

My dream - Investigation of an issue





Here is an issue of interest ...

1) How widespread is this? 2) Do we have this problem?"

1) "Looking at enforcement documents, recalls, adverse events — the issue first appeared 2 years ago and multiple companies have been affected, including 2 of your CMOs and 1 of your competitors"

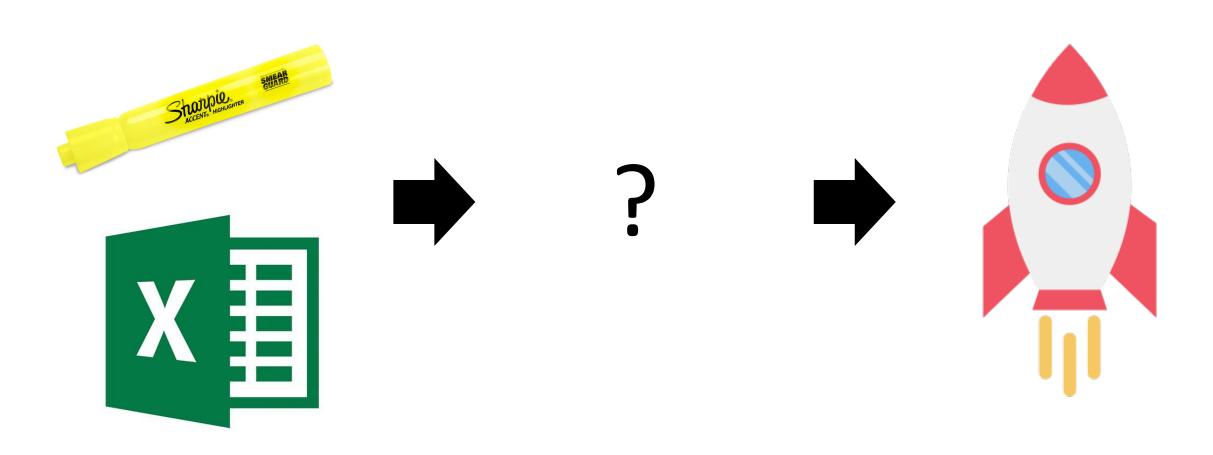
2) "Looking at batch records, deviations and investigations, complaints, and audit records – there is a high likelihood that you have this issue. Here is the highlighted evidence and my confidence intervals for each item..."

"Would you like alerts on this topic as it evolves? check all that apply a) Guidance b) proposed regulations c) enforcement d) recalls e) AE's f) deviations ...

Click ->

Click ->

How do we get there?



Hint: the answer is beyond ChatGPT...

But, isn't ChatGPT is "scary good"?

ChatGPT Prompt

"Write me a 10 line sonnet in iambic pentameter about auditing a drug facility in China, using issues found like - sterility issues, lack of a quality unit, and other items found in FDAs 21 CFR 211"



ChatGPT Output

In sterile chambers, flaws of cleanliness,
Revealed, where sterility's reign was weak.
A quality unit's absence, stark distress,
A void that hindered standards they should seek.

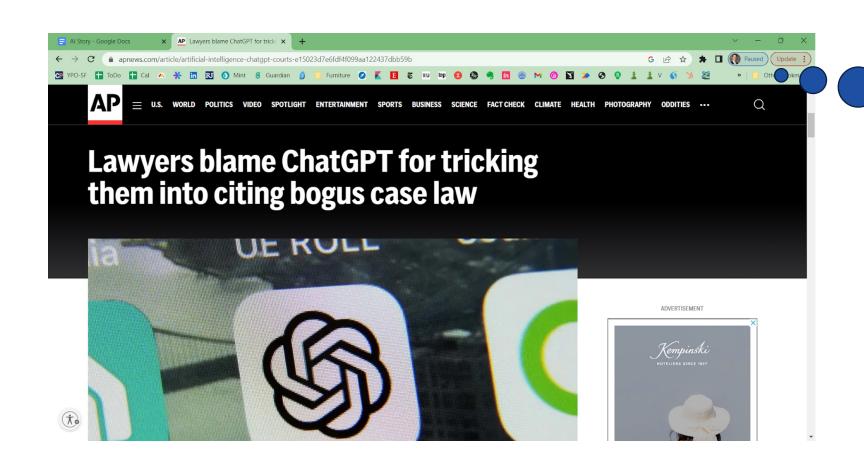
Within those walls, FDAs keenly roved, Unveiling faults in 21 CFR 211. Missteps, discrepancies, so much unproved, In regulations, compliance far undone.

