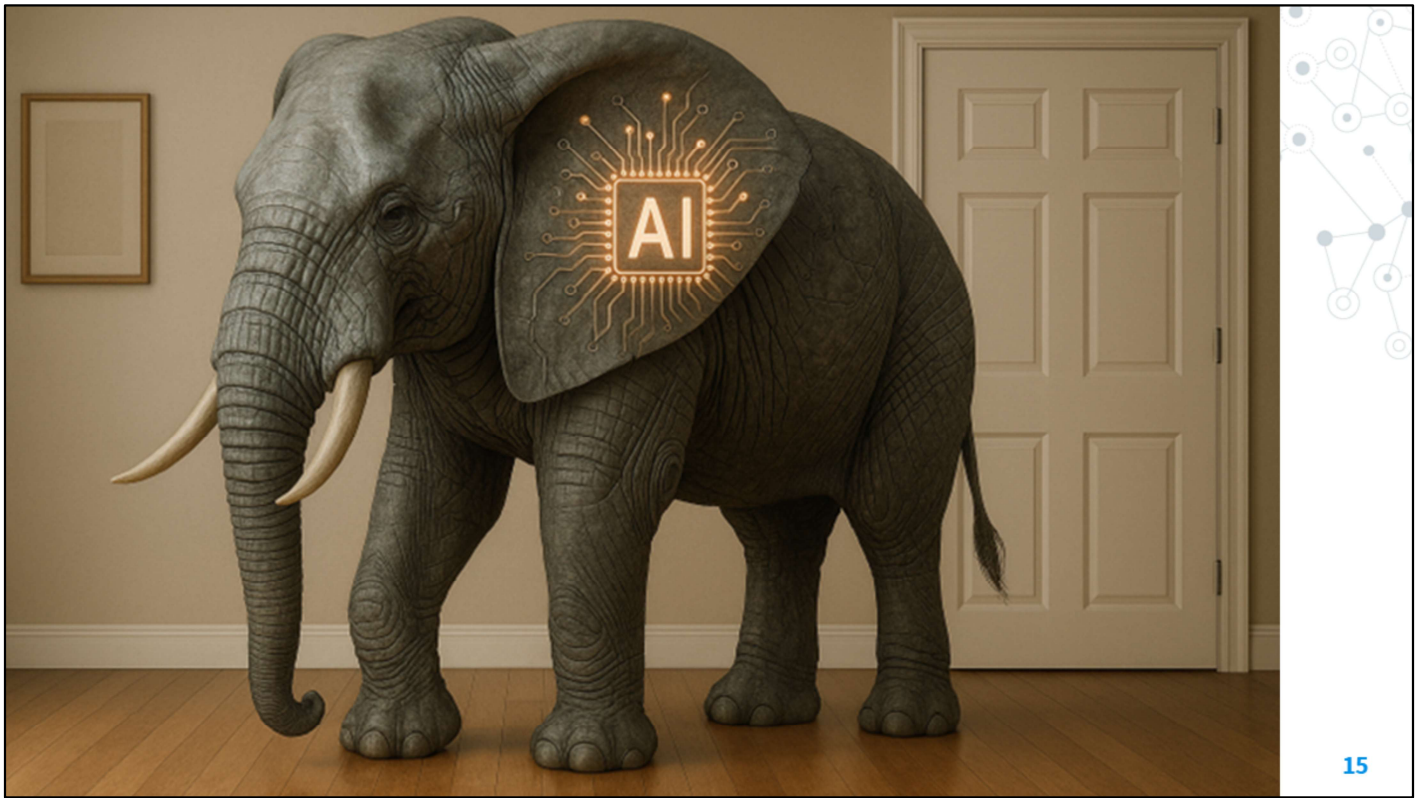
USER MESSAGE

capture this response as a markdown file.

- Scripture use survey
- Messy, in Swahili
- Summaries

Keyboard Studio (Still Unreleased)

- ◎ Creating quality Keyman keyboards requires you to be:
 1. a programmer
 2. a linguist
 3. a Windows user
 4. anal retentive
 5. a web designer
 6. and an English speaker
- ◎ Keyboard Studio seeks to knock most of those requirements out by adapting existing keyboards and giving good context-sensitive defaults.



