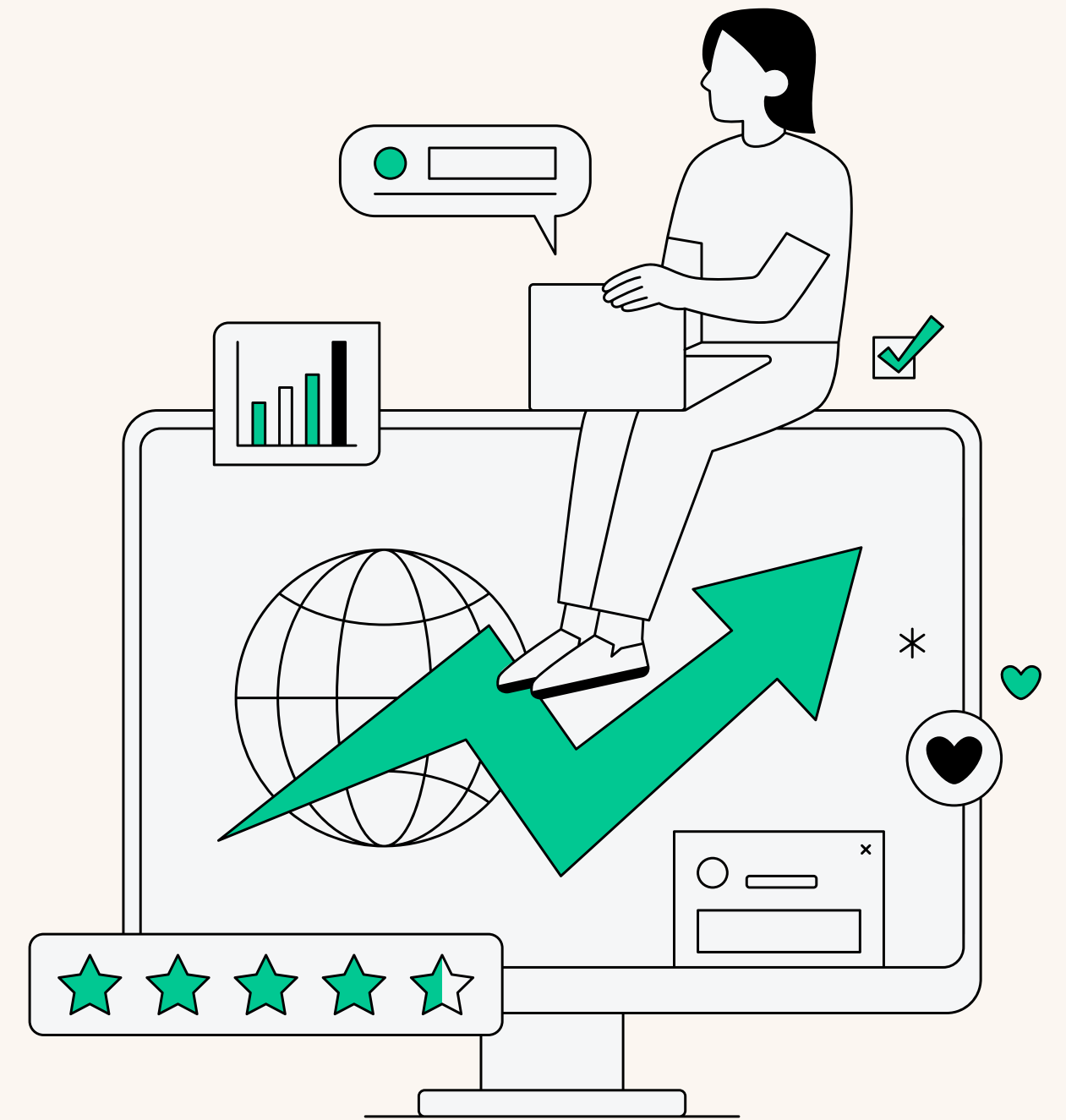


Presented by tubone24

ClickHouseと AWS CloudWatch RUMを統合して AIエージェントの 分析を加速させる



自己紹介

Yu Otsubo (tubone24)

ゆるふわエンジニア@KAG

妻と犬の3人暮らし

生成AI関連の案件をやっていて、AIエージェント(A-BOSS)開発をしています。
AIエージェントの開発ではLLMOps的なことも取り組んでいます。

最近執筆でいそがしかったため、いぬ、かまってあげられなかった。
ごめんよーーーー！



A-BOSSの紹介
YouTube




いぬのインスタ




Xアカウント





突然ですがみなさま
ClickHouseを“何に”
使っていますか...？

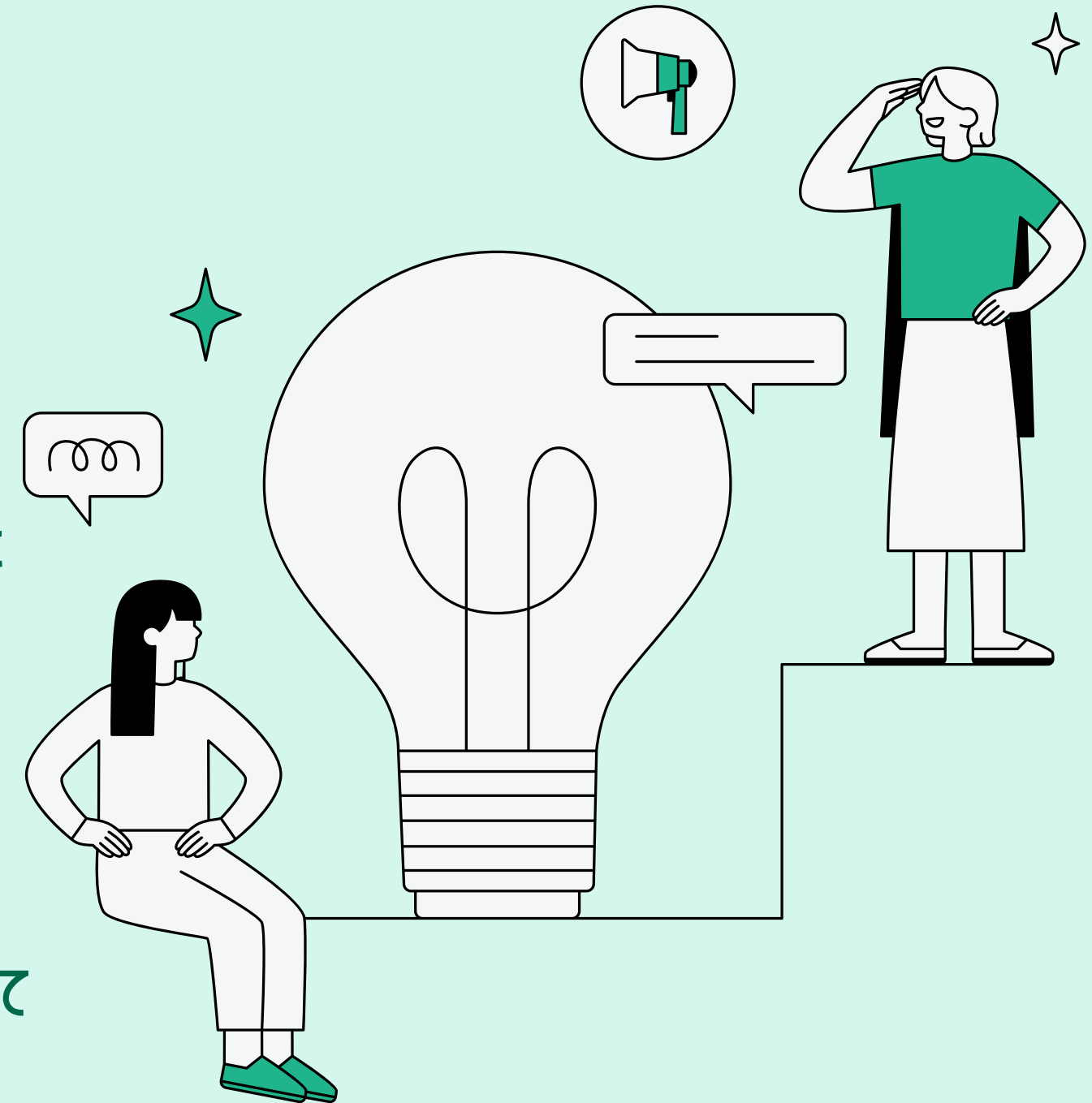


ClickHouse

ClickHouseは、OLAP（Online Analytical Processing）に特化した、高速でスケーラブルな列指向データベース管理システムです。

オープンソースソフトウェアとしても、クラウドサービスとしても利用できます。

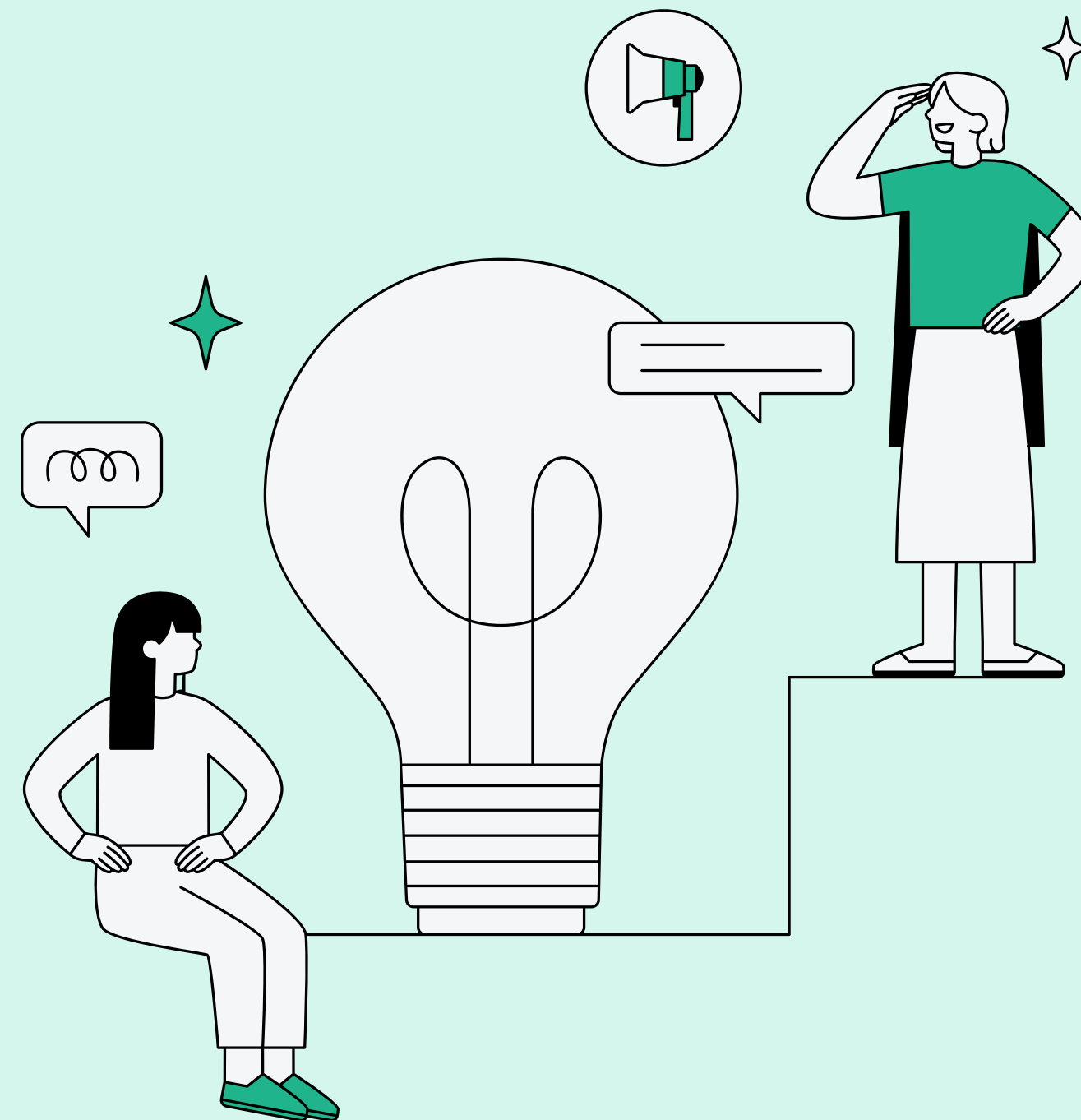
MySQLなどと比較して、特に大量データの分析処理において高い性能を発揮します。



