



Medicrawl: An example within Swissmedic and its digitalization

29. March 2022

Dr. Nicolás Andrés Pérez González

Schweizerisches Heilmittelinstitut
Institut suisse des produits thérapeutiques
Istituto svizzero per gli agenti terapeutici
Swiss Agency for Therapeutic Products

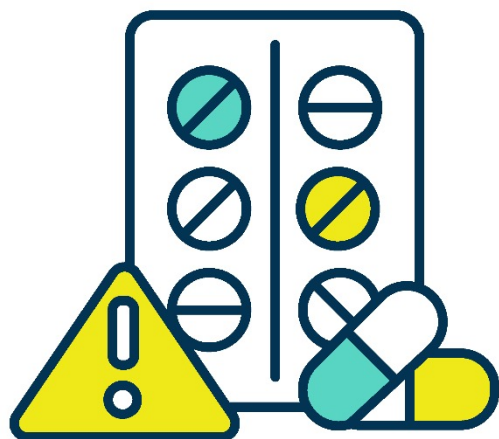
Hallerstrasse 7, 3012 Bern, Schweiz
www.swissmedic.ch



Swissmedic is the National
authorisation and supervisory
authority for drugs and medical
products

It ensures **high-quality, safe and**
effective medical products





HSA & International Counterparts Continue to Clamp Down on the Illegal Online Sale of Health Products



FDA NEWS RELEASE

FDA launches global operation to crack down on websites selling illegal, potentially dangerous drugs; including opioids

Illegally marketed opioids and drugs claiming to treat cancer and HIV are among the products targeted, along with the websites that sell them

Press release

Over 3 million medicines and devices seized in UK as part of global crackdown



[1] [HSA & International Counterparts Continue to Clamp Down on the Illegal Online Sale of Health Products](#)

[2] [Over 3 million medicines and devices seized in UK as part of global crackdown - GOV.UK \(www.gov.uk\)](#)

[3] [FDA launches global operation to crack down on websites selling illegal, potentially dangerous drugs; including opioids | FDA](#)

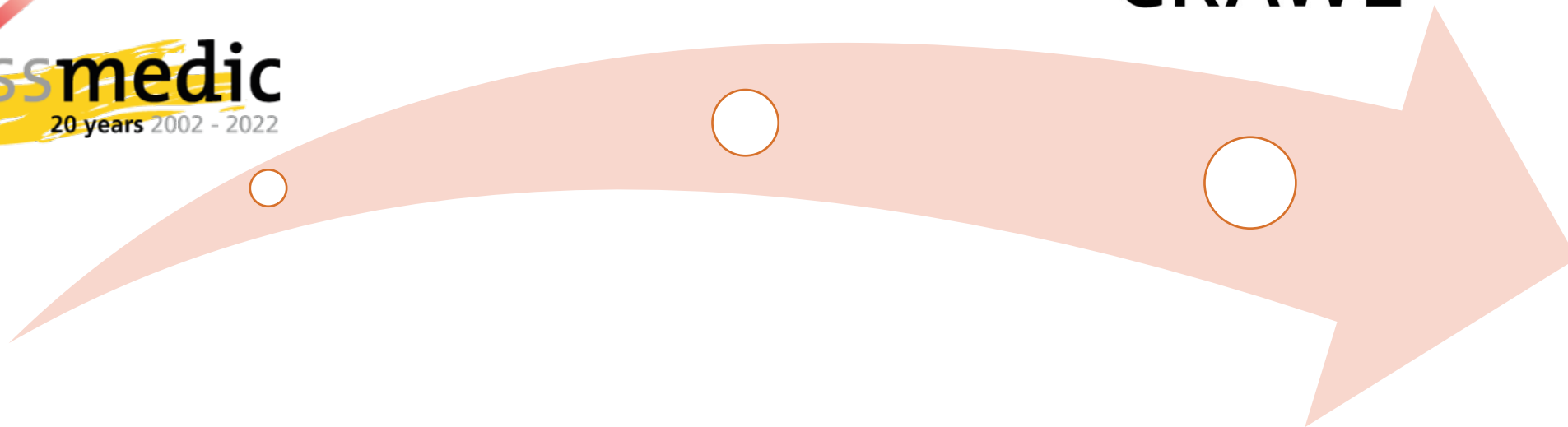


This affects everyone



**/MEDI
CRAWL**

Swissmedic 4.0



Among our core competencies

- the authorisation of medicinal products
- market monitoring of medicinal products and medical devices
- clinical trials and laboratory testing regarding the quality of medicines



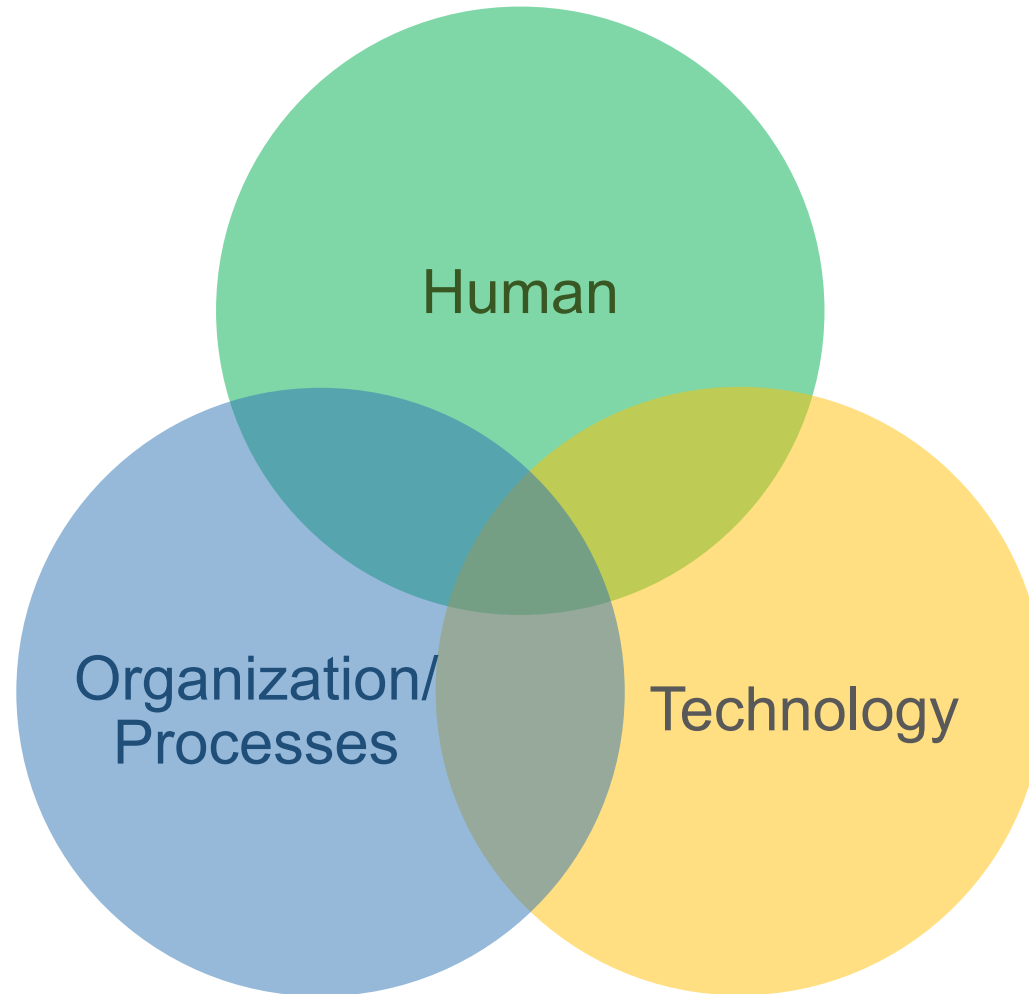


Swissmedic 4.0



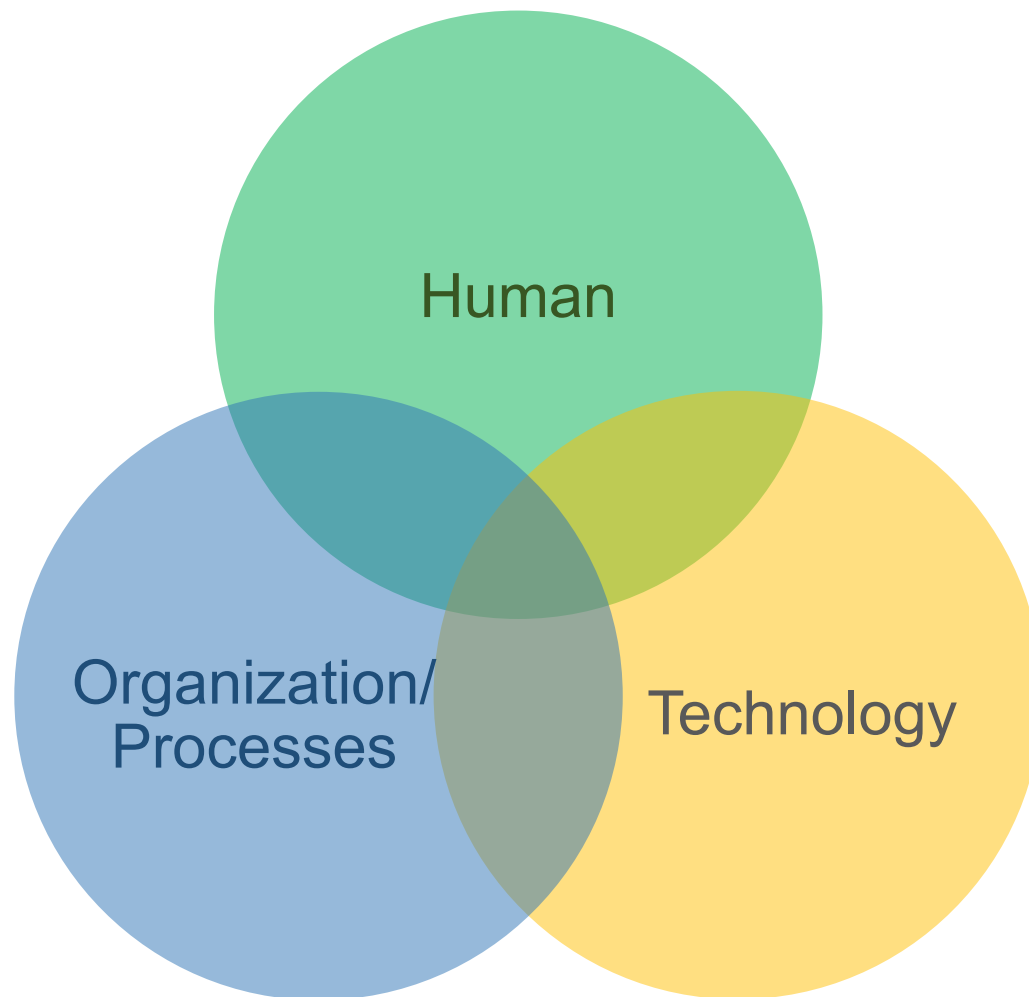
- Create a space for change and innovation
- Identify opportunities and risks, learn from mistakes and improve
- Transfer learnings in the organization