- I want to address the AI-lephant in the room. Some are fearful of AI, some are eager to take the easy route and sit back. I can understand when someone has been doing something one way for decades and they're reluctant to change before they retire, and they shouldn't have to change their ways, but I hope they'll be curious.
- Every generation fears being replaced — by immigrants, by outsourcing, now by AI. But for missionaries, replacement is the *goal*. Sustainability means training the people who take our place. Bible translation already made that move — from missionaries with local helpers to local translators with expat consultants — and AI is just the next handoff. If someone else now do what I used to do, with or without AI, I'm free to go do the next thing.
- I use AI every day to expand my capabilities, but I choose to stay involved in the process and verify the critical results.

Nearly Universal Change

- ◎ I asked Claude what jobs had been least affected by AI:
 - It suggested skilled trades, personal trainers, tattoo artists, and hairdressers.
 - Skilled trades use AI for estimation and modeling.
 - Personal trainers use AI to generate customized plans.
 - Tattoo artists are increasingly “installing” AI-created images.
 - So, maybe hairdressers are the final frontier for AI

Or not...



r/BetterOffline • 1y ago
lizgross144



My hair stylist is peeved about how AI is impacting her business

Today my hair stylist (small town, mid-tier salon) told me that every single day, at least one person brings in a picture of a haircut or style they want that is absolutely impossible to achieve, because the photo is AI-generated. Apparently, when you google specific types of cuts and styles, nearly all of the images that show up are AI. I checked today and I see lots of dead eyes, unnatural smiles, or gravity-defying hair.



Foundations in AI

I am a nerd, by the way.