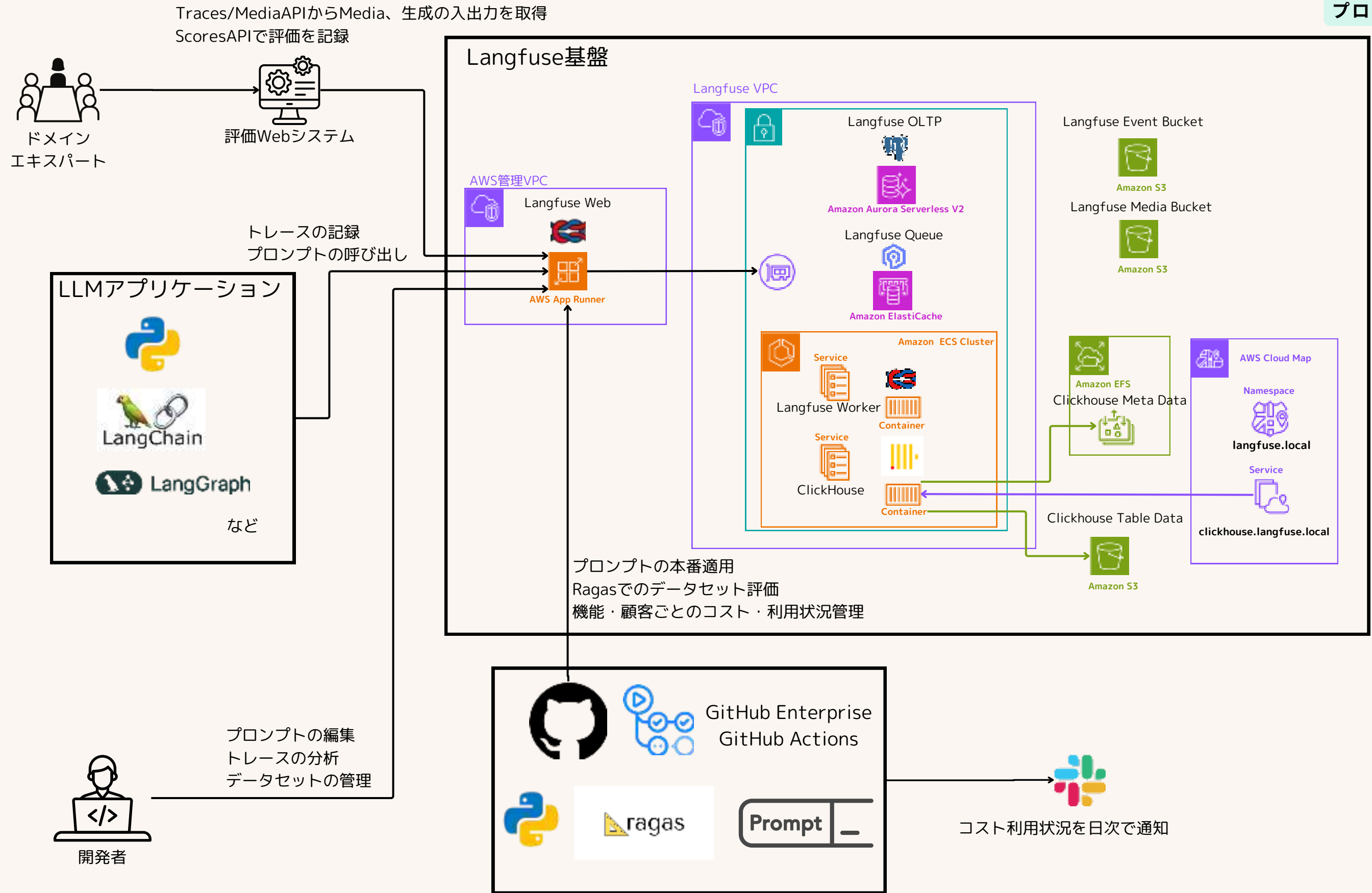
ぼくら生成AI peopleの

ClickHouse

LangfuseのOLAPです！！！！

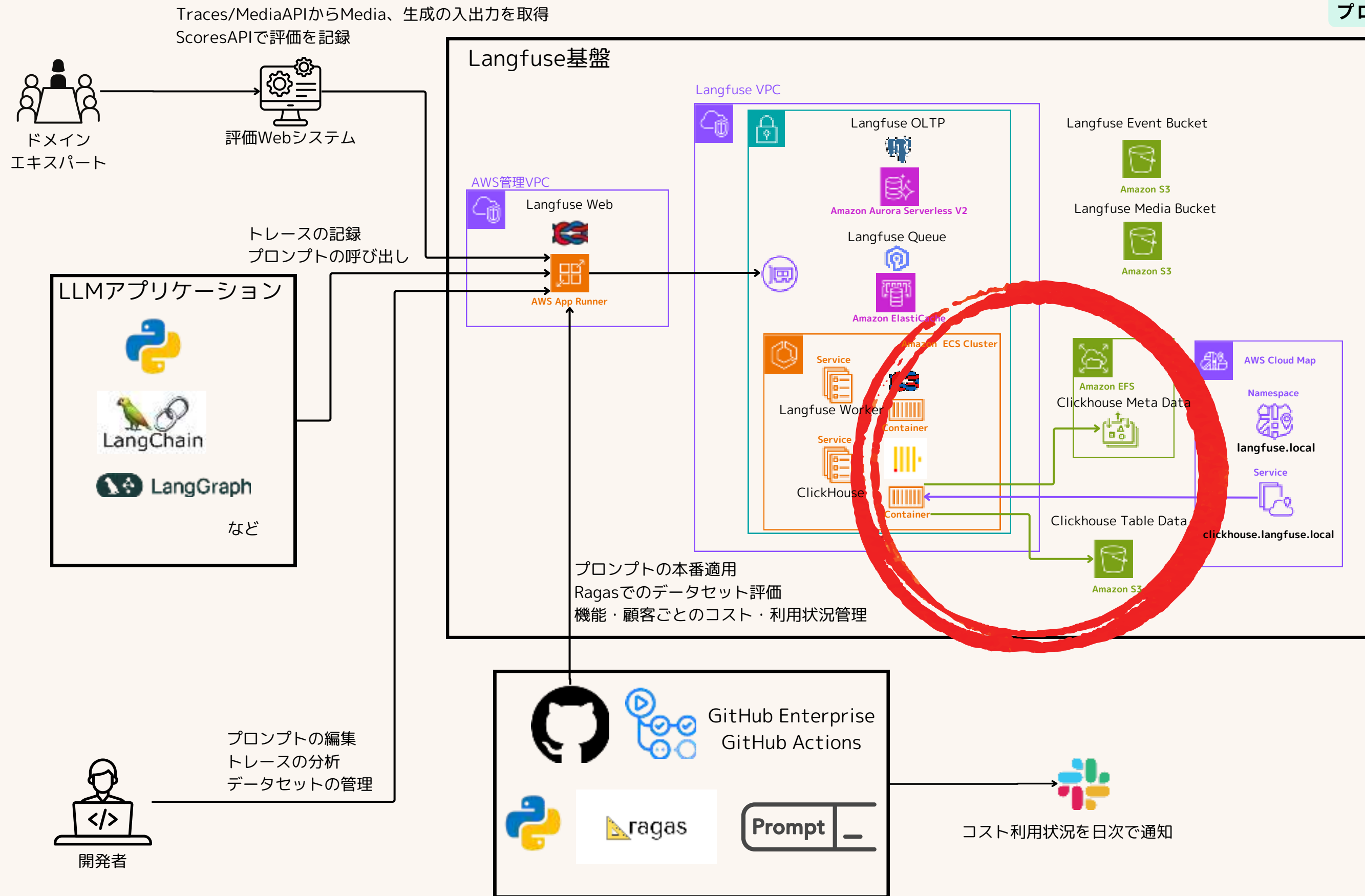


このプレゼンテーションにはプロモーションが含まれます



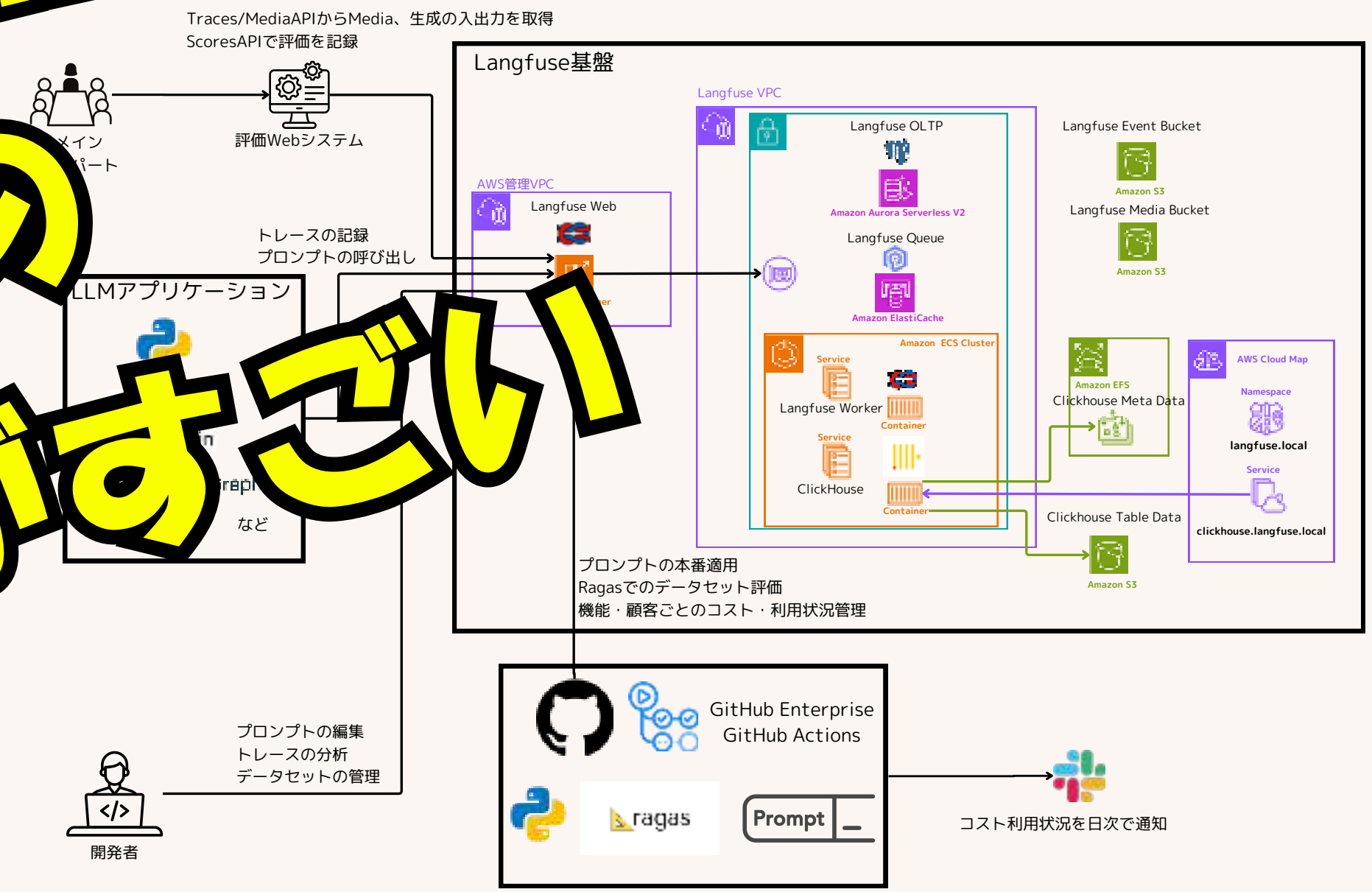
<https://findy-tools.io/products/langfuse/397/396>