Expired compounds, their potency decayed,
Improper labeling, a chaotic scene.
With findings stacked, concerns were not allayed,
A drug facility's flaws were unforeseen.

Yet audits shed light on shadows concealed, To rectify, ensure the best revealed.



Or is it?



"I did not comprehend that ChatGPT could fabricate cases,"

Chat GPT "hallucinated" precedent cases that "looked real"

Fact: Chat GPT loves to "hallucinate"

ChatGPT Prompt

Using the FDAs 21 CFR 211, what is being cited in this Observation

"OBSERVATION 3

Floors in the ~~~~(b)(4)~~~~ room, #(b)(4) in Building (b)(4) are not level. This was observed on Wednesday, April 30, 2014 during a walk-through of the facility ~~~~~(b)(4)~~~ was not ~~(b)(4)~~



"It is from 211.42, not 211.58. §58 doesn't even have lettered subsections."

chatGPT Output

The mary issue being cited in this observation is 21 CFR 211.58(c), which requires that the "buildings used in the manufacture, processing, packing, or holding of a drug product shall be of suitable size, construction, and location to facilitate cleaning, maintenance, and proper operations." The secondary issue being cited is 21 CFR 211.56(b), which requires that "buildings shall have adequate space for the orderly placement of equipment and materials to prevent mixups between different components, drug products, and to prevent contamination. The uneven floors in the could hinder proper cleaning and materials, which is a violation to see regulations.

"The correct citation should be 21 CFR 211.42(b)" "The content is great, seriously. It's just really weird that ChatGPT would just make up the citation"



Is it good or not? ... it depends

What it is great at

- Knowledge and Information: access to a vast amount of information and can provide accurate facts, explanations, and summaries on diverse subjects.
- 2. Language Fluency: It can generate coherent and grammatically correct sentences, making it proficient in constructing detailed responses.
- **3. Creativity:** can assist with creative writing, offer suggestions, and help develop ideas by generating imaginative and engaging content.
- **4. Conversation Handling:** It can understand context, maintain context across multiple turns, and respond appropriately, creating a conversational flow.

What it is not great at

This list pretty much describes the job of a regulatory and quality professional

- Factual Accuracy: can occasionally generate incorrect or incomplete responses- aka Hallucination
- 2. Explainability: responses are generated based on complex patterns learned from extensive training data, making it difficult to trace the exact reasoning or decision-making process behind each response.
- **3. Repeatability:** due to the inherent randomness in the model's architecture and the generation process, there might be variation in the responses.
- 4. Understanding Nuance and Context: it may still occasionally misinterpret or fail to grasp the subtleties, sarcasm, or context-dependent meanings in a conversation.
- Real-Time Information: ChatGPT's knowledge cutoff is in September 2021

The more specific, obscure, or recent the data needs to be, the more LLMs like ChatGPT struggle



How is Enterprise Al different from ChatGPT?

Types of Al

Supervised learning



"Labeling things" – text extraction, information extraction, translation, scoring

Unsupervised learning



"Clustering things" – trend and anomaly detection (e.g. GE airplane turbines)