*AI can do good work if you can
show it what good work looks like.*

Me, at least once weekly

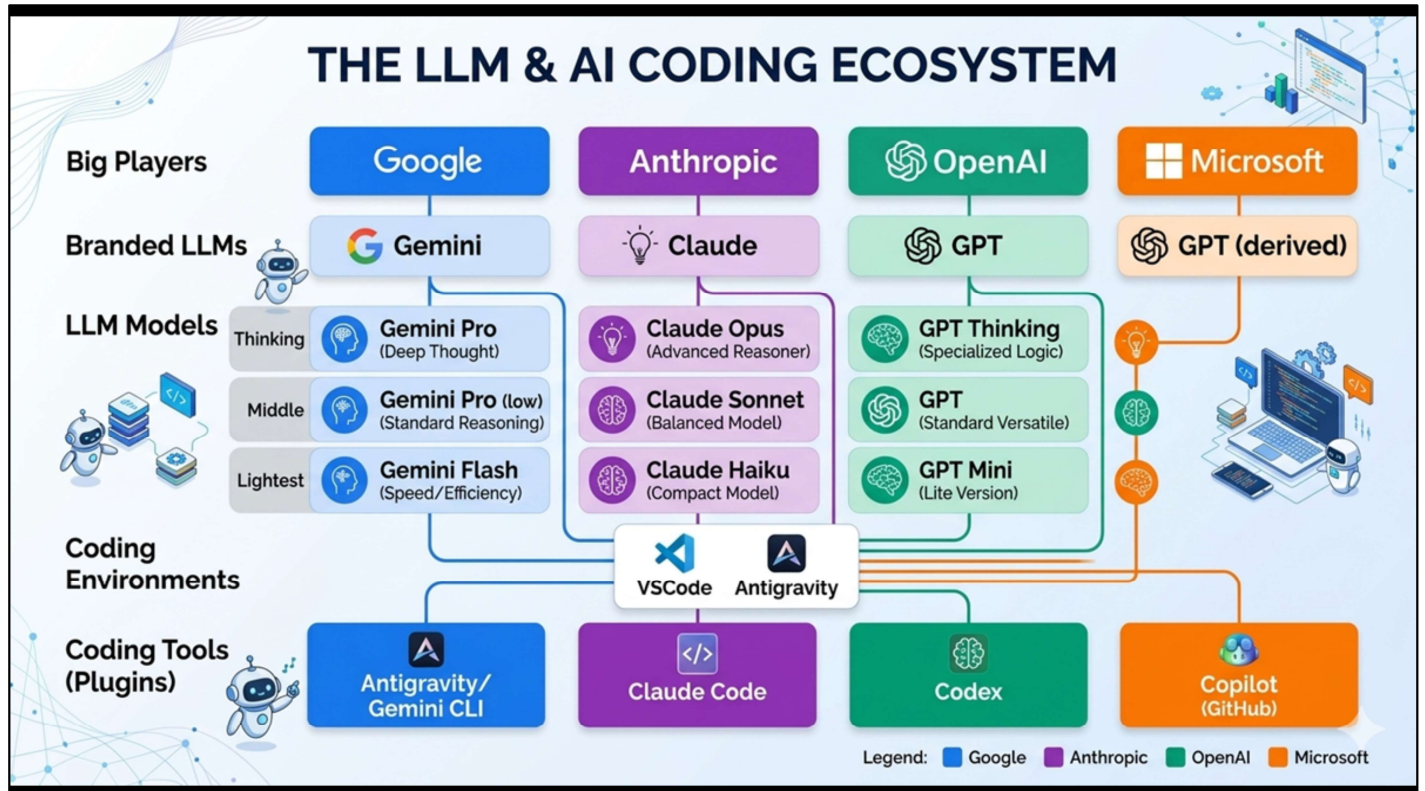
A Disclaimer

In AI, the only constant is change. I'll stick to what has held true for a little while.



- Claude, Gemini and ChatGPT are general LLMs (Large Language Models) trained on “the Internet”, predicting the next “reasonable” word from that training data, and quite a bit more.
- Artificial intelligence and machine learning cover MANY old and new tools or techniques, but gradually LLMs have become synonymous with “AI” to most non-specialists.

THE LLM & AI CODING ECOSYSTEM



- Most people who think AI is rubbish asked an LLM about their own specialty, then judged it by the first wishy-washy answer.
- This goes even more poorly if the data needed to construct the response was never made available digitally. It can “guess” but not manufacture data.

- Most AI “lies” (technically hallucinations) can be traced back to incomplete, false, contradictory, or vague training data.
- A “false” response in our software support agents can almost always be traced to a gap in the training literature and fixed by clarifying that literature.
- Directing those hallucinations is the AI equivalent to creativity and has given rise to AI art tools.

The Uni-cow (produced by Gemini)



- An LLM presents text as through a viewfinder.
- Zoom in for one concept in fine detail, zoom out for the blurry whole domain, never both in the same answer.
- So you decide where to point it.
- Ask a silly/vague question, get a silly/vague answer.
- Drill down and you will discover the detail together.

What can AI do?

◎ My answer:

- AI can do many things that you can do now nearly as well as you without a lot of guidance.
 - ◎ Translate a document.
- AI can do some general tasks better than you.
 - ◎ Write poetry. Summarize large documents.
- If it can't already do it, AI can do slightly more than you can do now, if you're willing to teach it.

Solving a difficult problem through programming.

LLM Strengths and Weaknesses (Text)

Strength

- ⦿ Find some spelling mistakes in this text. (Spot checks)
- ⦿ Estimate the number of words in this text.
- ⦿ Summarize this text.
- ⦿ Find another way to express this idea.

Weakness

- ⦿ Find ALL spelling mistakes in this text. (Thoroughness)
- ⦿ Count the number of words in this text.
- ⦿ Quote a text verbatim.
- ⦿ Keep *this* the same but change *that*.

Many people use LLMs like a search engine or Wikipedia.

- ⊙ What is the capital of Cameroon?
- ⊙ What is ergativity, anyway?
- ⊙ Who wrote “Les Misérables”?

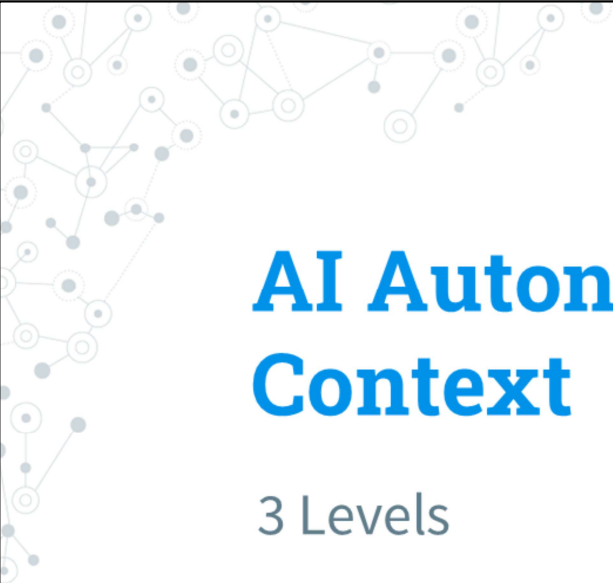
Three questions will get three “competent generalist” answers.

But a follow-up question is where LLMs will outshine search engines.

- ⊙ How has the movement of the Capital of Cameroon from Douala to Yaoundé by the Germans affected the country?
- ⊙ Ok, but how is ergativity represented in Austronesian languages?
- ⊙ Is Valjean's internal transformation a critique of institutional religion or an argument for personal spiritual regeneration outside formal church structures?