このプレゼンテーションには
プロモーションが含まれます



OSS版もあるので 自VPC上に構築できます

ClickHouse
の
ここがすごい



詳しくはこちらの資料で



Langfuse v3 使ってますか? - ページ 1/31

Langfuse Night #1で発表したLTです。

pages.dev



Langfuse V3を安くしたい - ページ 1/46

Langfuse Night #2 - 初心者から熟練者まで - でおはなししたものです。

pages.dev



今日はLangfuseの
裏側のClickHouseを
使い倒そう！という内容です

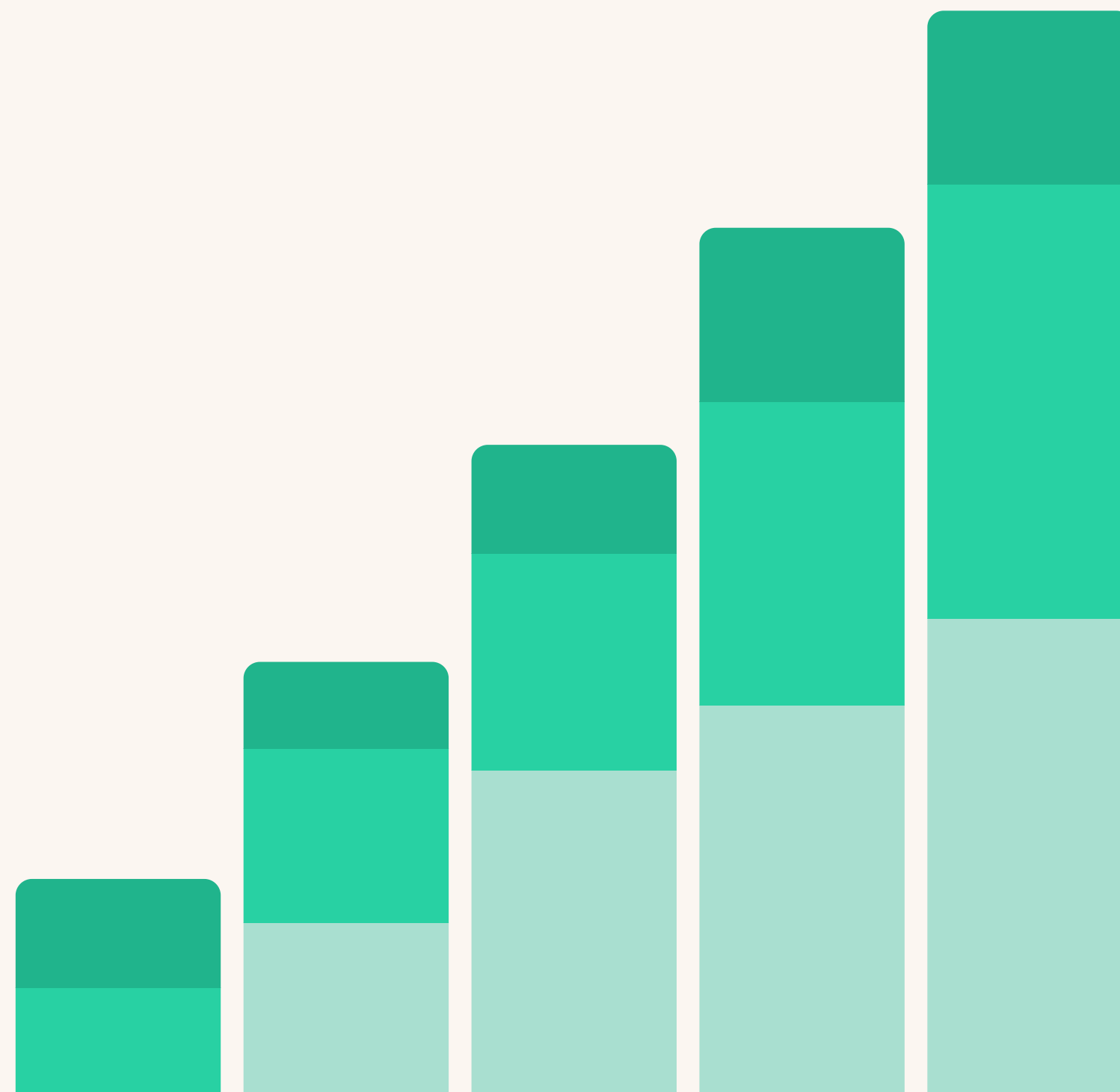


AIプロダクト 分析への誘い

あなたの作っているAIプロダクトはちゃんと使われていますか？

ちゃんと(安全に)“計測”していますか？

誰が“計測”していますか？

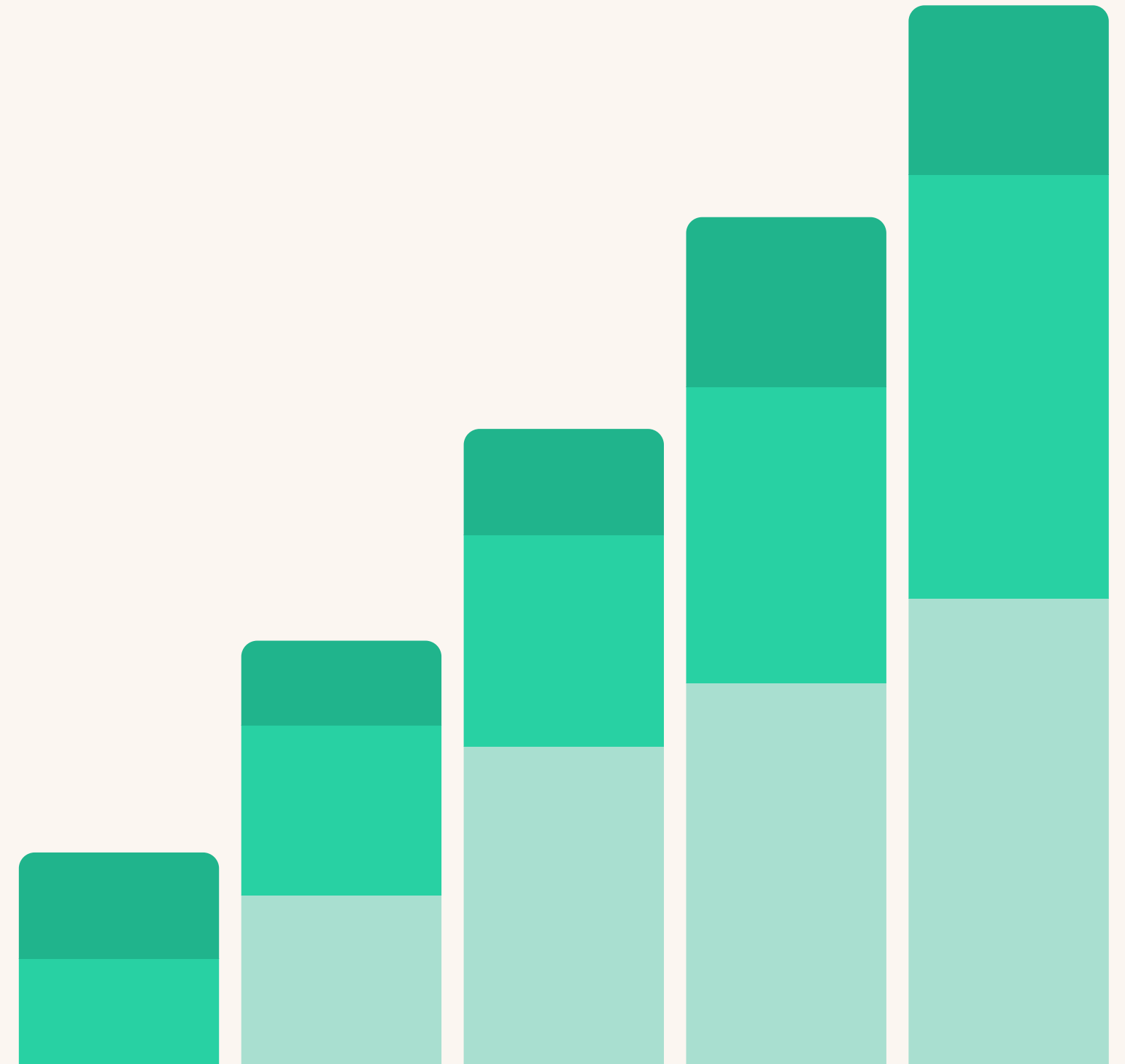


AIプロダクト 分析への誘い

あなたの作っているAIプロダクトはちゃんと使われていますか？

ちゃんと(安全に)“計測”していますか？

誰が“計測”していますか？