Reinforcement learning



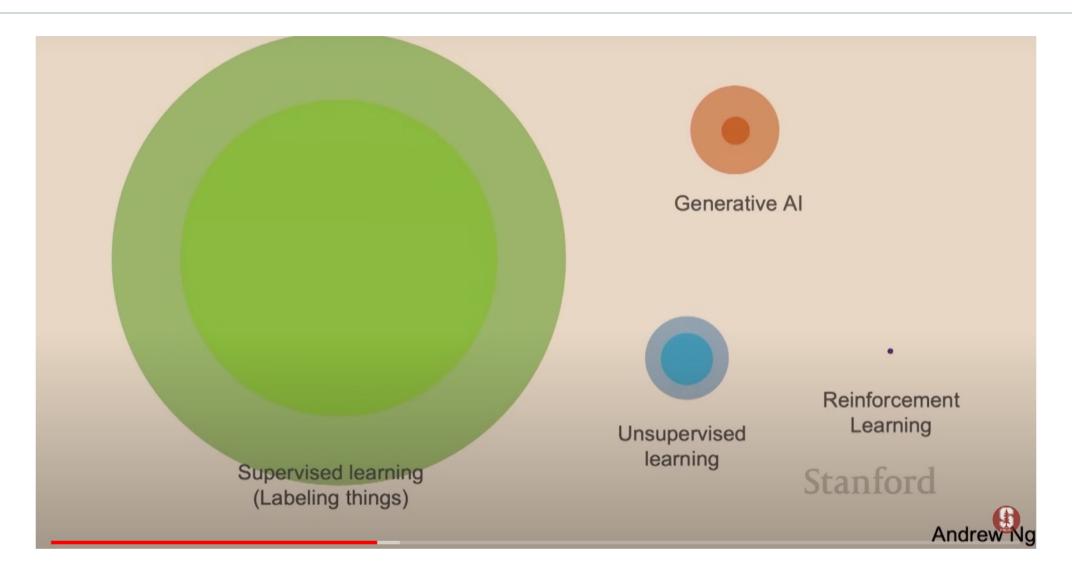
Optimization through rewards

Generative Al

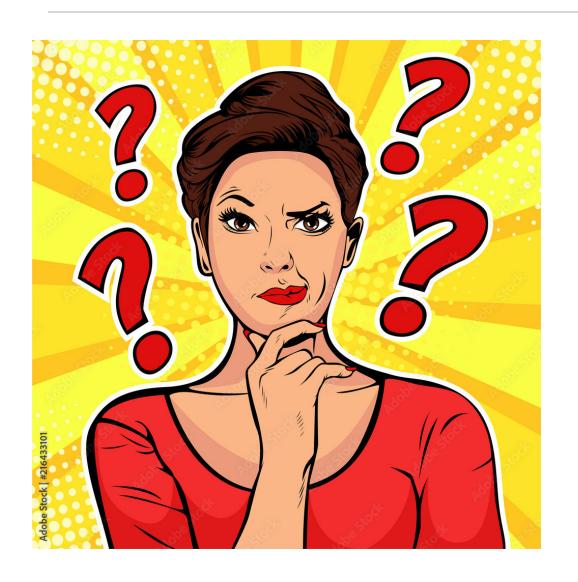


Text / image / song creation

Value to the Enterprise -> 3 year view



11



Most of the value comes from "Labeling things?"

A music example

Analog / Expert Driven

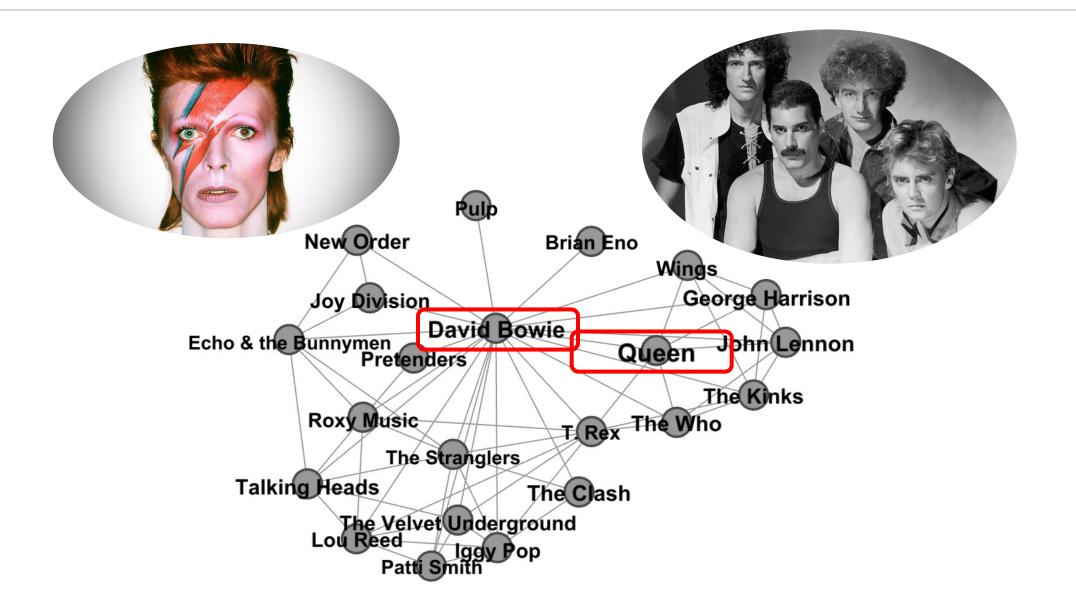


Digital / Al driven





It all started with Pandora...