Our approach to digitalization



/MEDI
CRAWL

Our approach to digitalization



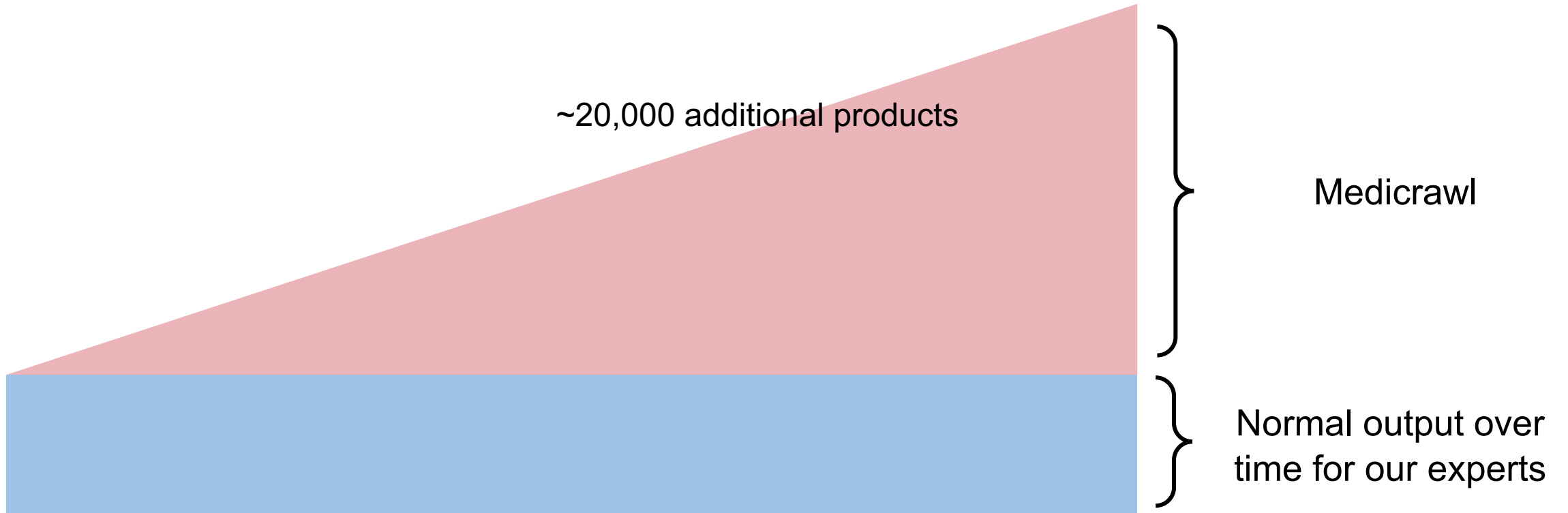
**/MEDI
CRAWL**



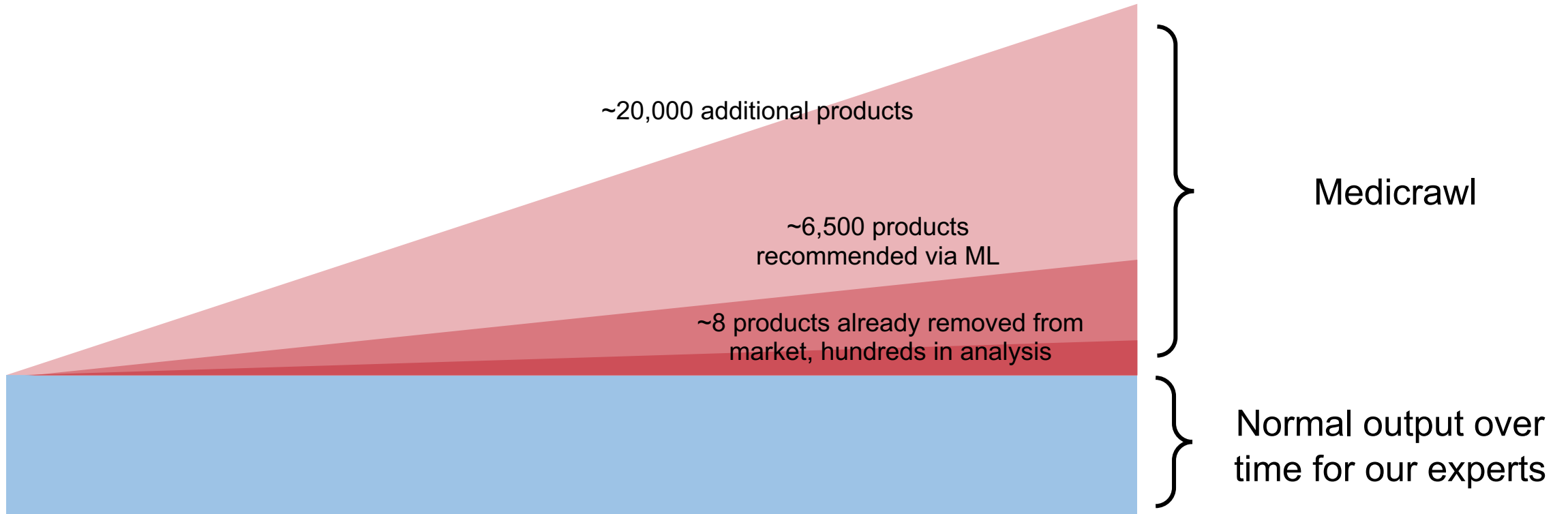
Medicrawl has a clear goal: **to look for illegal products in online markets.**

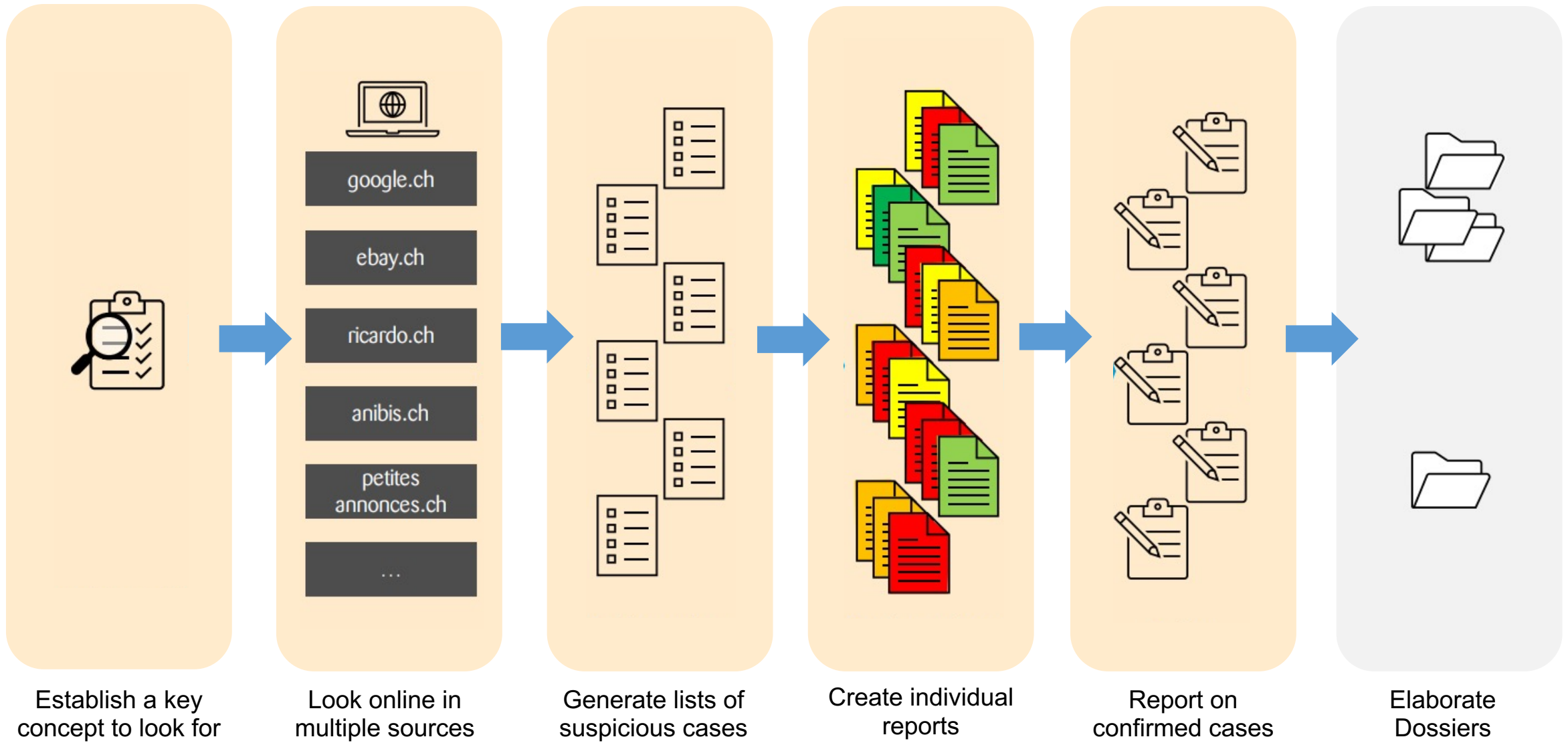
Medicrawl **evaluates if products** so that experts can establish if they are breaking the law