“誰が”計測して頂けますか？

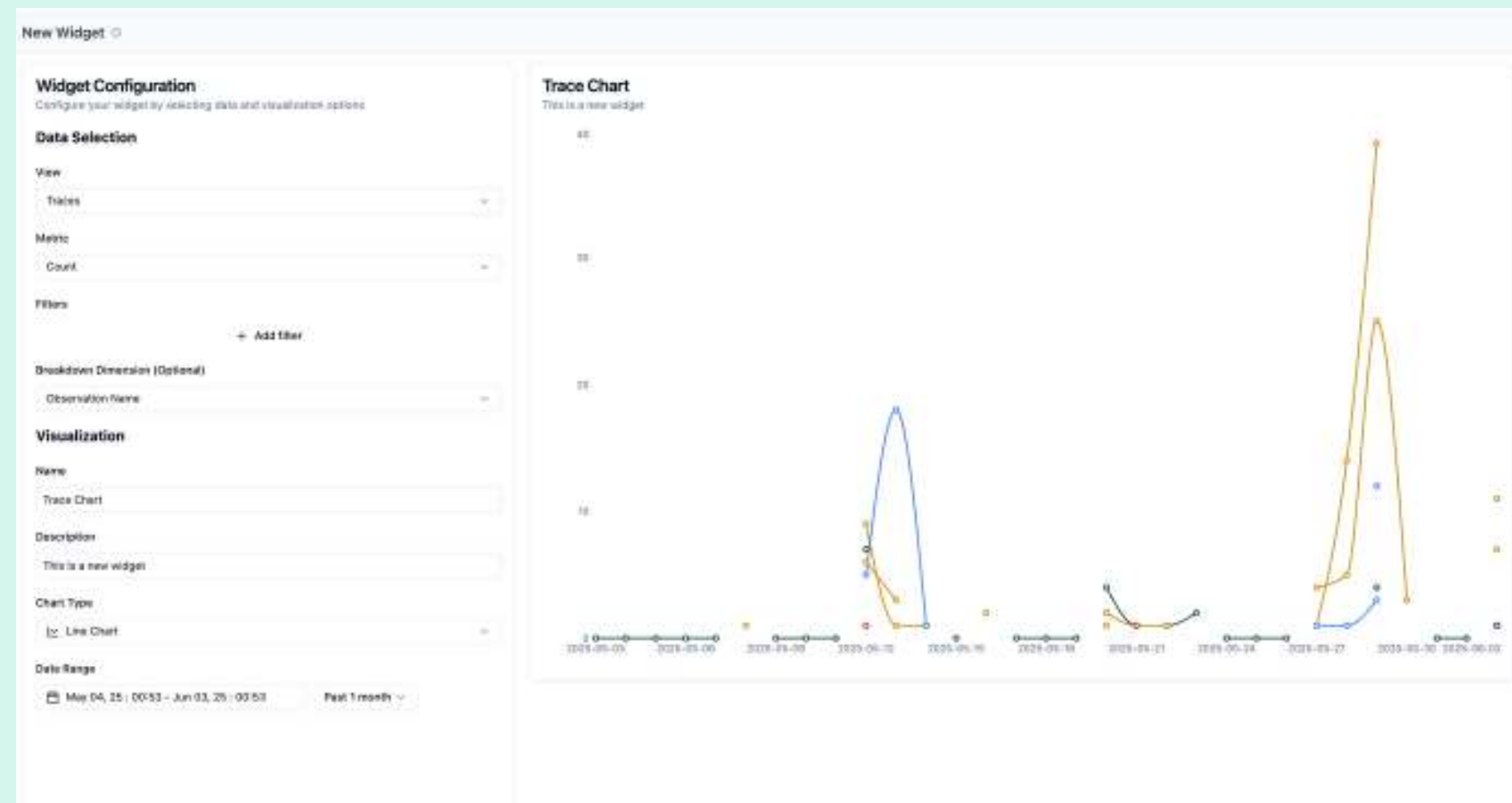
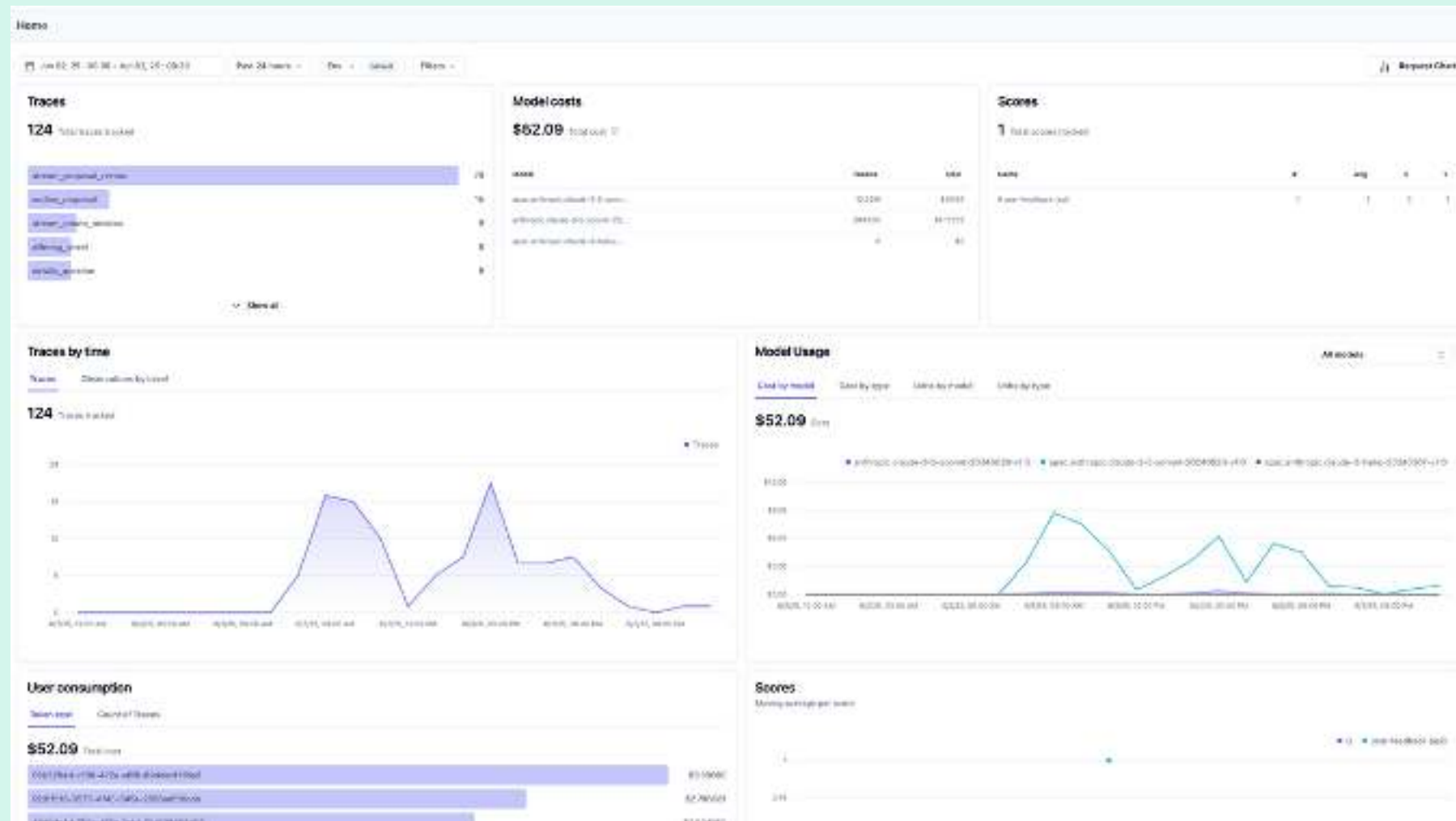
開発者？

プロダクトマネージャー？

データサイエンティスト？



Langfuseにも便利な ダッシュボードはありますが...



やっぱりSQLで分析したい

SQLのほうが
慣れているな

より高度な
集計をしたいな



プロダクトマネージャー

エンジニア





SQL叩くならClickHouseを
直接叩けばいいじゃない





というお話です



データ構造を正しく把握する

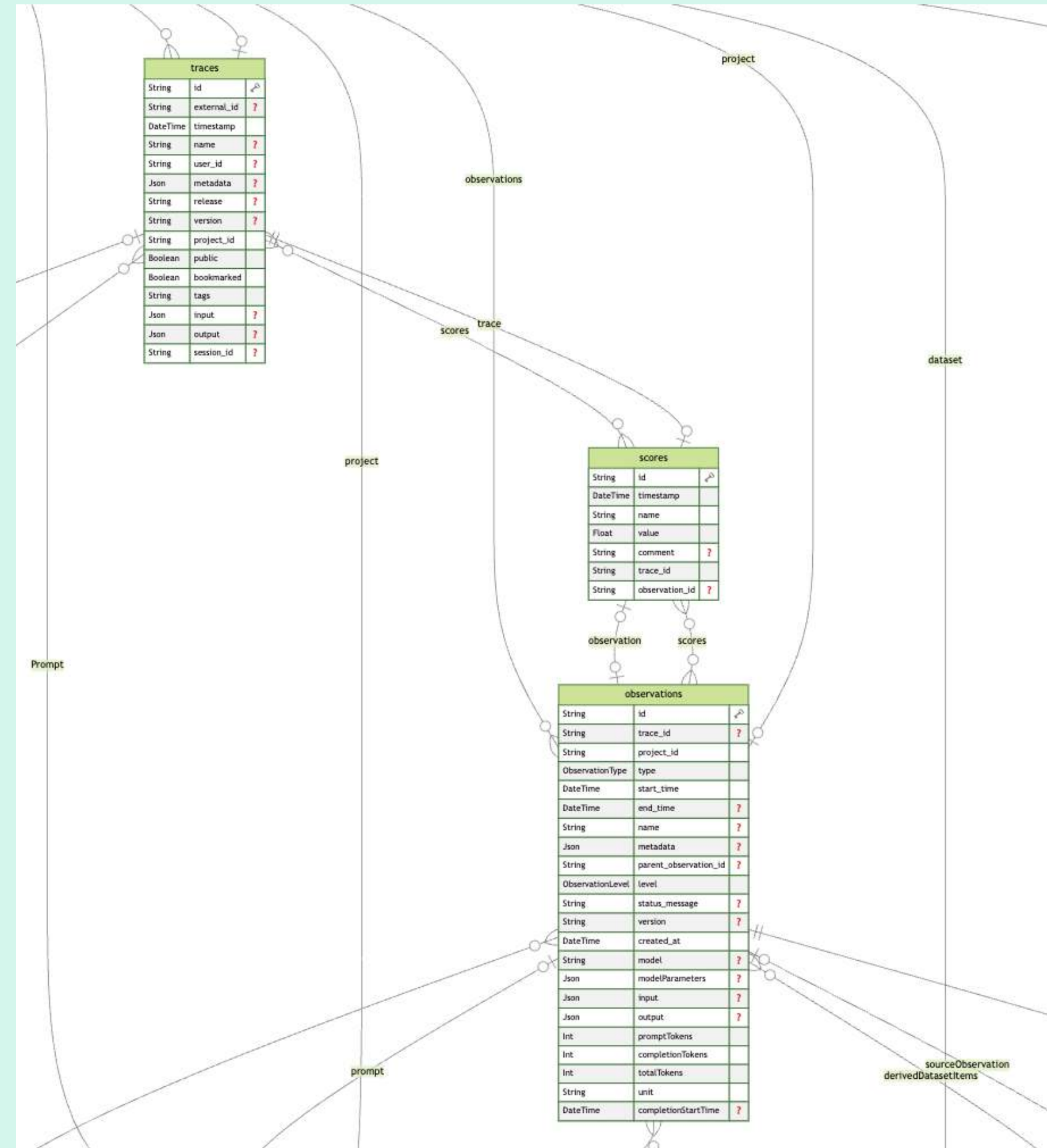
Langfuse公式に ER図あります

[langfuse / packages / shared / prisma / database.svg](#)

上記はPostgreSQLのER図なので、
ClickHouseにはこの内

- traces
- observations
- scores

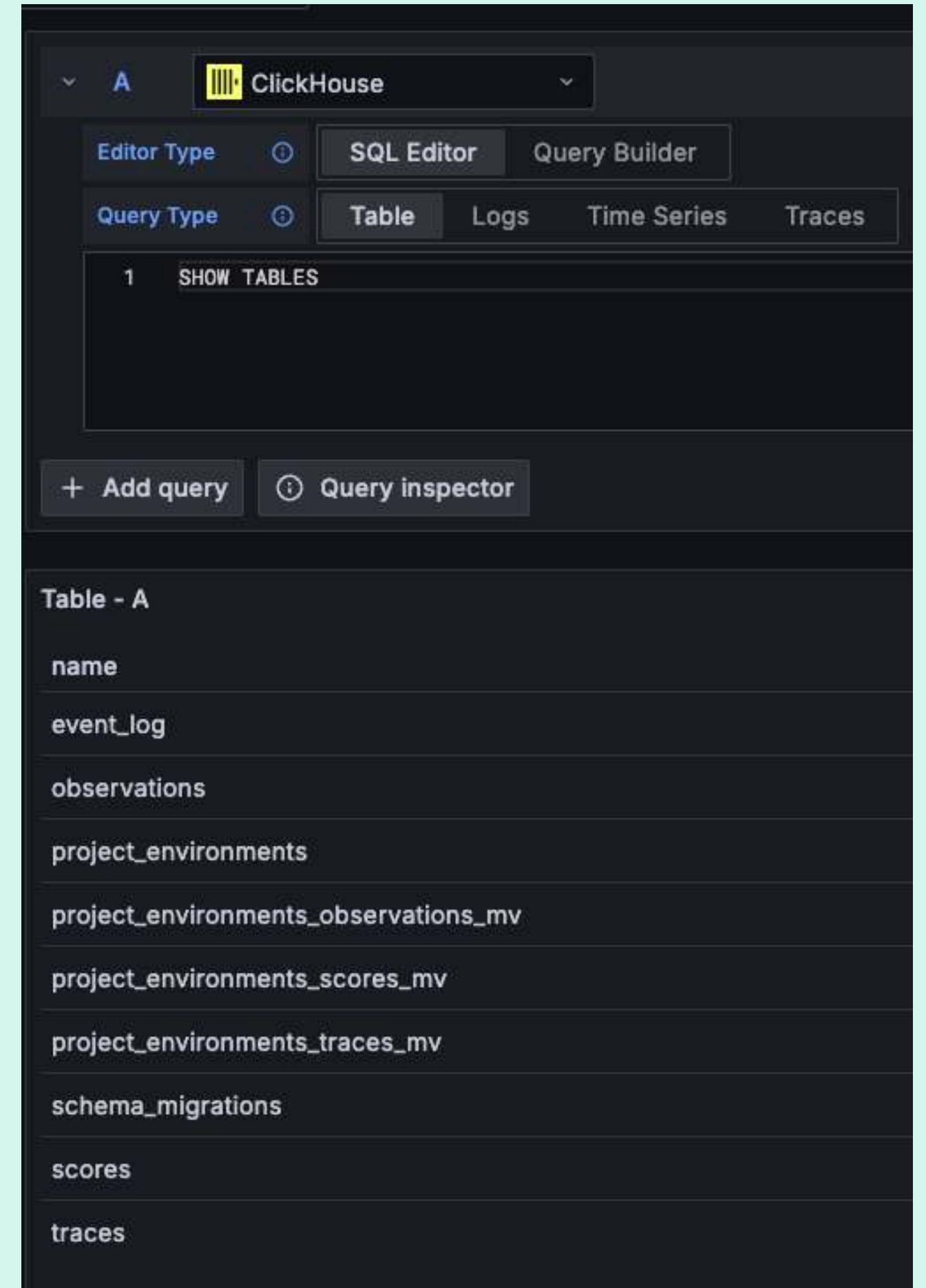
あたりを見ればOK!