30

- AI wrote that last one...I'm not that kind of nerd.



AI Autonomy and Context

3 Levels

- Most people's first interactions with AI have been through Web interfaces. Copying and pasting the text that you want to work on into a web page and getting a response back to questions.
- This would be ChatGPT, Gemini or Claude on the Web.

- The LLM generally responds from general knowledge, because it knows VERY little about your context.
- A “brain dump” about the situation, context, and constraints as well as what you bring to the table before asking the question will nearly always garner a better response than a “cold” question.

- Later, we started giving the AI access to folders of documents and code.
- The LLM can now search and supplement its knowledge to create a fix that is adapted to your context, and you don't have to remember to “paste” it in.
- Extra Connectors/Plugins point the LLM to knowledge where it lives (email, Slack, Google Drive, external archives).
- Suddenly, the LLM can sort through the noise, find what's relevant, and solve real problems.
- This would be Claude Cowork or Code, or Gemini Antigravity.

- Make a plan of what this feature must generally do.
- Interactively determine what it must and must not do.
- Make a plan for implementing this feature from start to finish.
- Divide that plan into tasks.
- Write tests that won't pass until the feature is done right.
- The AI implements the plan step-by-step until all tests go “green”, asking clarification questions along the way.

An Aside: Autonomous Agents

- ◎ You could give an always-on LLM goals and open access to information it needs, allow it to make new connections, to act on your behalf, to manage and prune its own memory and subgoals, and improve its skills.
- ◎ This is OpenClaw and Hermes (Not recommended by SIL)
- ◎ These are still pretty scary as they can find “creative” access methods that may not be kosher and can easily be tricked into sharing your secrets.

- Today's LLMs read billions of documents.
- Nearly all in a handful of world languages.
- Once a language is decoded, new concepts slot in easily.
- That is why they feel fluent in English, French or Spanish.
- The fluency comes from volume, not understanding.



*AI can do good work if you can
show it what good work looks like.*

Me, again

- More is often written about the language than in it.
- Early orthographies are fluid, so the text is inconsistent.
- Training data is scattered, unpublished or community-held.
- The largest document is often the New Testament.
- SIL AI Engineering works on exactly this gap.

- Know 9 words out of 10 and you can infer the tenth.
- Know none and there are no signposts at all.
- AI learns the same way: valid words, valid patterns, linked to concepts it already holds.
- So Scripture Forge needs many well-aligned books before drafts are usable, often a New Testament.

- It needs to see each word dozens of times, in different grammatical situations.
- It learns by example how creative it is allowed to be.
- You want creativity inside the grammar, not invented words or ungrammatical sentences.
- “Unicow” is not a real animal, but compounding rules and related concepts made the request comprehensible in English.

SIL AI Tools <https://ai.sil.org/>

◎ **Alpha2**

- Alpha2 is a translation service aimed at non-scripture text and audio going out into a wide range of languages.

◎ **AERO: AI Enhancements Responsive to Orality**

- AERO is a set of tools for tool developers and users that meets the needs of oral language communities. (Text to Speech and Transcription)

SIL AI Tools <https://ai.sil.org/>

◎ **AQuA: Augmented Quality Assessment**

- AQuA is a quality assurance copilot tool which assists translation teams in reviewing local language drafts by bringing attention to potential issues and helping them gain insights into the quality of the translation.

◎ **Faithbridge**

- Faithbridge is a chatbot that allows users to interact with a set of ground-truth Bible translation resources.

SIL AI Tools <https://ai.sil.org/>

◎ **FLExTrans**

- FLExTrans is a specialized machine translation tool designed to assist Bible translators in translating between related languages.

◎ **Scripture Forge**

- Scripture Forge is an online app for team-based Bible translation, featuring AI help and tools for community review.

- Non-determinism is a liability when you want one repeatable answer, but it is an asset when you want options.
- Ask for twenty ideas rather than one; the weak ones clarify what you actually want from the good ones.
- The model never gets bored, tired, or defensive, so you can push back as many times as you like.
- Judgement stays with you: it generates the options, you pick the winner.

- LLMs are non-deterministic:
- Solve problem A for me... You get a reasonable response.
- Ask it in a new chat to solve problem A. You get a different reasonable response.
- Ask it in a new chat to solve problem A. You get a different reasonable response.
- Sometimes a single long chat will unintentionally change things that worked earlier.