Our goal: to expand coverage on online markets



Our goal: to expand coverage on online markets







7

cyclotest® Chlamydien-Schnelltest

Produktdetails / Pflichtangaben Weitere Produktinformationen Wirkstoffe
Chlamydien-Testzone Dieser **Test** dient zum Nachweis des Bakteriums **Chlamydia trachomatis**. **Infektionen mit Chlamydia trachomatis** sind die häufigsten sexuell übertragenen **Infektionen**. Unbehandelte **Infektionen** führen häufig zu **Unfruchtbarkeit** und erhöhen das Risiko von Eileiterschwangerschaften und Frühgeburten. Zur Probenahme vom **Gebärmutterhals**, ausschließlich zur Anwendung durch Frauen geeignet. Testergebnis einfach ablesbar bereits nach 20 Minuten Hohe diagnostische Sicherheit von bis zu 98,3 % Wie wende ich den cyclotest® **Chlamydien** Schnelltest an? Der **Test** ist für die Anwendung zuhause. Dazu muss ein Abstrich vom **Gebärmutterhals** genommen werden. In Verbindung mit den Lösungsflüssigkeiten wird bereits nach 20 Minuten ein Ergebnis auf der Testkassette angezeigt. cyclotest® **Chlamydien** Schnelltest ist ausschließlich zur Anwendung bei Frauen geeignet. Treffen Sie keine medizinisch wichtigen Entscheidungen ohne Ihren **Arzt!**

[Show less](#)Source: shop-apotheke.ch

Bookmark

In Progress

Do not show again



1

CYCLOTTEST Chlamydien-Schnelltest 1 St

cyclotest Chlamydien-Schnelltest Abstrichtest für Frauen zum Nachweis von **Chlamydien**. Einziger Schnelltest für die Anwendung zu Hause. Dieser **Test** dient zum Nachweis des Bakteriums **Chlamydia trachomatis**. Zur Probenahme vom **Gebärmutterhals**, ausschließlich zur Anwendung durch Frauen geeignet. Nicht für die Anwendung durch **schwängere** Frauen geeignet. Nicht während und bis 3 Tage nach der Periode oder während eines Harnwegsinfektes durchführen! Testergebnis einfach ablesbar bereits nach 20 Minuten Hohe diagnostische Sicherheit von bis zu 98,3 %