Langfuse公式に ER図あります

ぜひ、SHOW TABLESしてみてください！

ぜひ、SELECT * してみてください！

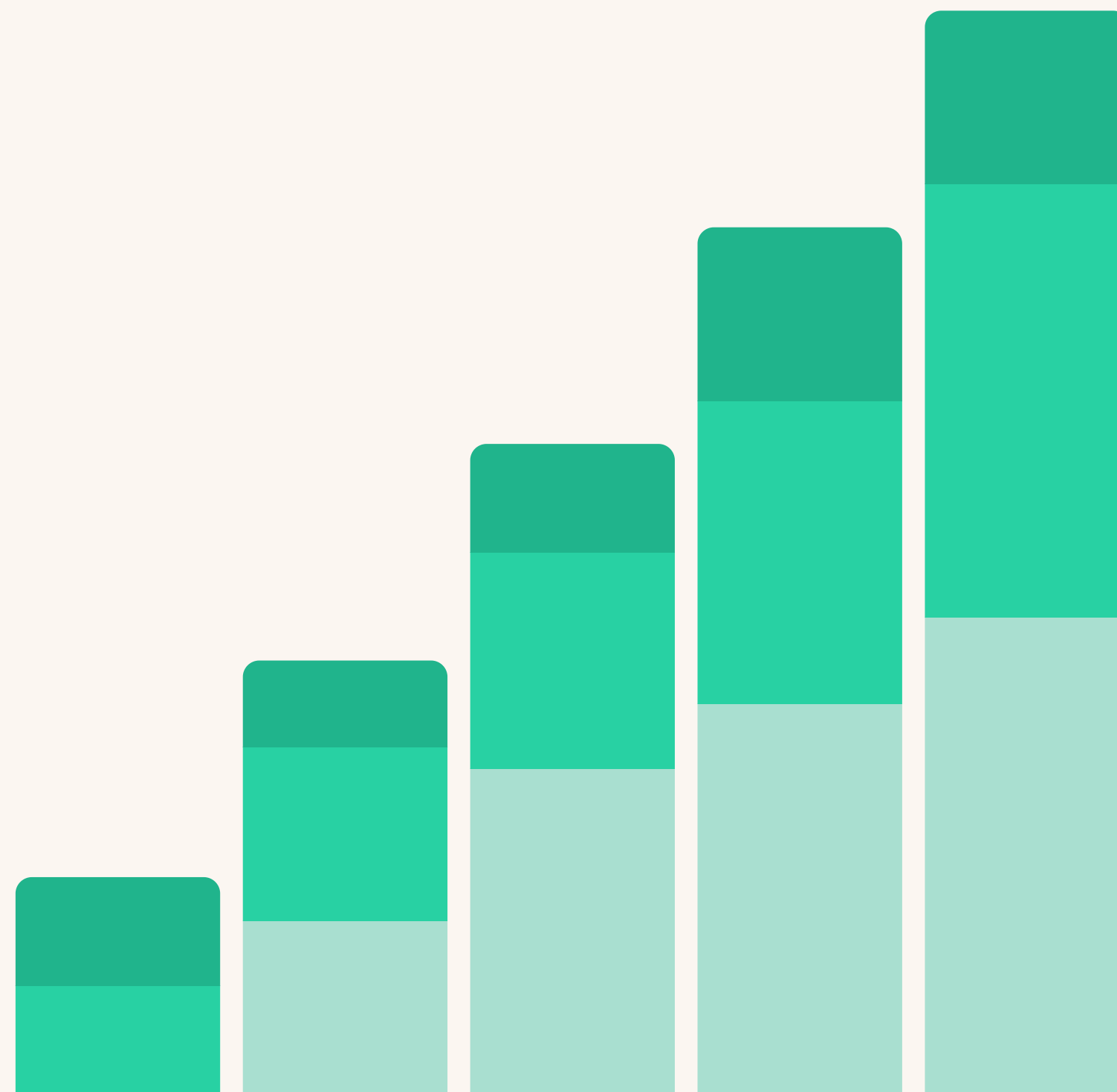



AIプロダクト 分析への誘い

あなたの作っているAIプロダクトはちゃんと使われていますか？


ちゃんと(安全に)“計測”していますか？

誰が“計測”していますか？

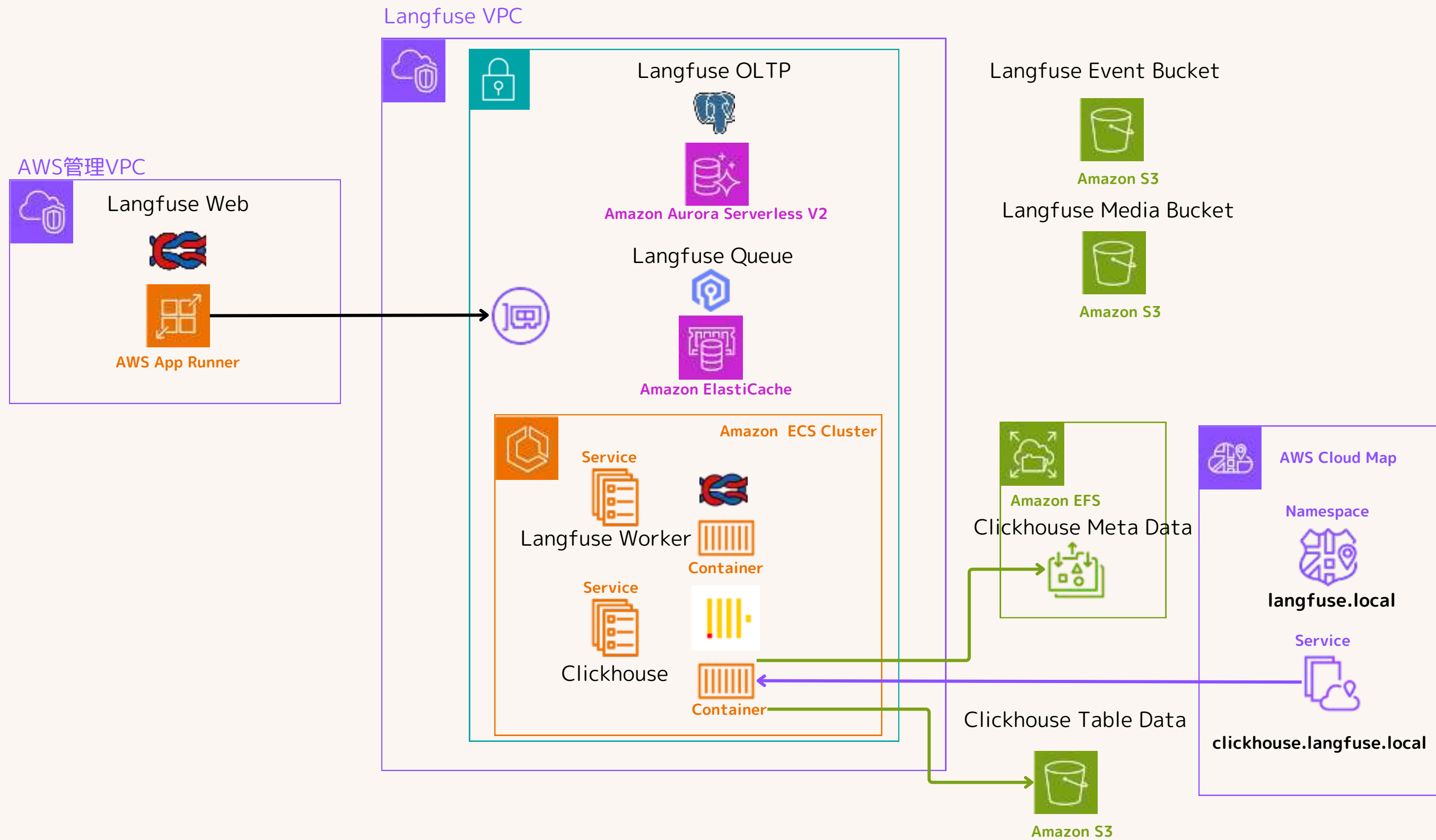




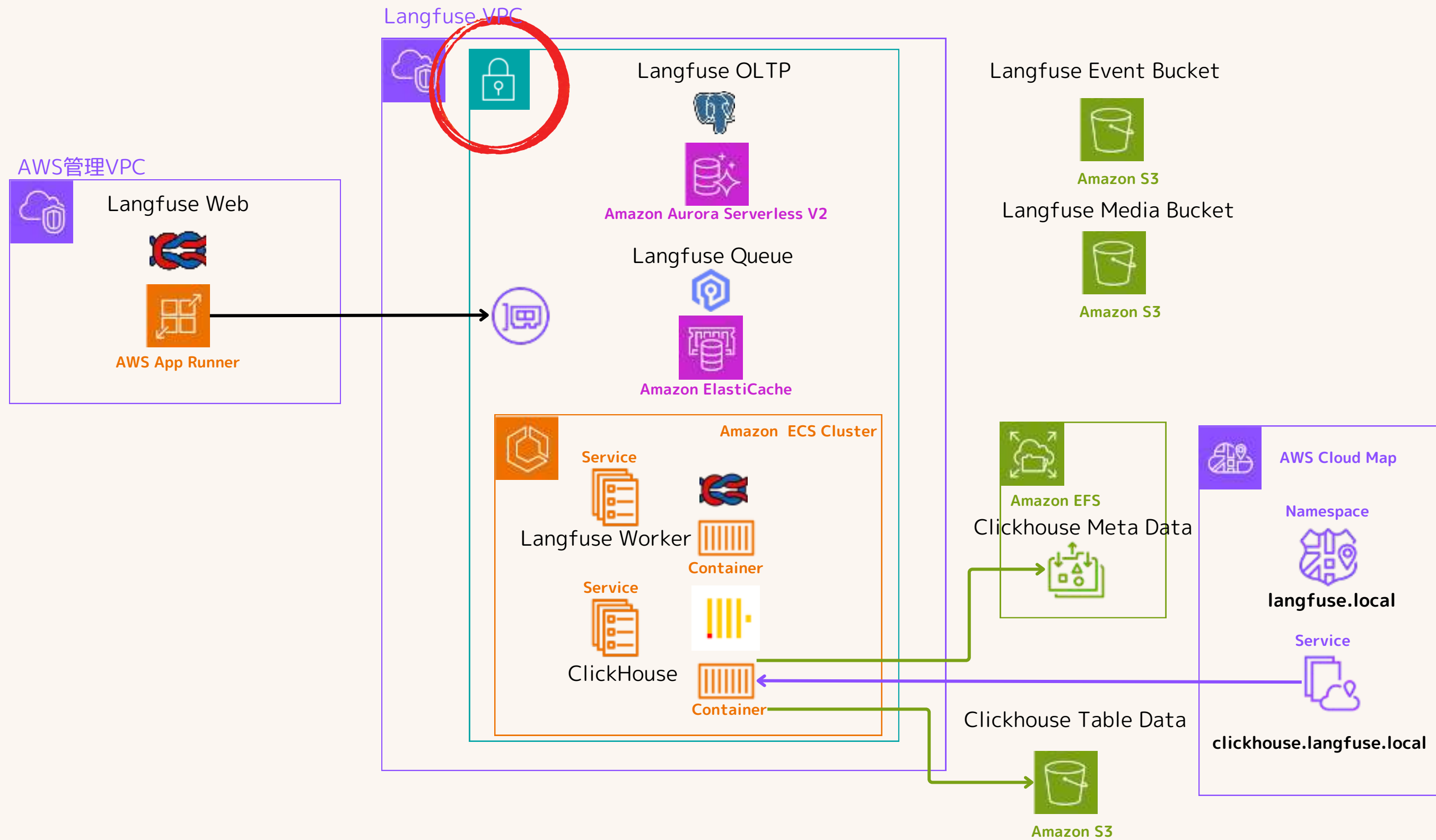
ClickHouseの クエリランナーを作る




Langfuse 基盤




Langfuse 基盤





ClickHouseに
インターネット経由で
クエリできない
(させたくない)