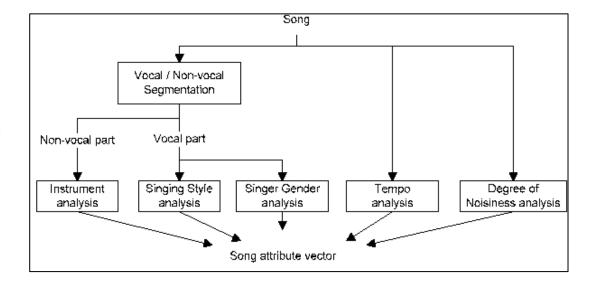


Pandora – the "Music genome project"

attribute	description	type	example
song ID	a unique identifier	nominal	101
title	the title of the song	nominal	Dancing Queen
artist	the name of the performing artist	nominal	ABBA
album	the name of the album of the song	nominal	Greatest Hits
track	the track number of the song on the album	numerical	1
duration	the duration of the song (in seconds)	numerical	232
file size	the size of the file of the song (in MB)	numerical	3.63
genre	the genre to which the song belongs	categorical	popular
year	the year the song was released	numerical	1975
tempo	the tempo of the song (in beats per minute)	numerical	100
tempo marking	the description of the tempo of the song	ordinal	medium
mood	the overall mood of the song	categorical	happy
key	the tonality of the song	categorical	G major

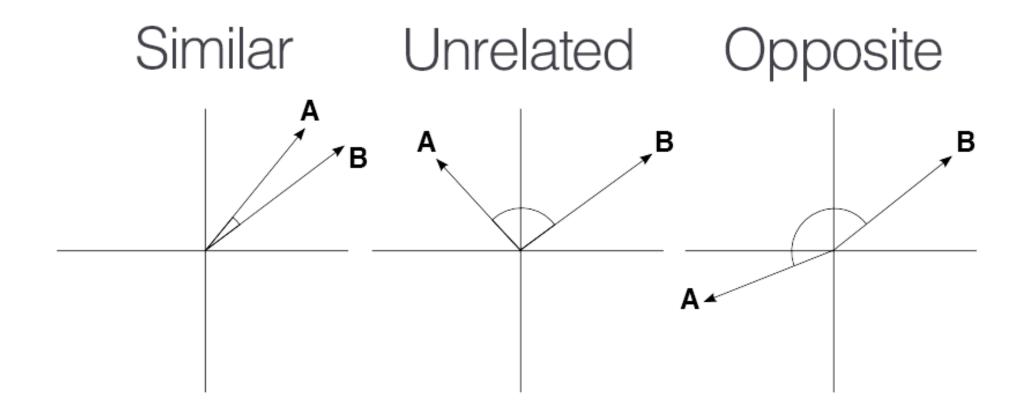
Attribute	Appearances
Minor Key tonality	11
Hard rock roots	9
A subtle use of vocal harmony	9
Mild rhythmic syncopation	8
Extensive vamping	6
A vocal-centric aesthetic	5
Repetitive melodic phrasing	4
Electric rock instrumentation	4





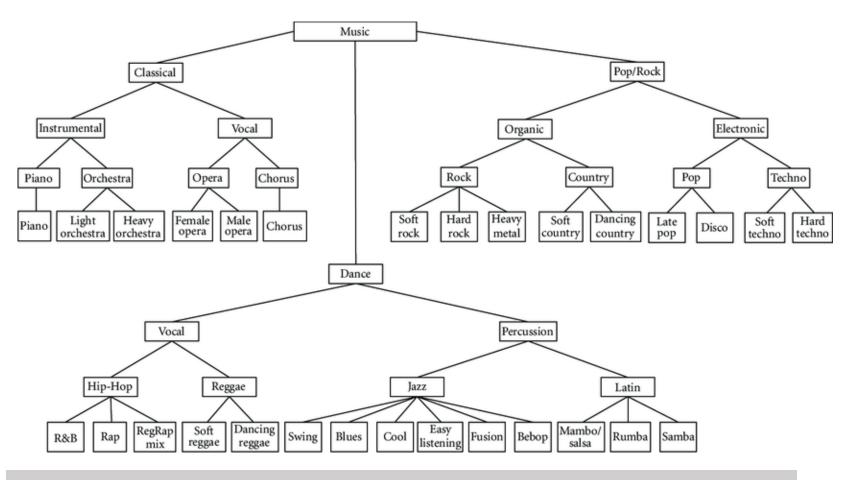
Started with dozens of humans categorizing and labeling tens of thousands of songs ...