[Show less](#)Source: freiburg-im-hbf.easyapotheken.de

Bookmark

In Progress

Do not show again

☰ **Sorting:** Highest Relevance☰ **Filter**

Search

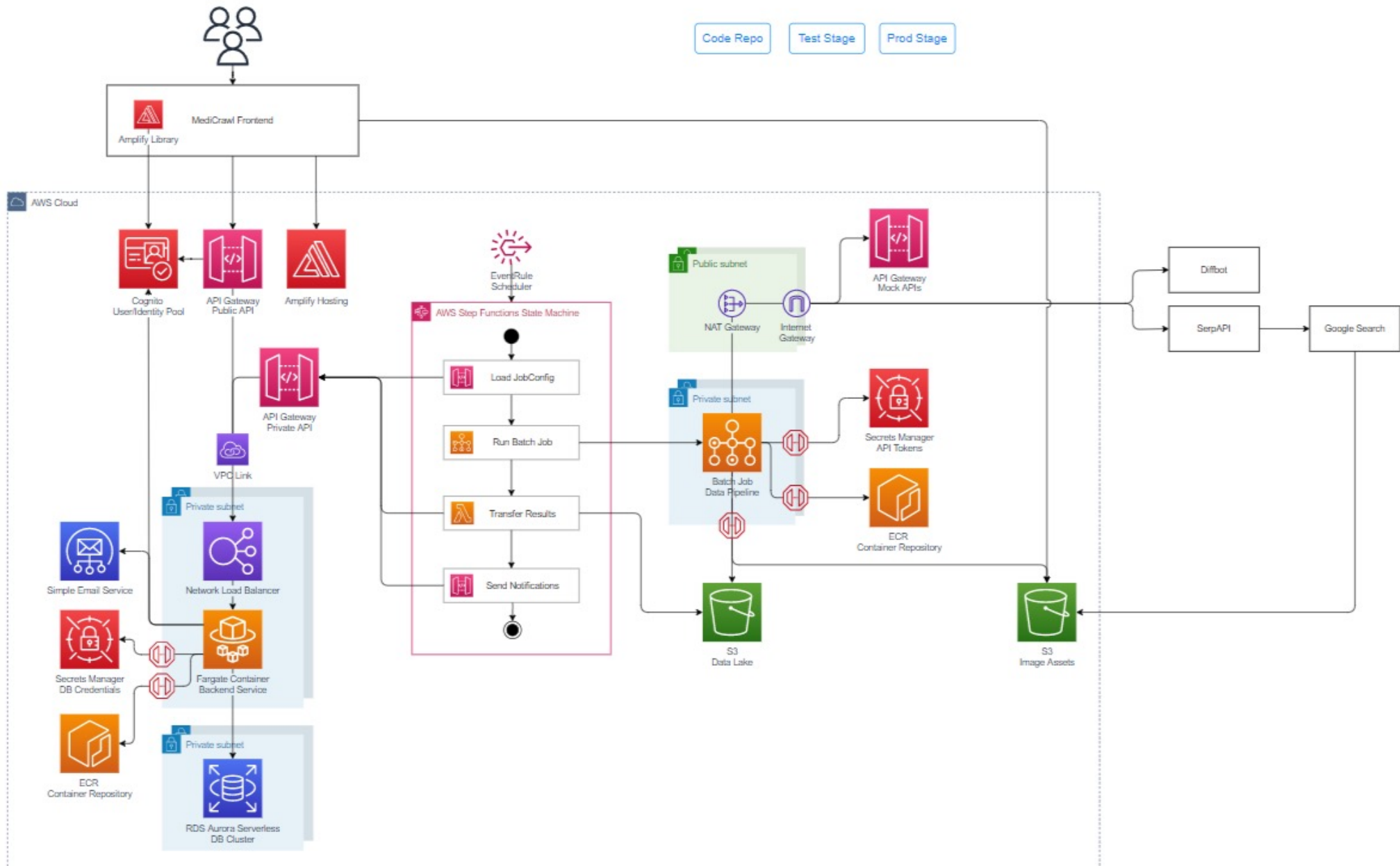
☰ **Action for all results**

Bookmark

In Progress

Do not show again

PDF-Export

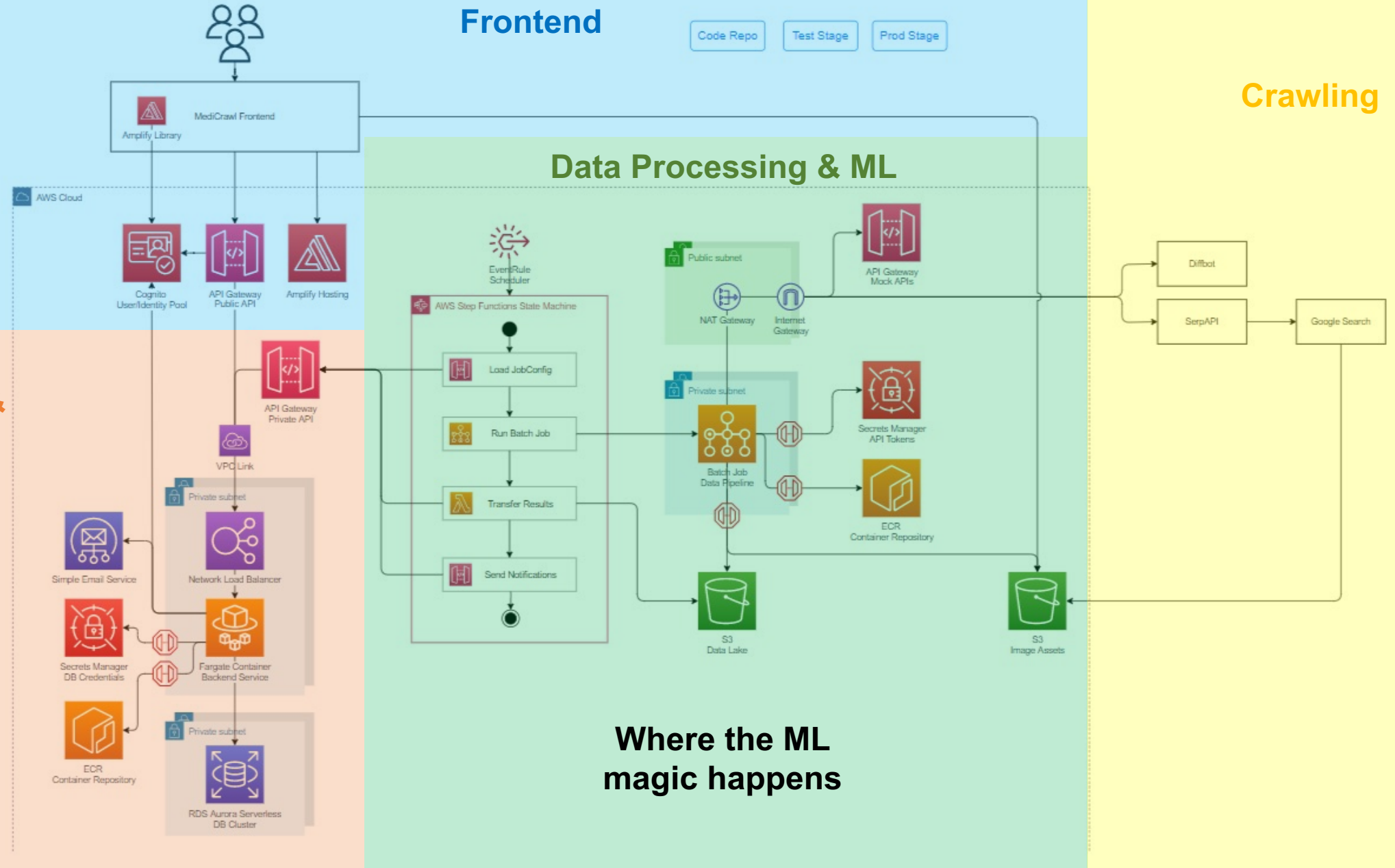


Frontend

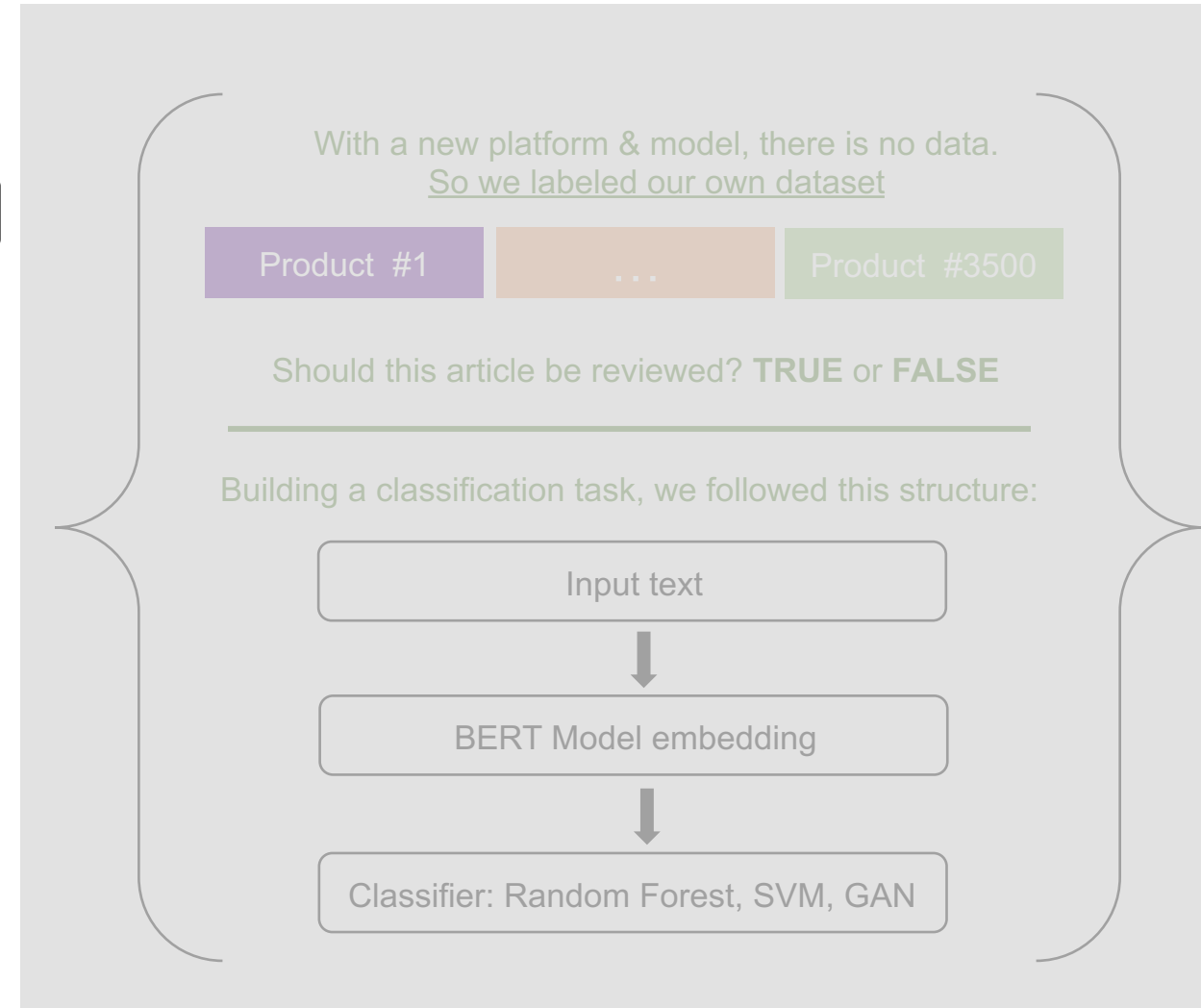
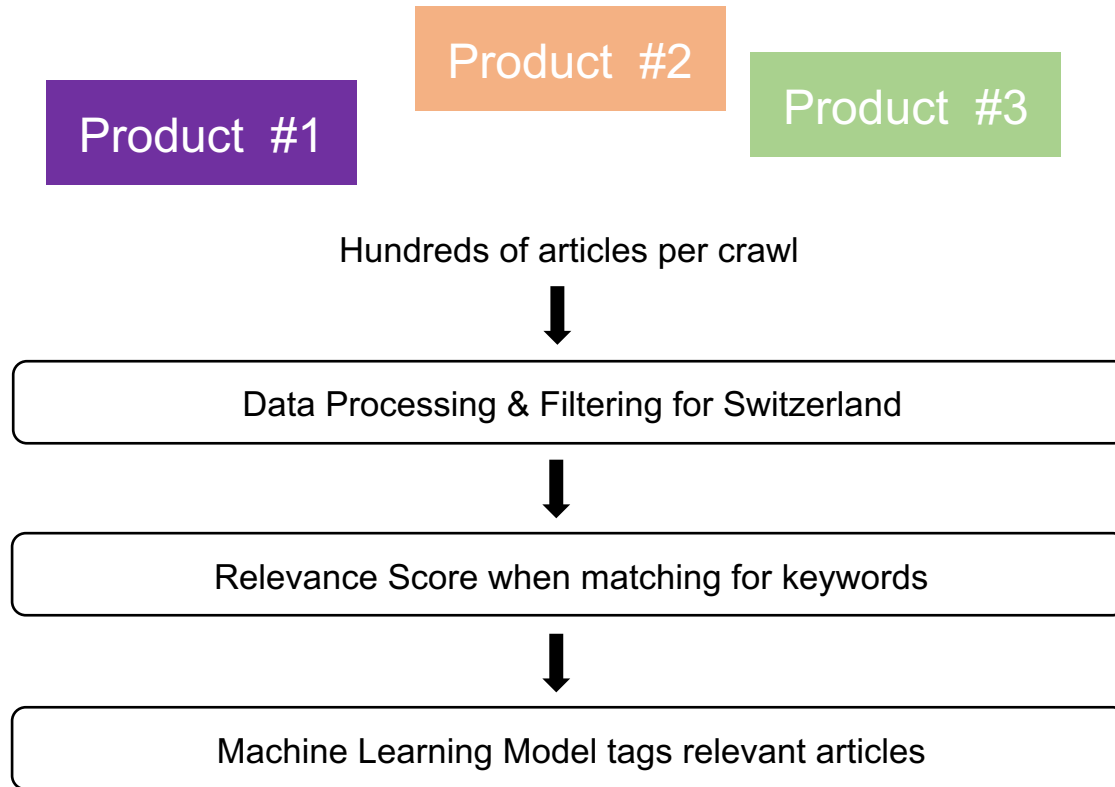
Code Repo Test Stage Prod Stage

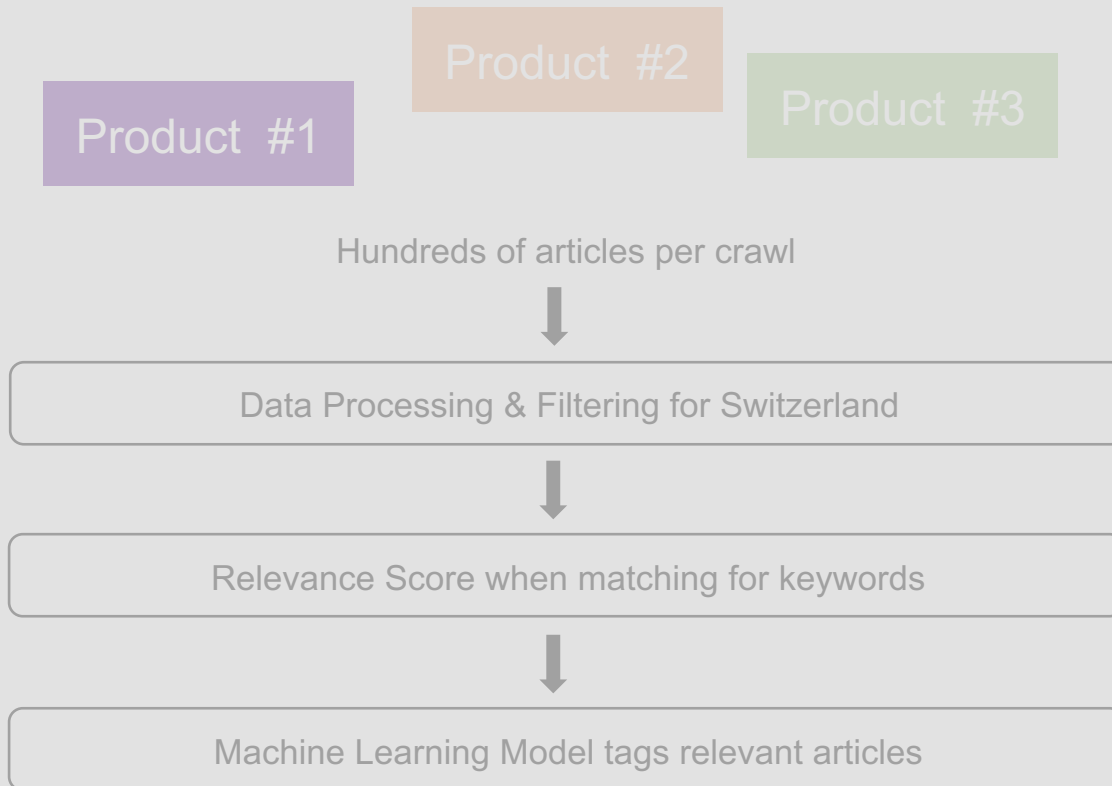
Crawling

Database & Others



Where the ML magic happens





With a new platform & model, there is no data.
So we labeled our own dataset

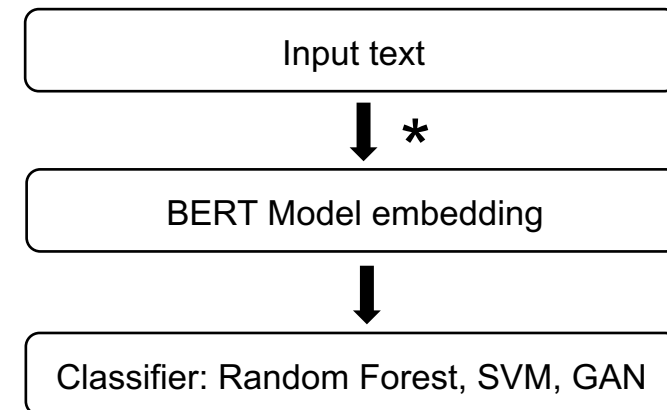
Product #1

...