- “Let's analyze this problem and build a deterministic process or tool for determining a solution.”
- You get code or text that you can test, refine and tweak as far as you desire.
- In my experience: the first draft will get you 80% of the way there. The next cycle will add 10%. The next cycle adds 5%. With diminishing returns.
- But you can use the created script dozens or hundreds of times. No AI needed for further runs.

- Write an R script that determines the confidence of this conclusion.
- Maintain an audit trail of changes and transformations to the raw data.
- Build a survey wizard tool that helps users adapt an existing Keyman keyboard for their language.

- It's nice to have an answer, but better to have proof.
- This process of "showing your work" makes processes measurable, improvable, and auditable.
- Responses drawn from the "black box" of AI aren't reproducible research in the same way you can't cite Wikipedia.

Global Diaspora Services

- ◎ Your team is making connections across worldviews.
- ◎ You are innovators and making connections.
- ◎ You're already enabling activities that might not happen in our lifetimes without "outside insider" help.
- ◎ A little AI assistance might make some tasks you're doing more efficient.
- ◎ We haven't seen the limits of your imagination, and I can't wait to see what diaspora leaders can dream up.

50

- I wish that I had taken the time to get to know you and your challenges to make targeted suggestions, but we haven't

Diaspora

- ◎ The insider (emic) view sees what an outsider never fully will, yet a broader education sometimes fills real gaps and enables unique connections. That's how diaspora can stand in the gap and bridge both worlds: rooted on the inside, widened from the outside. AI can work the same way, offering one more avenue for closing gaps, not a wall that closes us out.

**Do you work with
documents, data,
media, presentations
or communication?**



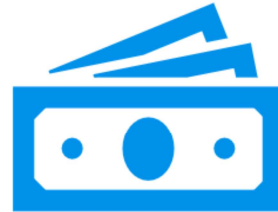
Read Wikipedia in your language									
1,000,000+ articles									
العربية	English	فارسی	Italiano	Nederlands	Polski	Sinugboanong	Svenska	Tiếng Việt	中文
Deutsch	Español	Français	مصرى	日本語	Português	Binisaya	Українська	Winaray	Русский
100,000+ articles									
Afrikaans	বাংলা	Ελληνικά	हिन्दी	Latviešu	Bahasa	O'zbekcha /	Slovenščina	Татарча / Tatarça	粵語
Shqip	Беларуская	Esperanto	Hrvatski	Lietuvių	Minangkabau	Ўзбекча	Српски / Srpski	తెలుగు	
Asturianu	Català	Euskara	Bahasa Indonesia	Magyar	မြန်မာဘာသာ	Қазақша /	Srpskohrvatski /	ភាសាខ្មែរ	
Azərbaycanca	Čeština	Galego	עברית	Македонски	Norsk (bokmål •	Qazaqşa / قازاقشا	Српскохрватски	Тоҷикӣ	
Български	Cymraeg	한국어	ქართული	Malagasy	nynorsk)	Română	Suomi	تۆرکجه	
閩南語 / Bân-lâm-gú	Dansk	Hausa	Ladin	मराठी	Нохчийн	Simple English	Kiswahili	Türkçe	
	Eesti	Հայերեն	Latina	Bahasa Melayu		Slovenčina	தமிழ்	اردو	
10,000+ articles									
Bahsa Acèh	Bikol Central	Føroyskt	Ilokano	Kurdi / کوردی	Ming-dêng-ngû /	پښتو	ChiShona	chiTumbuka	नेपाली
Alemannisch	বিষ্ণুপ্রিয়া মণিপুরী	Frysk	Interlingua	کوردی ناوەندی	閩東語	Piemontèis	Sicilianu	Basa Ugi	नेपाली टिप्पणी
አማርኛ	Boarisch	Fulfulde	Interlingue	Кыргызча	Монгол	Plattdüütsch	සිංහල	Vèneto	
Aragonés	Bosanski	Gaeilge	Ирон	Кырык мары	Napulitano	Qaraqalpaqsha	سنڌي	Volapük	
Արեւմտահայերէն	Brezhoneg	Gàidhlig	Íslenska	Lëtzebuergesch	नेपाल भाषा	Qırımtatarca	Ślůnski	Walon	
Bahasa Hulontalo	Чӕсӕхмал	گيلکی	Jawa	Uigure	Nordfriisk	Runa Simi	Soomaaliga	文言	
Basa Bali	Dagbanli	ᲫᲗᲚᲘᲗ	ಕನ್ನಡ	Limburgs	Occitan	Русиньскый	Basa Sunda	吴语	
Bahasa Banjar	الدارجة	Hak-kâ-ngî / 客家	Kapampangan	Lombard	Олык марий	संस्कृतम्	Tacelhit	שׂיב	
Basa Banyumasan	Diné Bizaad	ꨀꨂꨛꨍꨔ	ᱫᱟᱲᱟ	मेथिली	ଓଡ଼ିଆ	ᱥᱟᱱᱛᱟᱲ	Tagalog	Yorùbá	
Башҡортса	Emigliàn-	Hornjoserbsce	کٲشُر	മലയാളം	অসমীয়া	سرائیکی	தேவநாட்டி	Zazaki	
Беларуская	Rumagnòl	Ido	Kotava	მარგალული	ਪੰਜਾਬੀ	Саха Тыла	ተፎርታፍላት ተፎላፍቲ	Žemaitėška	
(тарашкевіца)	Fiji Hindi	Igbo	Kreyòl Ayisyen	مازونی	پنجابی (شاہ مکھی)	Scots	tolışi	isiZulu	