ClickHouseを叩ける クライアント一覧

公式でもサードパーティー製の
Visual Interfaceを紹介しています。

<https://clickhouse.com/docs/interfaces/third-party/gui>

日本語のマニュアル！

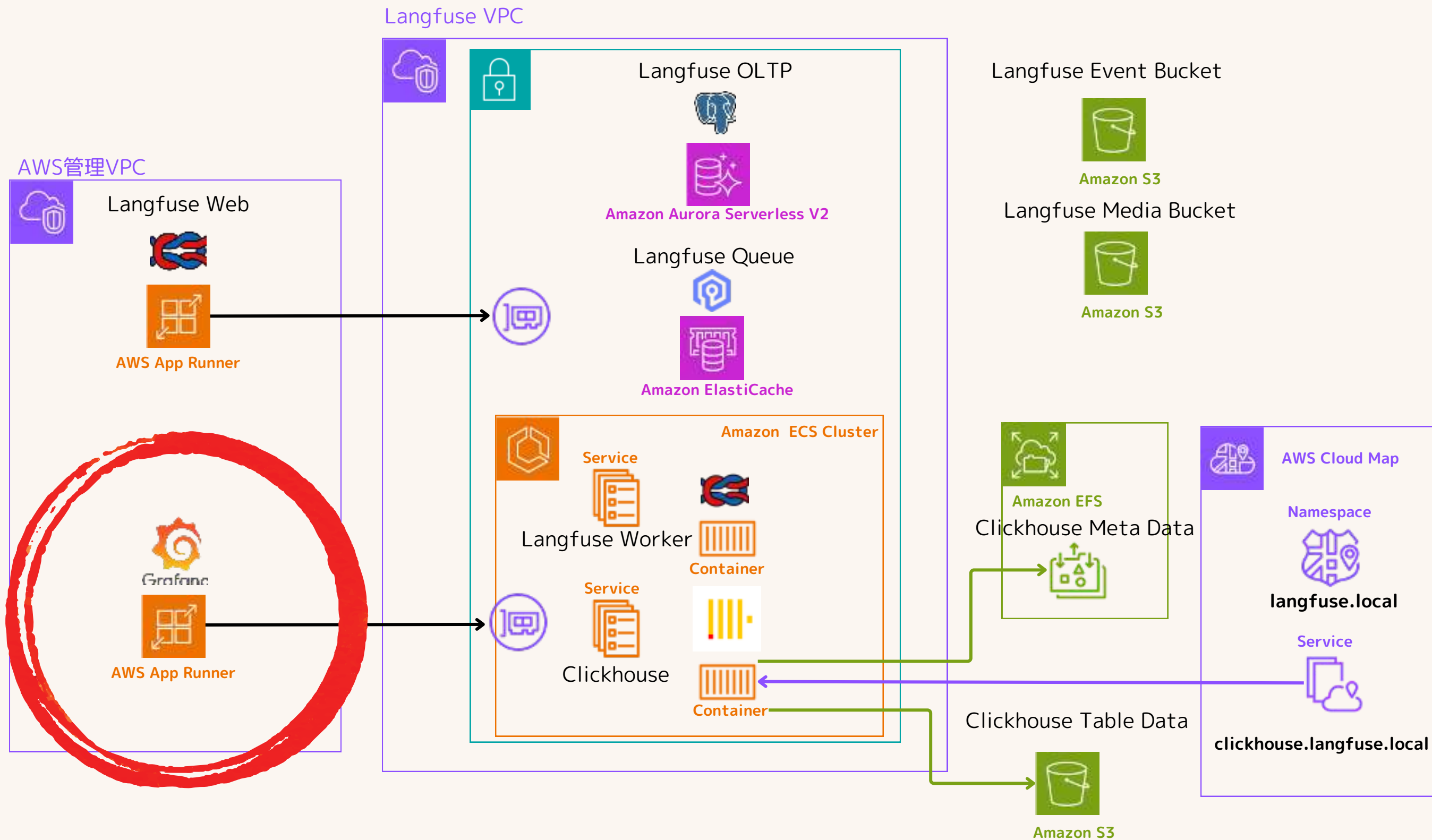
ClickHouse
の
ここがすごい

The screenshot shows the ClickHouse website's navigation menu and a document page. The navigation menu includes '製品', 'ユースケース', 'ドキュメント', 'リソース', '料金', and 'お問い合わせ'. The document page is titled '第三者開発者によるビジュアルインターフェース' (Visual Interface by Third-Party Developers). A notification box at the top of the document states: 'このドキュメントはAIの助けを借りて翻訳されています。' (This document is translated with the help of AI). Below the notification, there is a button '問題を報告する' (Report a problem). The main content area contains a list of third-party integrations: 'agx', 'データ取り込み', 'ツール', 'SQLクライアント', 'データ統合', and 'その他'. The 'agx' section is expanded, showing a description: 'agx は、ClickHouse の埋め込まれたデータベースエンジン (chdb) を使用してデータを探索およびクエリするためのモダンなインターフェースを提供する、Tauri と SvelteKit で構築されたデスクトップアプリケーションです。' (agx is a modern desktop application built with Tauri and SvelteKit that provides an interface for exploring and querying data using ClickHouse's embedded database engine (chdb)). Below the description, there are four bullet points: 'ネイティブアプリケーション実行時に ch-db を活用。' (Use ch-db when running native applications), 'ウェブインスタンスを実行しているときに Clickhouse インスタンスに接続できます。' (You can connect to Clickhouse instances when running web instances), 'モナコエディタで、親しみやすいインターフェースを提供。' (Provide a familiar interface using Monaco Editor), and '複数の進化するデータビジュアライゼーション。' (Multiple evolving data visualizations). The page also shows a sidebar with 'Wa' and 'Try for' buttons.



Grafana

Langfuse 基盤





このプレゼンテーションには
プロモーションが含まれます

でのクエリ実行

Outline ClickHouse

Query history Share

Split Add Last 1 hour Run query

Queries Table

Editor Type SQL Editor Query Builder

Query Type Table Logs Time Series Traces

```
1 SELECT * FROM traces WHERE timestamp > '2025-05-01' LIMIT 10
```

+ Add query Query inspector

Table - A

id	timestamp	name	user_id	metadata	release	versi	project_id	environment	public	bookmarked	tags	input	output	sess
97f6f6e4-306b-414a	2025-05-30 18:47:44		6f	{ "company_id":...	d9058c!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
0678cbe3-ec87-489	2025-05-01 16:15:04		f9	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
0f321057-a1ad-4a4a	2025-05-01 16:19:51		f6	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
11679139-39ff-4174-f	2025-05-01 13:58:28		fe	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
12b822b7-61b4-4ef9	2025-05-01 09:31:00		2v	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
139e206f-5750-4d9c	2025-05-01 22:05:31		ef	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
16f5aff2-1233-460e-	2025-05-01 17:35:51		bi	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
18b273a7-697d-4b0f	2025-05-01 09:00:36		ef	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
1a5d04b1-cc93-4a9c	2025-05-01 12:12:31		Bc	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[
1b9dad3e-f844-42ff-	2025-05-01 12:08:12		Bc	{ "company_id":...	36874d!	0.0.0	cm0xjtrro0009f1p5a	default	false	false	[

Grafana経由で アクセス

AWS App Runner で建てたGrafanaをVPCコネクタで
ClickHouseに接続（ClickHouseはCloudMapで名前解決）

Ch-uiなど、多くのWeb UIではClickHouse HTTPエンドポイントをフロントエンドから直接、つまりブラウザから直接アクセスできないと使えないが、
GrafanaはClickHouseからデータを取る経路だけ確保されていればHTTPエンドポイントをブラウザから直接叩けるようになる必要はない。

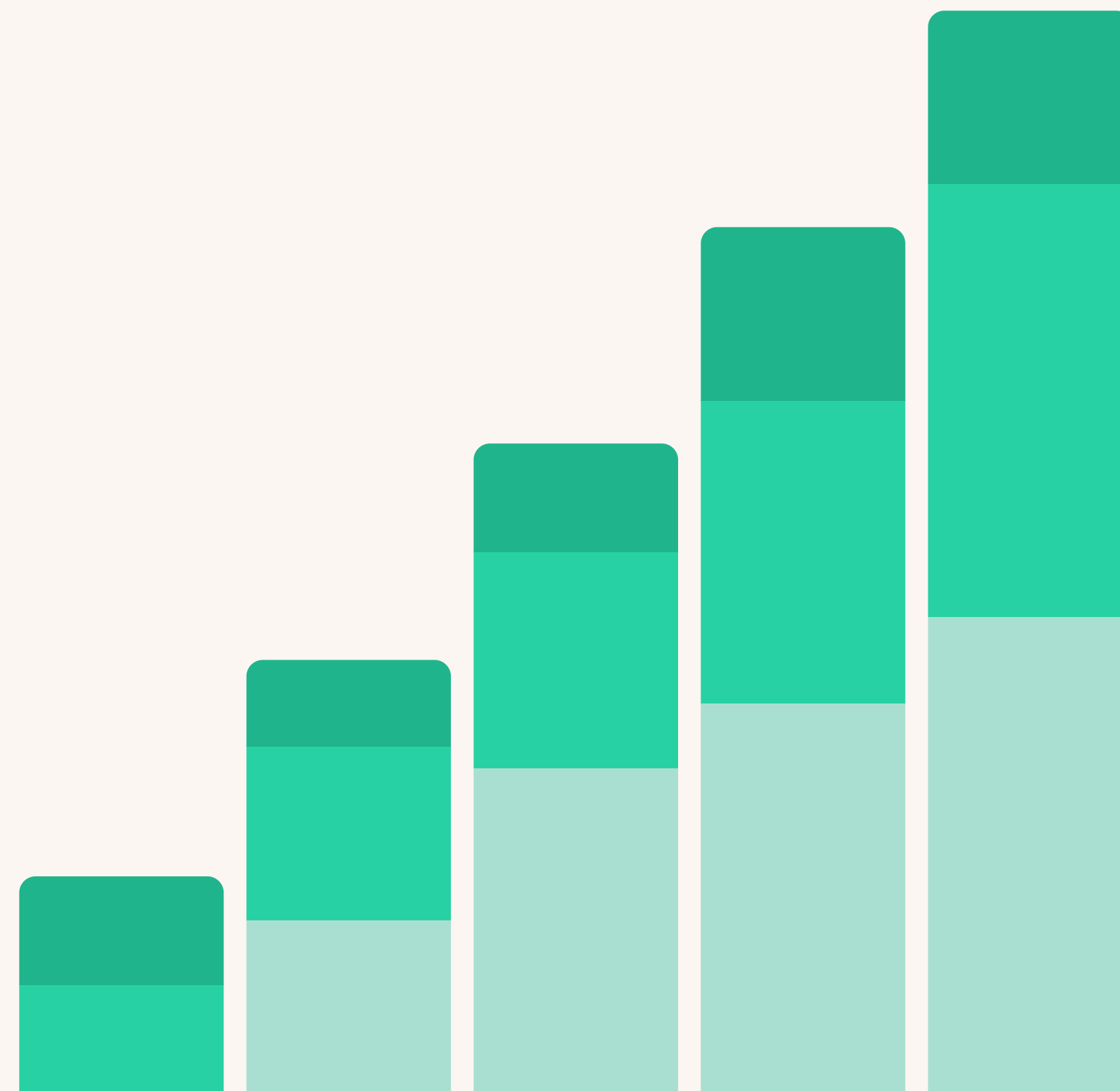


AIプロダクト 分析への誘い

あなたの作っているAIプロダクトはちゃんと使われていますか？

ちゃんと(安全に)“計測”していますか？

誰が“計測”していますか？





より使えるデータにする



AIエージェントの 効果測定

Langfuseで取っている集計だけで十分ですか...?