Product #3500

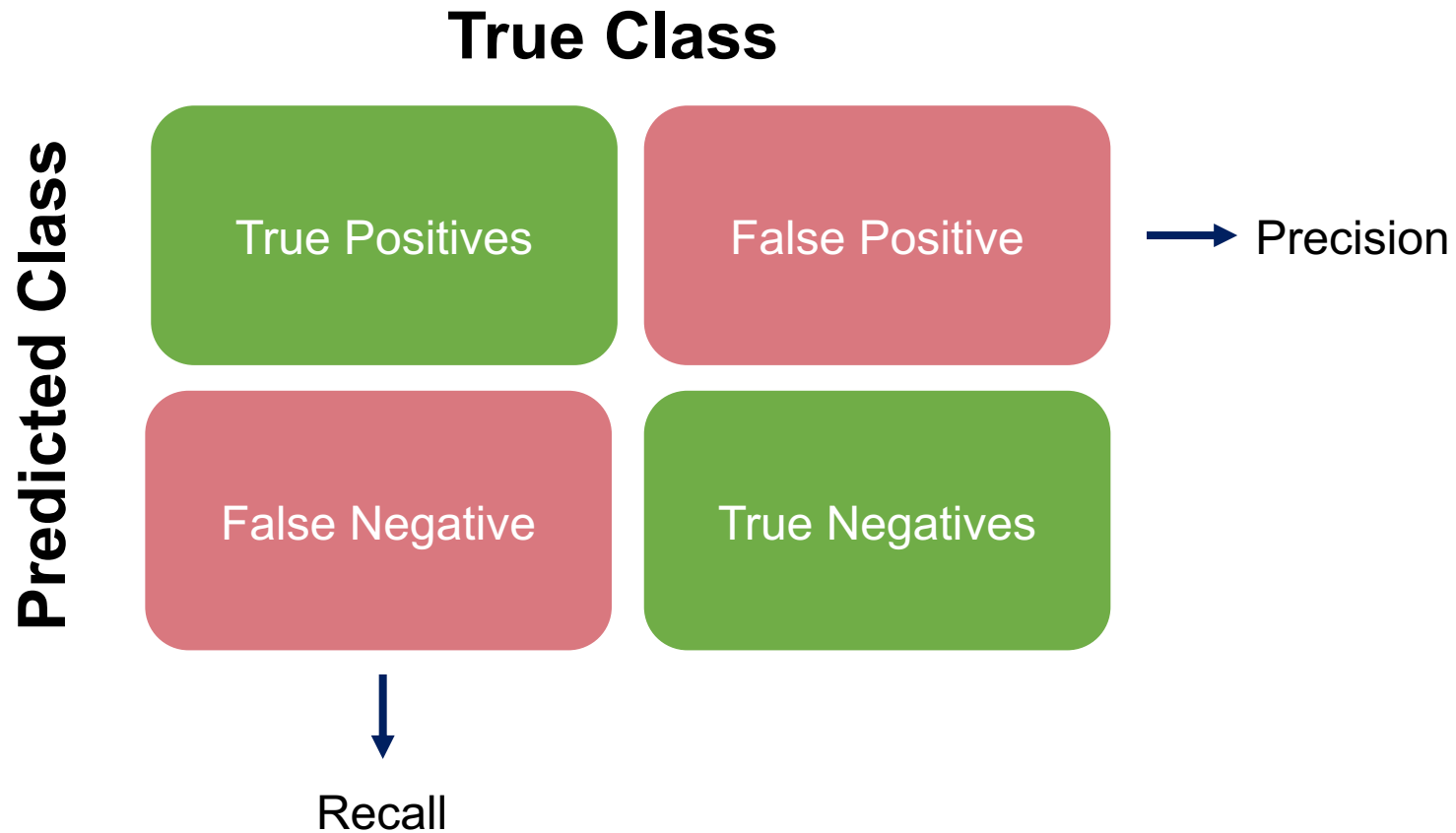
Should this article be reviewed? **TRUE** or **FALSE**

Building a classification task, we followed this structure:

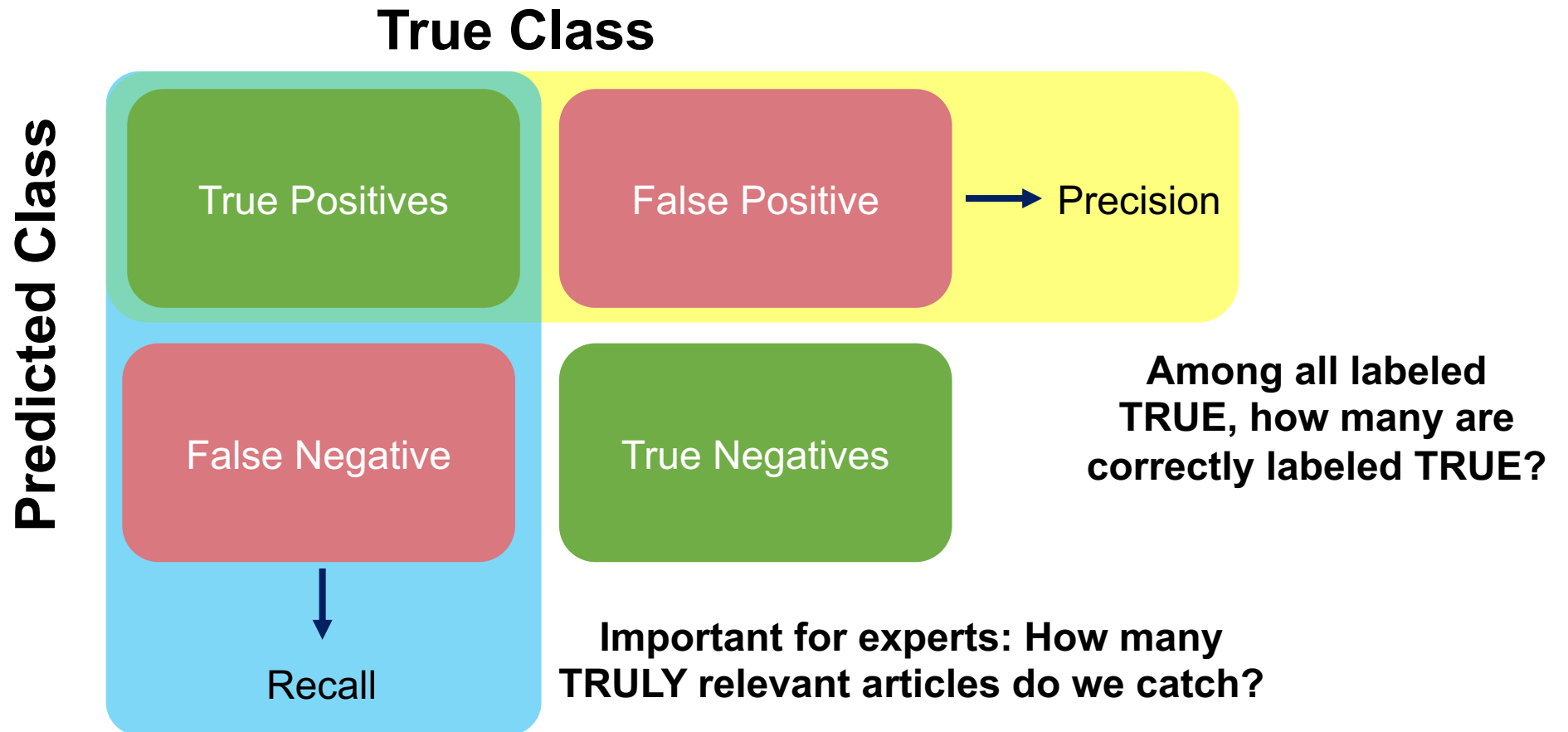


*lots and lots of processing

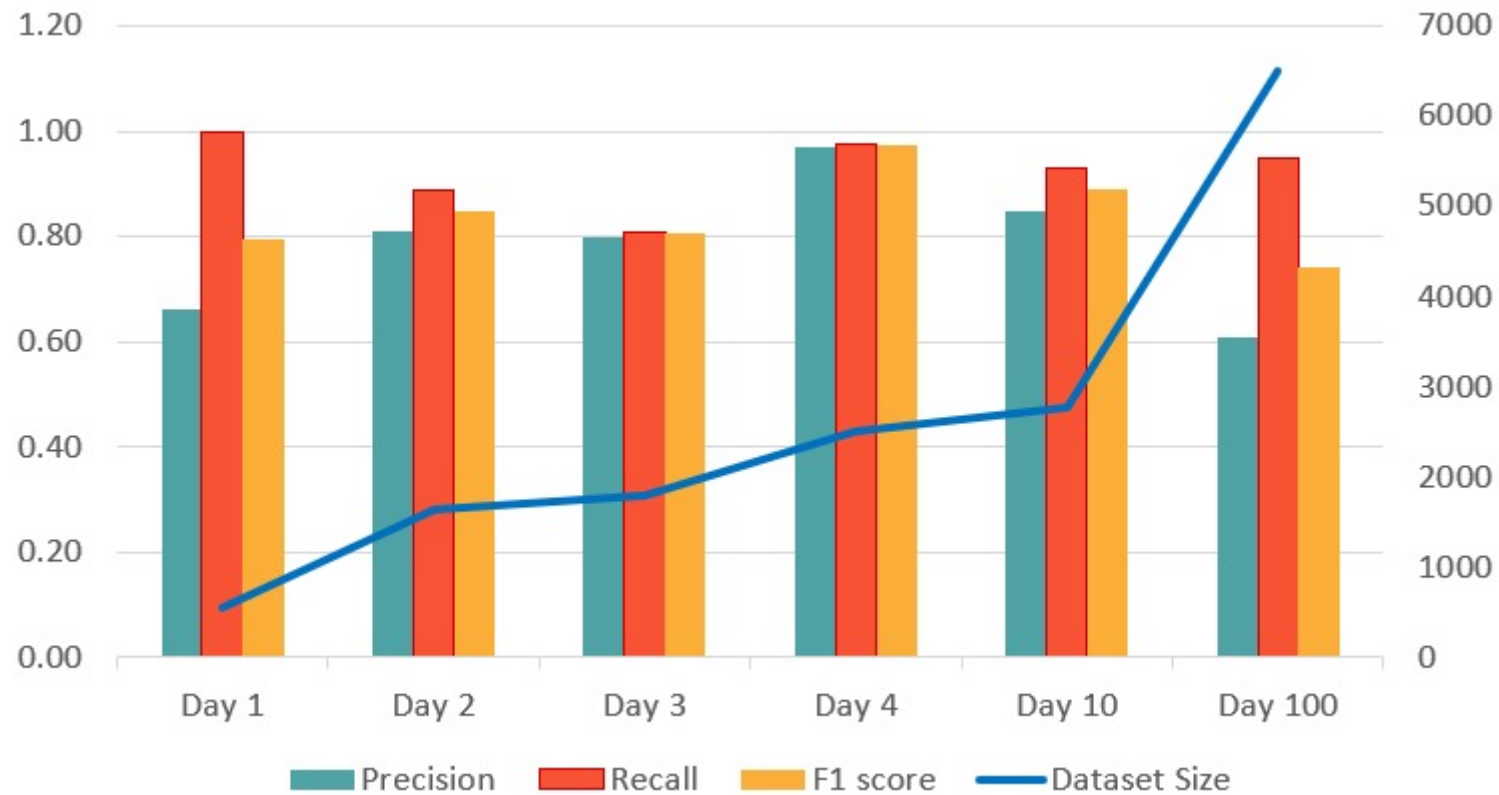
Recall over precision matters to us



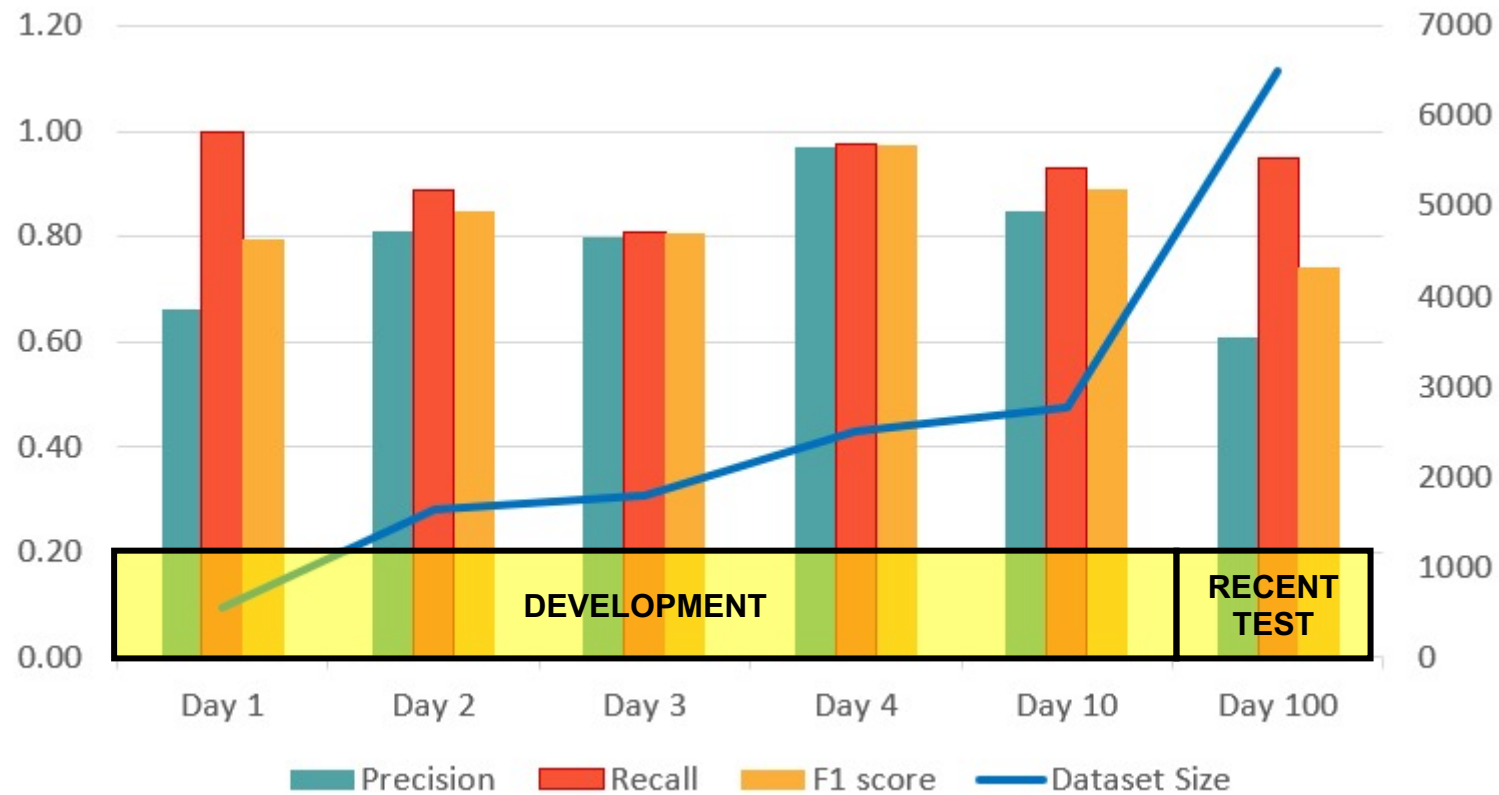
Recall over precision matters to us



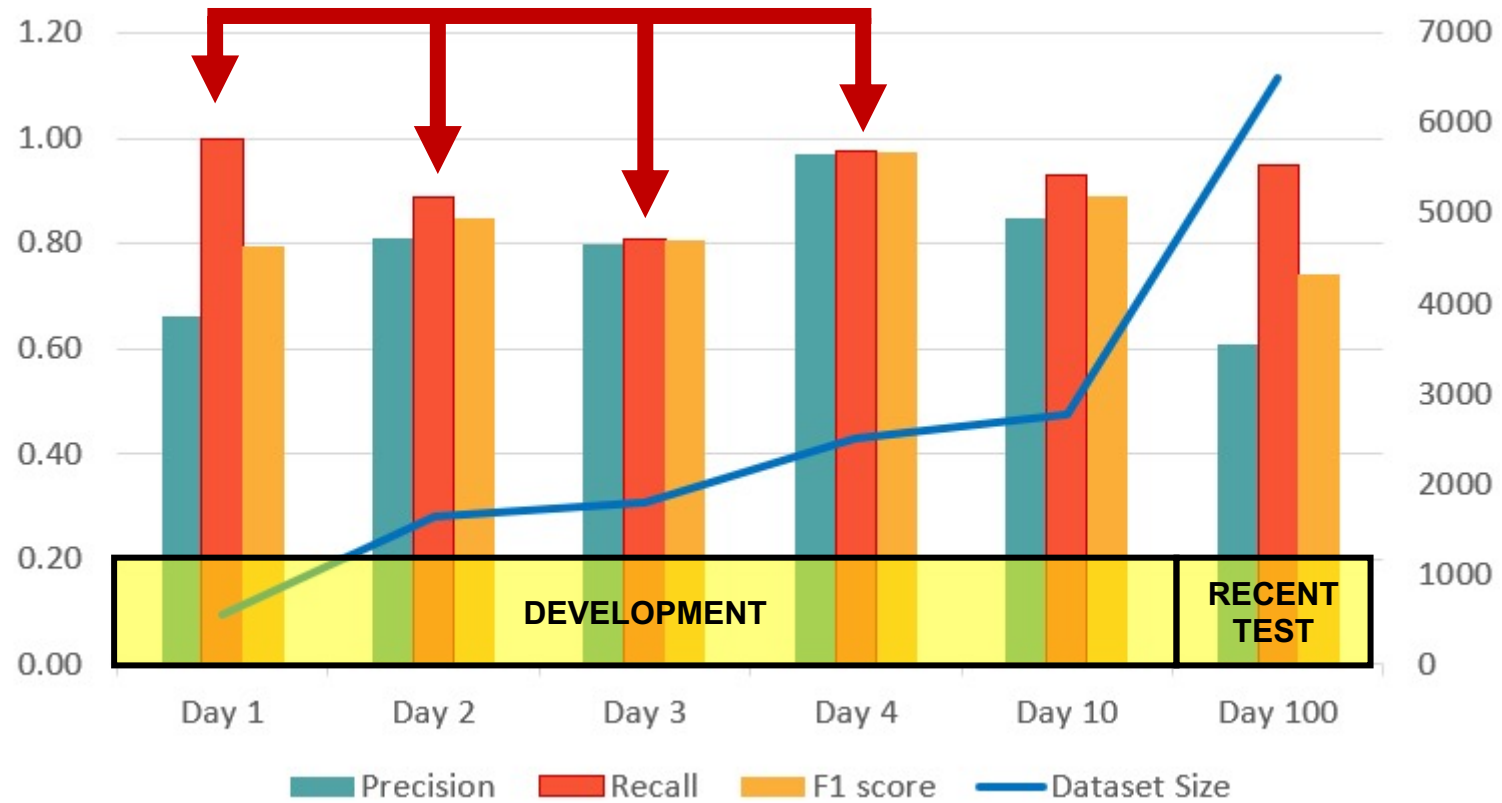
Real usage shows promising results