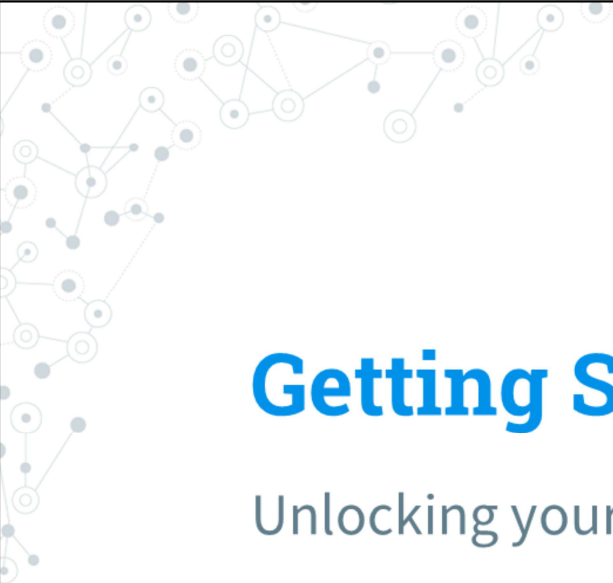
- How many of you work, at least partly, in some of these languages?

**Do you have
ideas you want
to implement?**



**Is it worth \$10 a
month to
accomplish
more?**





Getting Started

Unlocking your AI assistants

Let's Talk Security

There are many service providers on the internet, but if you pay nothing or SIL has no agreement, **you and your data are the product.**

For sensitive SIL and partner Data, we must stick to APPROVED DATA PROCESSORS and corporate accounts.

Google, Claude, Microsoft
(OpenAI still being Negotiated)



Two easy ways in today

Gemini (Google)

Available through SIL, but limited in what it will do (even with a paid account). Useful for quick drafting and short summaries.

Subproducts:

- ⦿ Gemini Notebook (NotebookLM)
- ⦿ Google Vids

Claude (Anthropic)

Available to staff at a deep discount (\$10/\$50). Better for long documents, real work, and survey write-ups.

Subproducts:

- ⦿ Claude Chat
- ⦿ **Claude Cowork**
- ⦿ **Claude Code**

<https://licensemanager.sil.org/catalog>

- Unlock Gemini on your SIL Account
- Subscribe to a Claude Account



My Subscriptions

Catalog

AI

Backup

CRM

Creative

Email Accounts

IT

Meetings

Storage

Task/Project Management

Action Center

License Catalog

AI



Claude - Premium

Anthropic · AI
\$50 per seat/month

Request



Claude - Standard

Anthropic · AI
\$10 per seat/month

Request



Gemini

Google · AI
\$0 one-time

Request



Google Delegated Account

Google · Email Accounts
\$180 per year

Request



Partners

Commercial Vendors (Approved Data Processors)

- ◎ Anthropic (Claude)
- ◎ Google (Gemini)
- ◎ Microsoft (Copilot)

Who SIL is working with on AI



Missional AI

SIL sponsors the summit with ETEN's Innovation Lab, Biblica and American Bible Society.



XRI Global

A partnership on AI language technology built for less-resourced languages.



Faith Comes By Hearing

An oral-first drafting workflow, piloted with SIL.



ETEN

A resource platform supporting Bible Translation.



Greek Room

AI-assisted checking that flags spelling, consistency and wording issues.



Aquifer

An open library of trusted Bible resources for other tools.





Thanks!

Questions?

Matthew Lee, Language Technology
Consultant

Matthew_Lee@sil.org