そもそもAIエージェントを使おうと思って離脱しているユーザーはいませんか...?
様々なデータセットとつなぎ込みをしたほうが良いのではありませんか...?

Amazon CloudWatch RUM

Web(フロント)のパフォーマンスデータ

いわゆるGoogle Analytics的なやつ

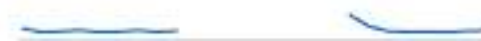
(AWSで使えるから別途GAの社内申請とかいらない！これ×リット！)

ユーザーセッションを収集し、ページの読み込み速度やエラー、Web導線を確認
このユーザーIDを“Langfuse”のユーザーIDと統合しましょう！というお話

300 +1.6k%



1kミリ秒 -34.6%



122



時刻別の平均ロード時間

時刻別のロード時間

6.2kミリ秒

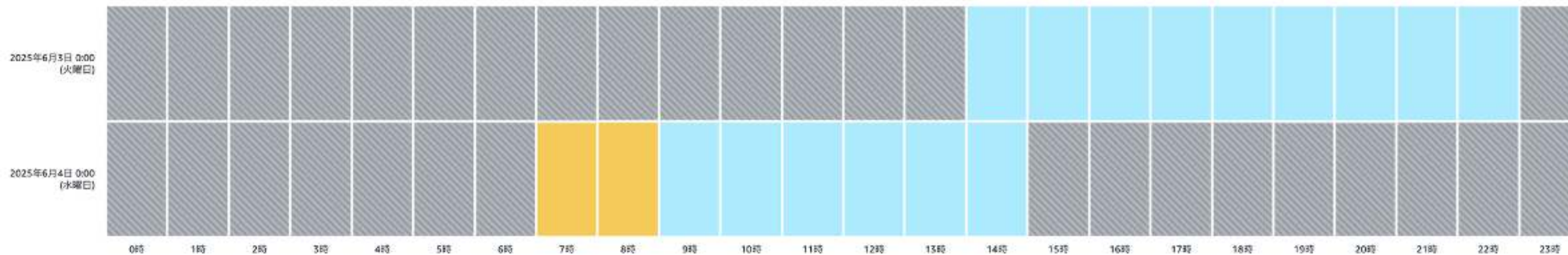
122

レイテンシーが最も長い

エラー

エラーを表示

● 利用可能なデータがありません ● ポジティブ (< 2kミリ秒) ● 許容できる (2kミリ秒 - 8kミリ秒) ● イライラする (> 8kミリ秒)



ウェブバイタル 情報

統計 P75

ウェブバイタルは、ウェブページのコンテンツのロード方法を表します。ウェブバイタルの詳細を表示

Choose a point on the graphs to see correlated events.

New! Choose a point on the graphs to see correlated events.

最大コンテンツの描画

ロードパフォーマンスを測定



累積レイアウトシフト

最初の入力遅延

インタラクティブ性を測定



Interaction to next paint

RUM統合3ステップ

01.

RUMをWeb
フロントエンドに入れる

02.

ユーザーIDを発行し、
RUMとLangfuseにそれぞれ
設定する仕組みを作る

03.

めっちゃAWSの仕組みを使っ
てClickHouseと同じように
Grafanaで可視化する

ユーザーIDを発行し、 RUMとLangfuseにそれぞれ 設定する仕組みを作る

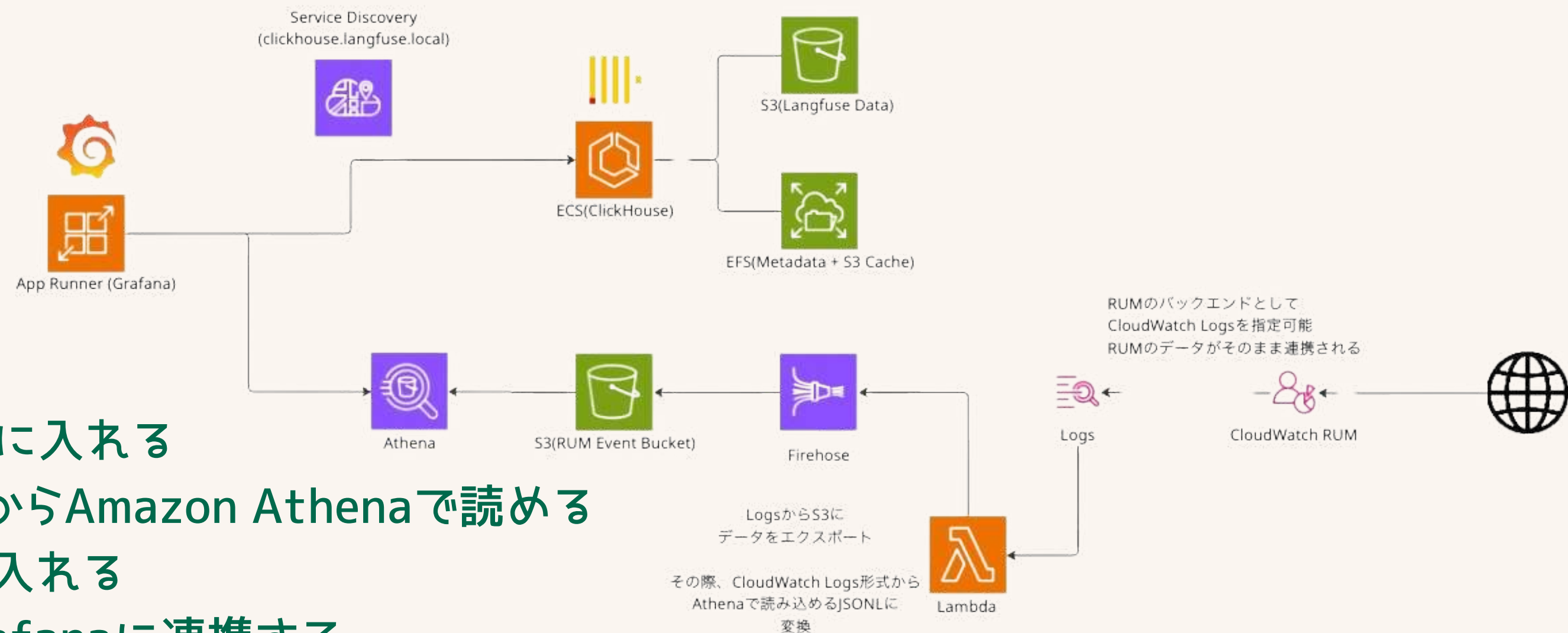
```
55  
56 useEffect( effect: () : void => -{  
57   if (!userId) {  
58     const newUserId : string = generateUserId();  
59     setUserId(newUserId);  
60   } else {  
61     // userIdが設定されたらAWS RUMを初期化 | tubone, 2025/04/18 17:40 • AWS RUMでA-BOSSのUserIDを取得で  
62     const rumInst : Orchestration | null = execute_aws_rum(userId);  
63     if (rumInst) {  
64       setRumInstance(rumInst);  
65     }  
66   }  
67   }, deps: [userId]);  
68  
69   const value : {...} = {  
70     userId,  
71     setUserId,  
72     clearUserId,  
73     rumInstance,  
74   };  
75  
76   return <UserContext.Provider value={value}>{children}</UserContext.Provider>;  
77 };
```

ユーザーIDの取り回しの
仕組みの中でRUMにも
ユーザーIDをつけちゃう

やろうと思えば2日くらいでできそう

ここではユーザーIDのCONTEXTがあるので
そこで初めてRUMのインスタンス初期化をしていますが一例の
実装です。

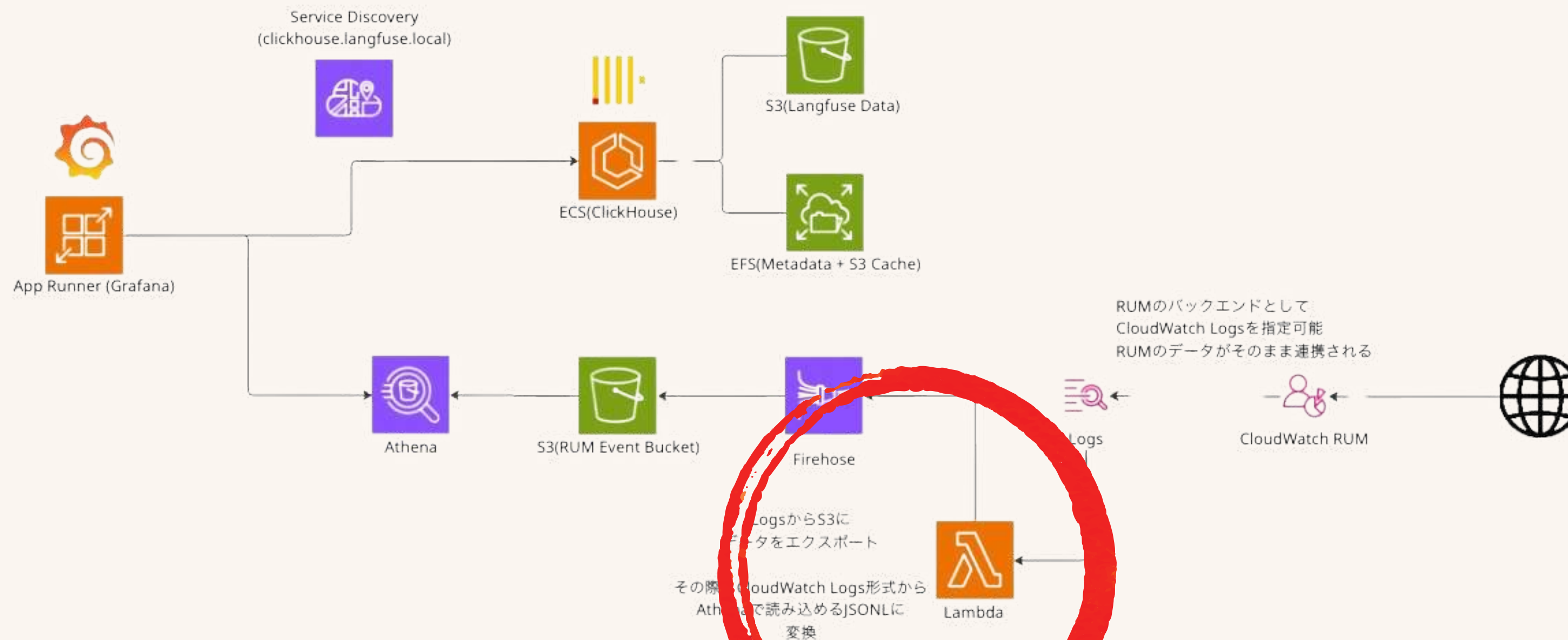
めっちゃAWSの仕組みを使ってClickHouse と同じようにGrafanaで可視化する



- RUMをCloudWatch Logsに入れる
- Kinesis FirehoseでLogsからAmazon Athenaで読めるJSONL形式に変換しS3に入れる
- Athenaを先程つくったGrafanaに連携する

やるうと思えば1週間でできそう

ちなみにここでClickHouseに入れることもできそう！



このプレゼンテーションにはプロモーションが含まれます

ClickHouse SQL Editor

```
1 SELECT * FROM traces WHERE timestamp > '2025-05-01'
```

Run Query

Athena Query Editor

Region: default (us-east-1) | Data source: default (AwsDataCatalog) | Database: default (a_boss) | Table: Choose | Column: Choose | Format data frames as: Table

```
1 SELECT
2   metadata.pageid AS page_id,
3   metadata.ABossId AS user_id,
4   COUNT(*) AS view_count
5 FROM
6   a_boss.rum_events
7 WHERE
8   event_type in ('com.amazon.rum.page_view_event', 'com.amazon.rum.qos_event')
9   AND event_timestamp >= CAST(to_unixtime(date '2025-05-01') * 1000 AS bigint) -- 集計開始日
10  AND event_timestamp <= CAST(to_unixtime(timestamp '2025-05-20 23:59:59') * 1000 AS bigint) -- 集計終了日
11  AND dt >= '2025-04-30'
12  AND dt <= '2025-06-01'
13 GROUP BY
```