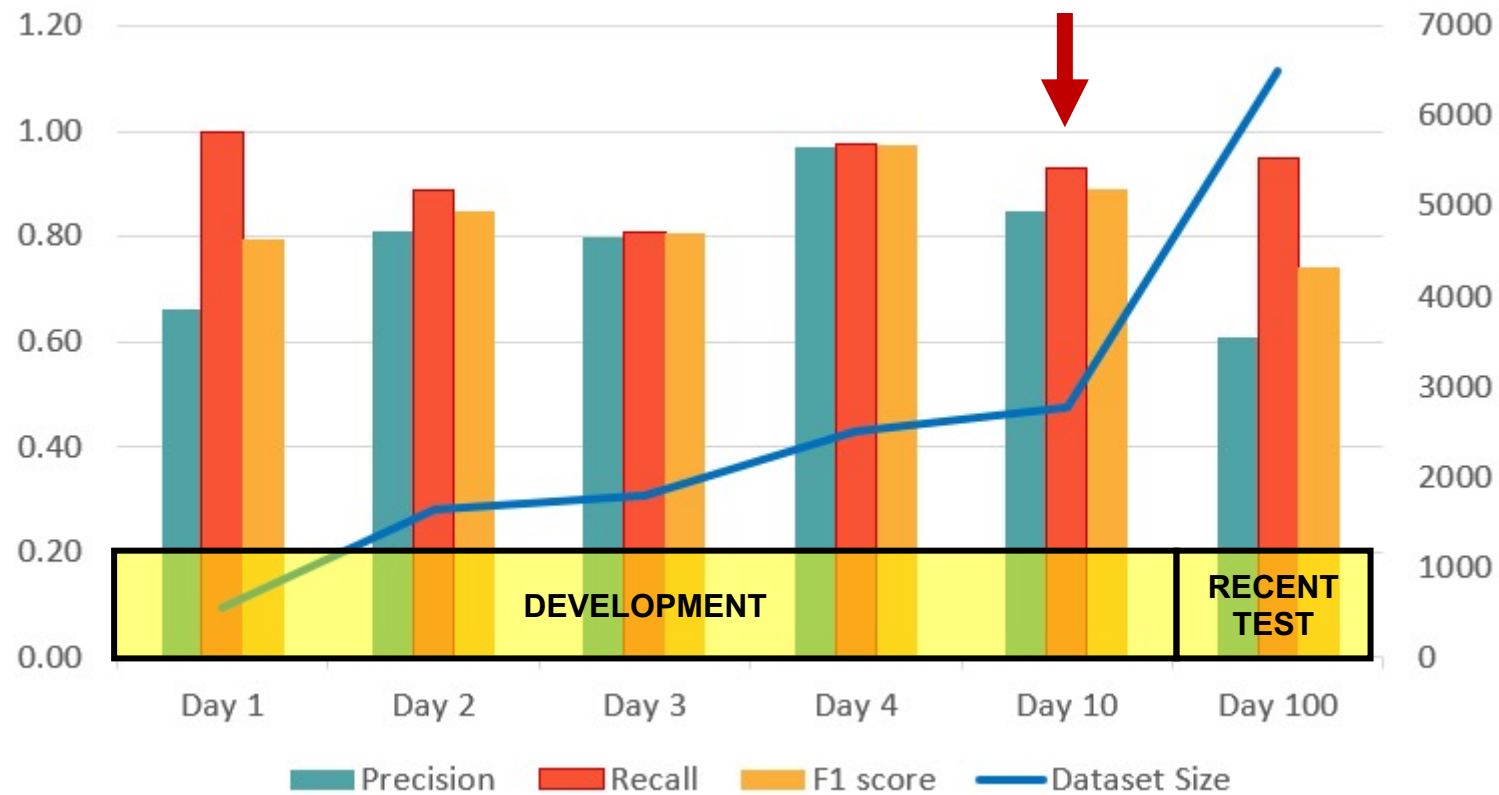
Real usage shows promising results



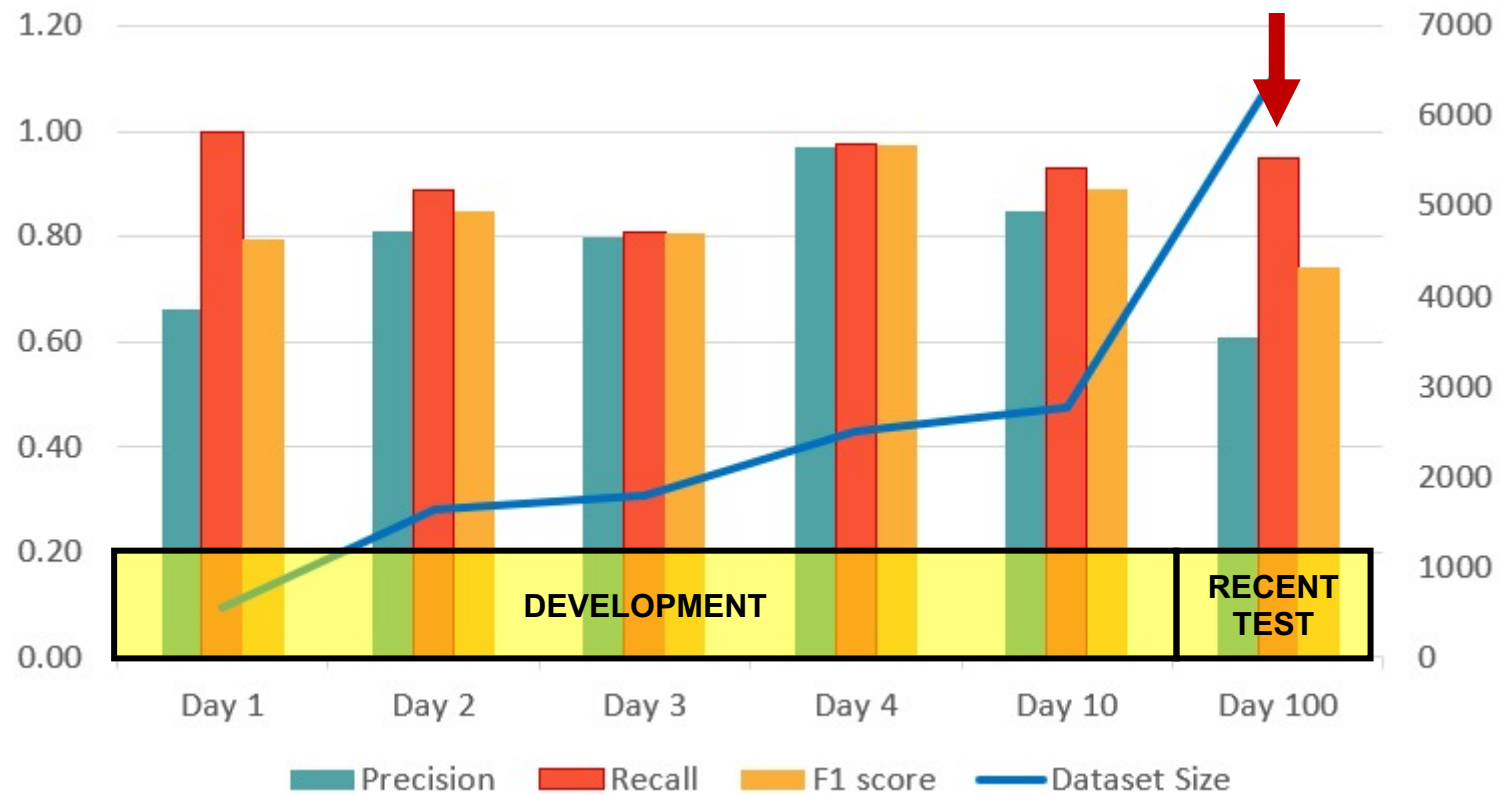
How big should the training dataset be?



To avoid bias, we modified our dataset

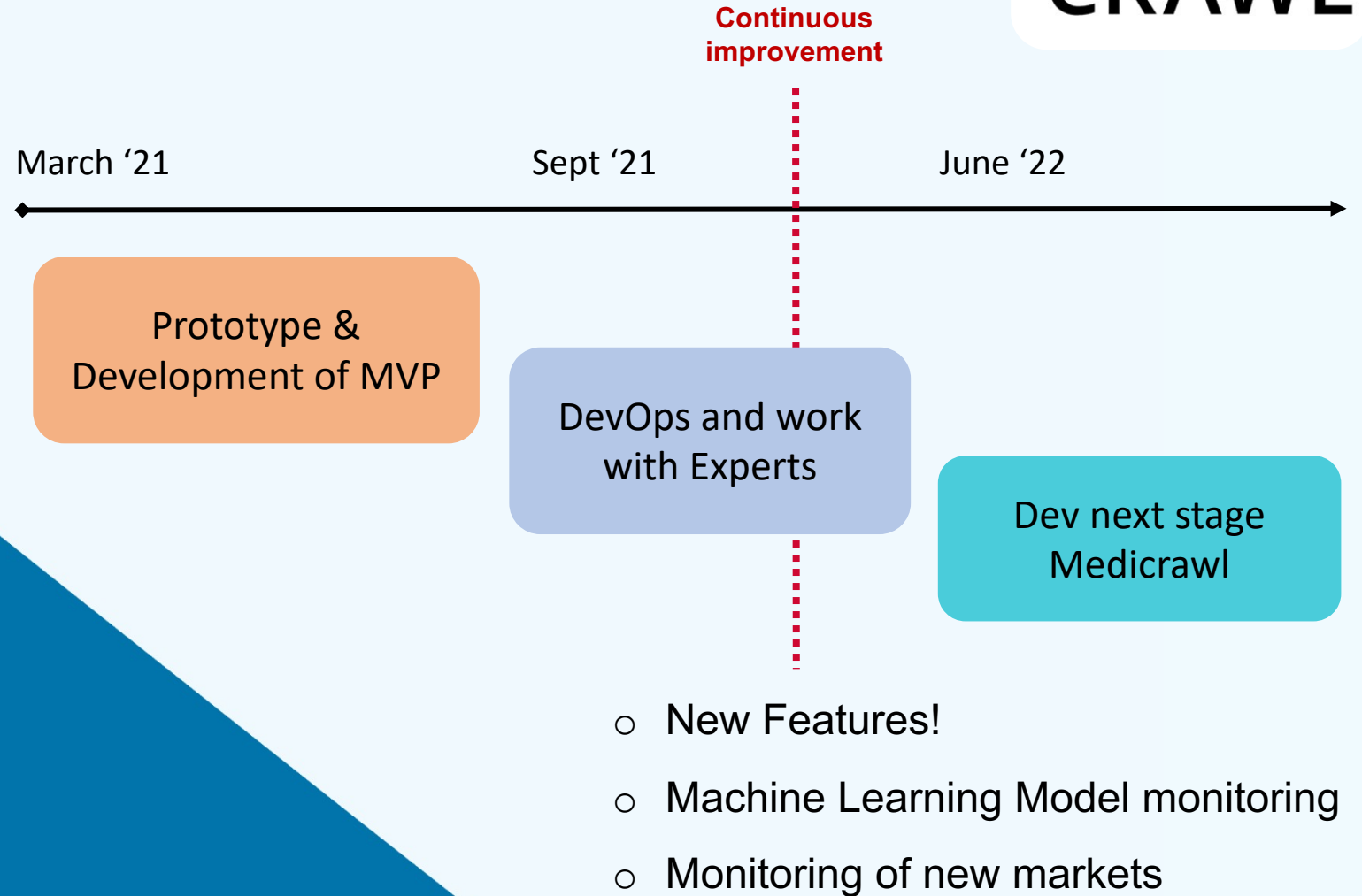


Today's performance has gone down overall, but stayed high where it matters!



What's the future of Medicrawl?

**/MEDI
CRAWL**

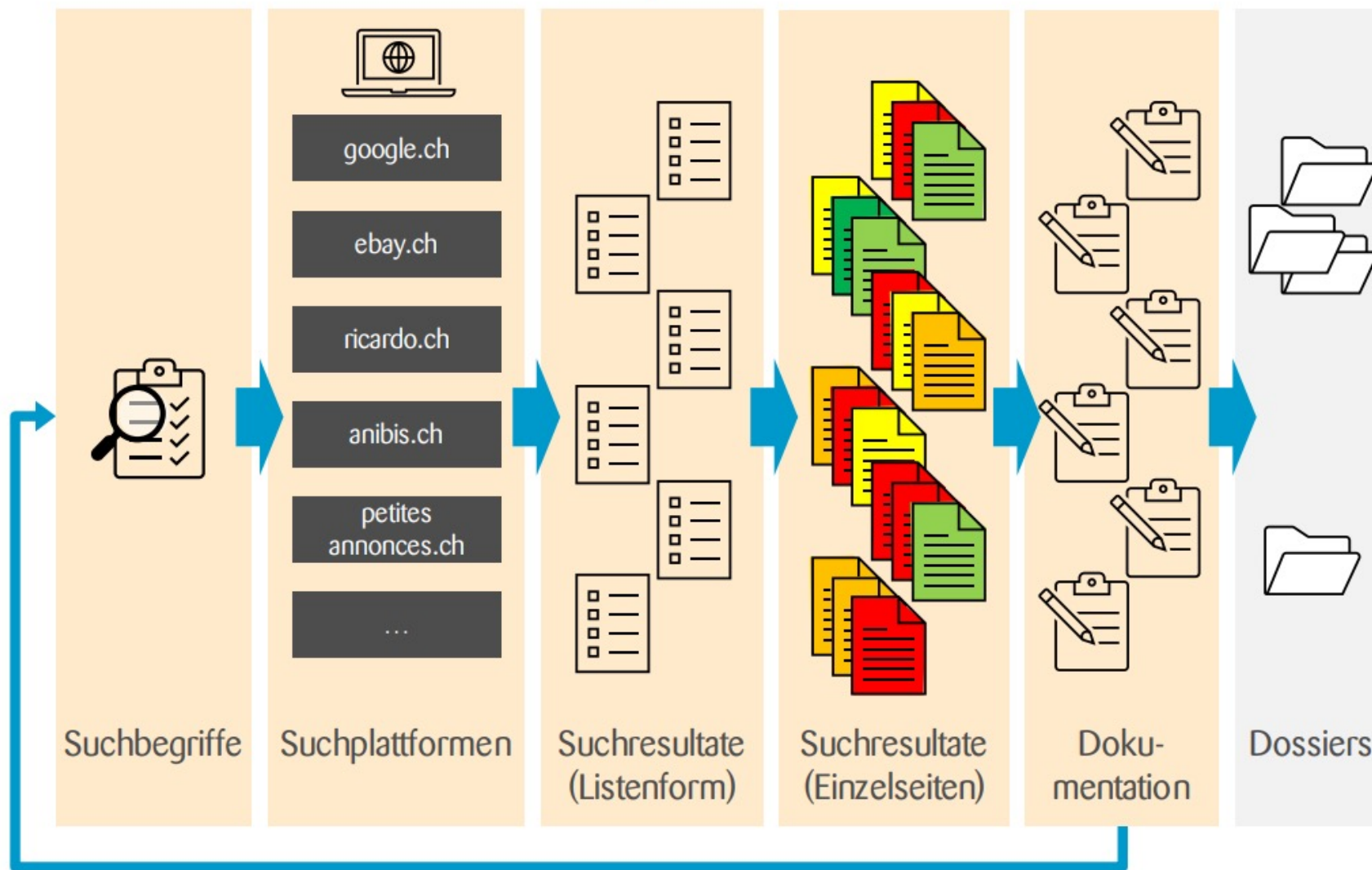




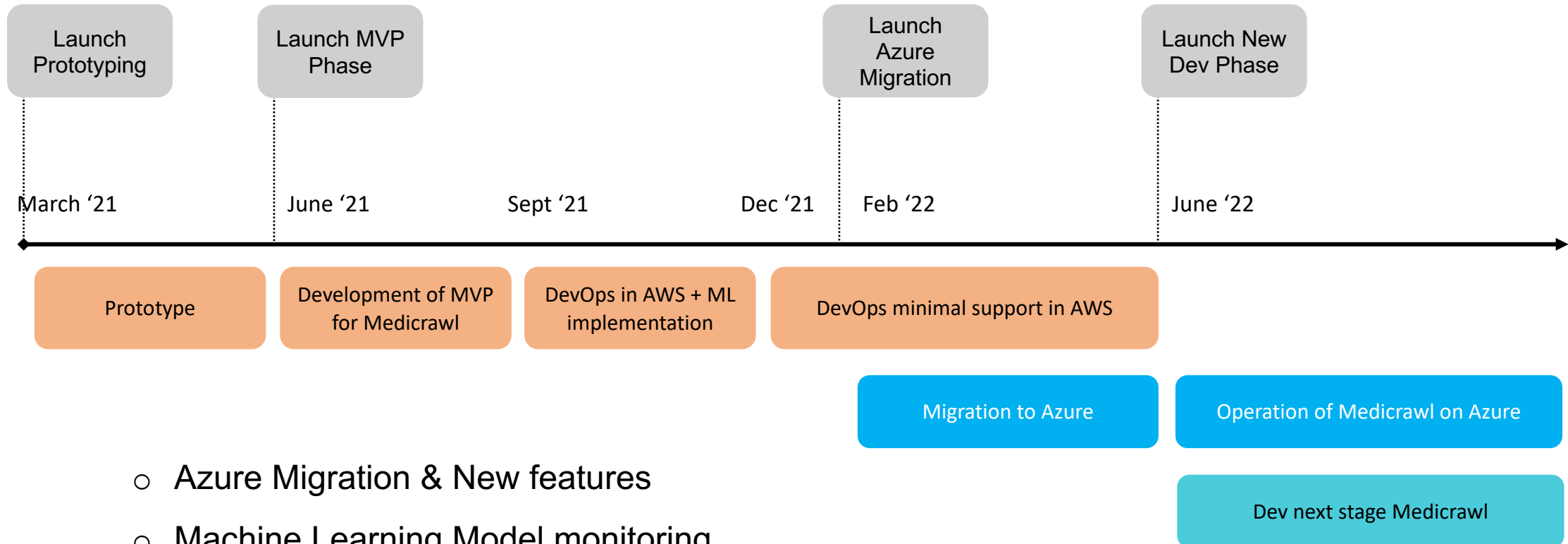
Questions? Cool ideas to share?



Contact me at nicolas.perezgonzalez@swissmedic.ch or find me on 



What's the future of Medicrawl?



- Azure Migration & New features
- Machine Learning Model monitoring
- Monitoring of new markets

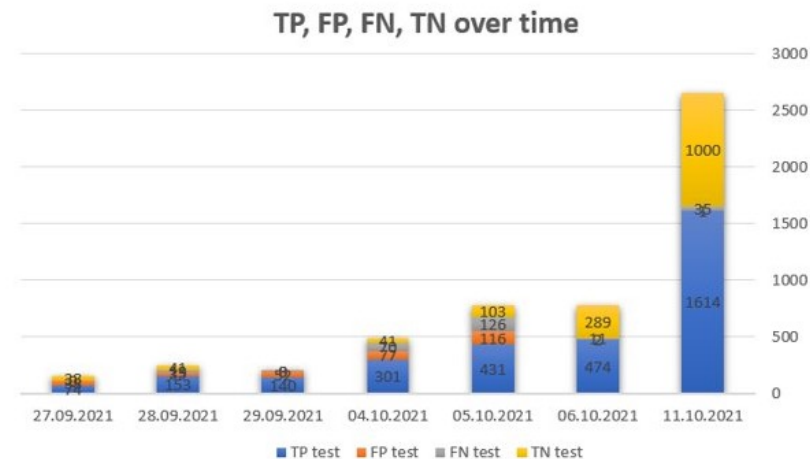
How did the model training go?

First few data collections showed promise with training



Date	Status	Files	Accuracy	Precision	Recall	F1 score	TP	FP	FN	TN	Samples
27.09.2021	Training		0.88	0.92	0.79	0.85	23	2	6	7	400
27.09.2021	Test	mavira_crawling_französisch_	0.75	0.66	1	0.80	74	38	0	38	150
28.09.2021	Test	mavira_crawling_französisch_	0.77	0.75	0.47	0.58	145	48	159	552	900
28.09.2021	Training	mavira_crawling_französisch_	0.85	1	0.93	0.96	87	0	6	99	1392
28.09.2021	Test	mavira_crawling_2021-09-27_	0.78	0.81	0.89	0.85	153	35	19	41	248

Up to 3,500 were collected to reach high performance



But that lead to overfitting! Most recent test lead to an accuracy of 80% in new crawled articles!