Table Showing only 20 columns

version	project_id	environment	public	bookmarked	tags	input	output	session_id	created_at	updated_at	event_ts	is_deleted	page_id
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-05-30 18:47:46	2025-05-30 18:50:36	2025-05-30 18:50:36		/sp
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 21:03:59	2025-06-01 21:04:29	2025-06-01 21:04:29		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 15:54:02	2025-06-01 15:54:26	2025-06-01 15:54:27		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 21:04:13	2025-06-01 21:08:06	2025-06-01 21:08:06		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 15:53:00	2025-06-01 15:53:33	2025-06-01 15:53:33		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-02 08:47:31	2025-06-02 08:48:16	2025-06-02 08:48:16		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-02 08:57:41	2025-06-02 08:59:41	2025-06-02 08:59:41		/search
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 15:50:37	2025-06-01 15:52:27	2025-06-01 15:52:28		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 15:55:23	2025-06-01 15:57:15	2025-06-01 15:57:16		
0.0.0	cm0xjtrro0009f1p5ap	default	false	false	[2025-06-01 15:11:54	2025-06-01 15:14:08	2025-06-01 15:14:09		/search

ちゃんと分析したら色々わかってきた

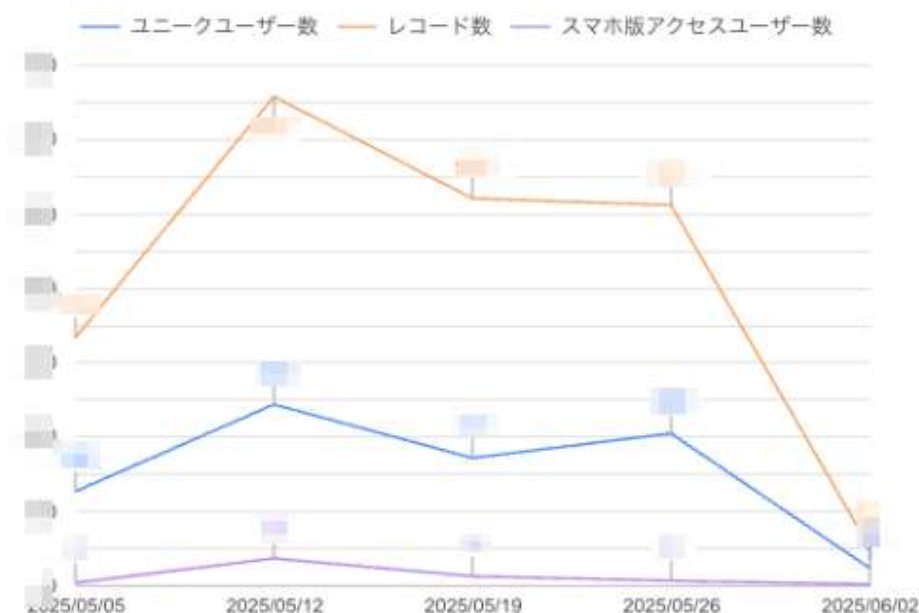
弊社PdM主導のもと、AIエージェント（A-BOSS）の利用傾向を調査

AIエージェントを使う前に離脱してしまったり、継続的に利用しているユーザー傾向を洗い出したり、いくつかの施策の効果測定を実施したりと色々わかってきました。

アクティブユーザー推移

全体の利用者数を可視化

アクティブユーザー推移



アクセスユーザー推移

来訪者数を可視化

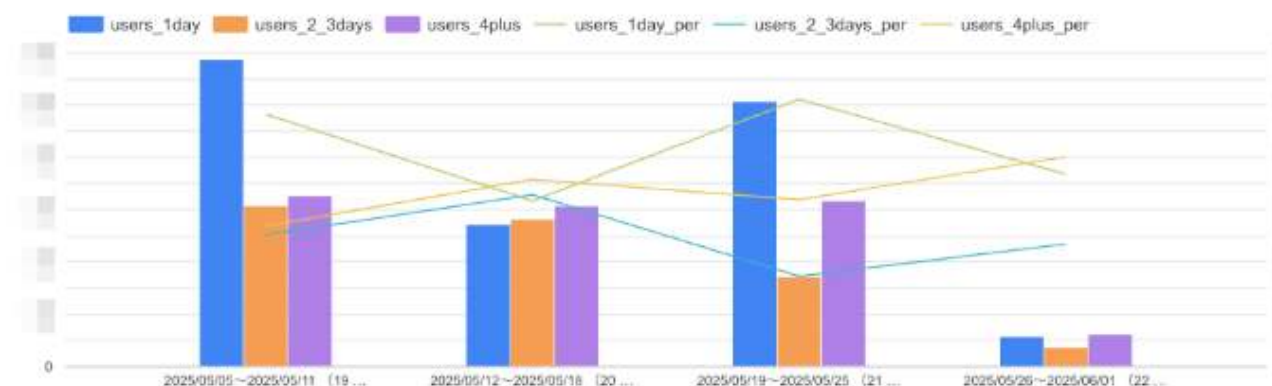
アクセスユーザー推移



セグメント別（継続日数）推移

ユーザーの利用継続性を可視化

アクティブユーザー数・セグメント別

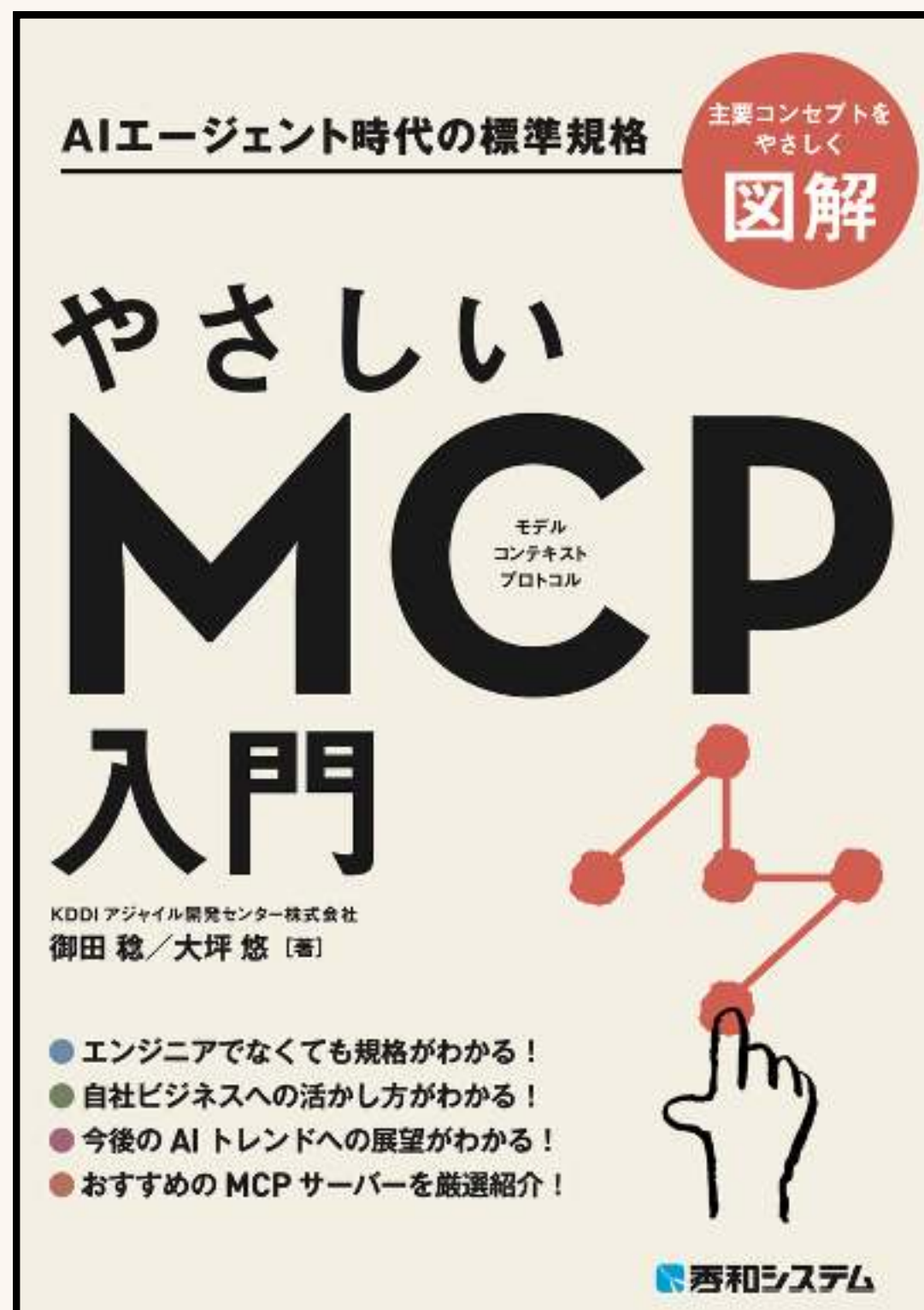


おわり

おわり？

**このプレゼンテーションには
プロモーションが含まれます**

最後に宣伝(唐突!)

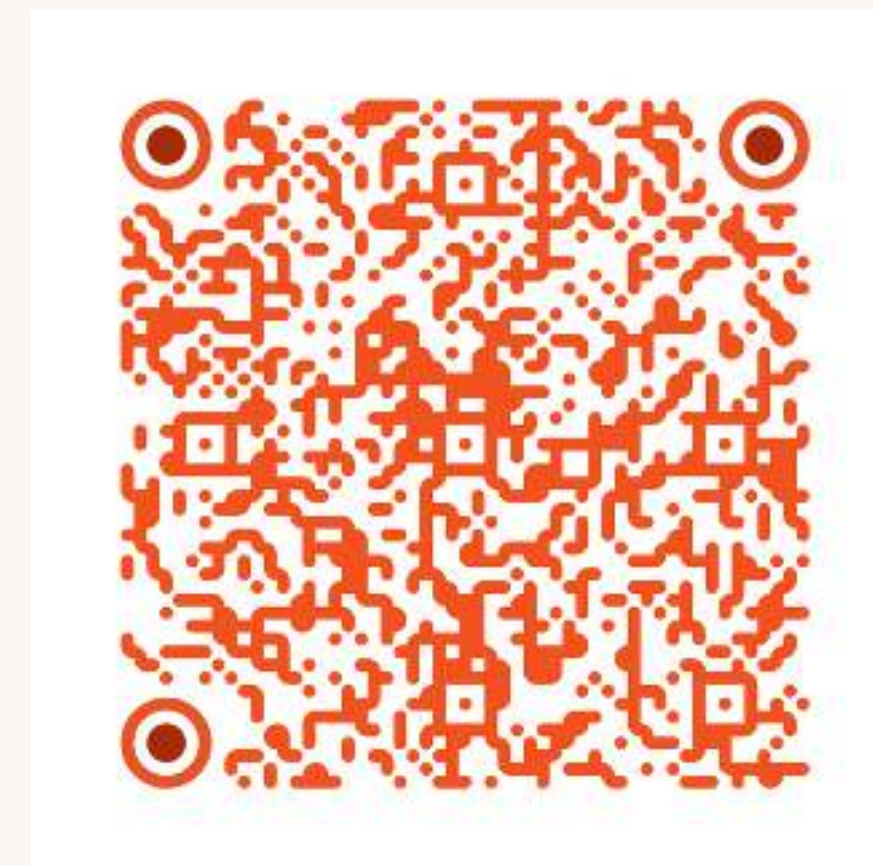


これからの生成AI時代に欠かせない技術になりつつあるMCP

まだ入門できてないエンジニアの皆さんだけでなく、**ビジネスサイド・経営者の皆さん**にも面白い内容がふんだんな読みやすい一冊です。

いぬの機嫌を犠牲にして書いた一冊なのでぜひお手に取っていただければと思います。

**Amazonで
予約受付中!**
7.1発売



Presented by Yu Otsubo

Thank you very much!

<https://tubone-project24.xyz>