Turning features into statistics (aka vectors)



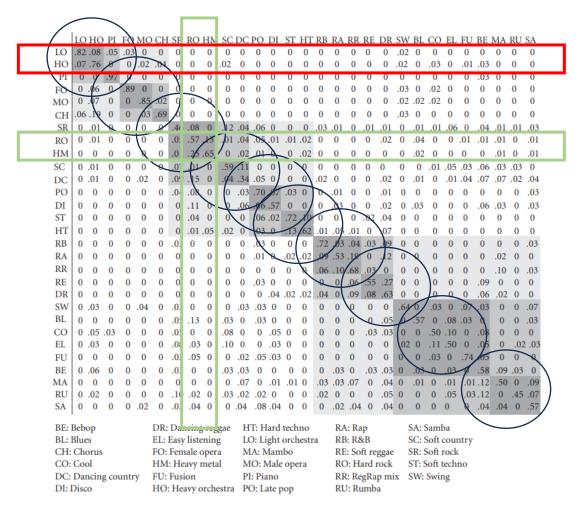
Then the statisticians and engineers created mathematical representations of each label (Machine Learning)

Genre categorization (clusters)



Mathematical categorization of "clusters" of music

The math behind the clusters (similarity)

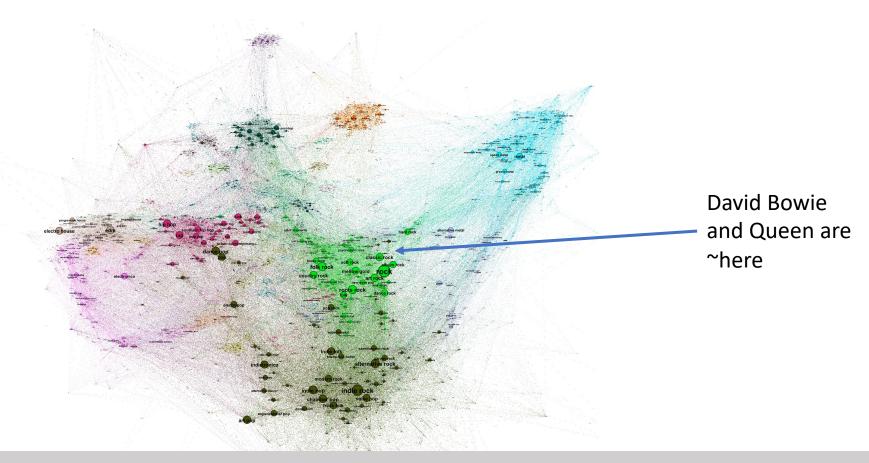


Heavy metal has no similarity to "heavy orchestra"

Hard rock and Heavy metal are quite similar

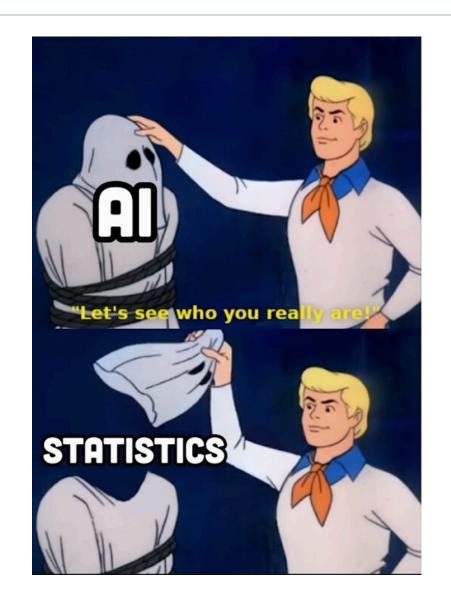
FIGURE 3: Confusion matrix.

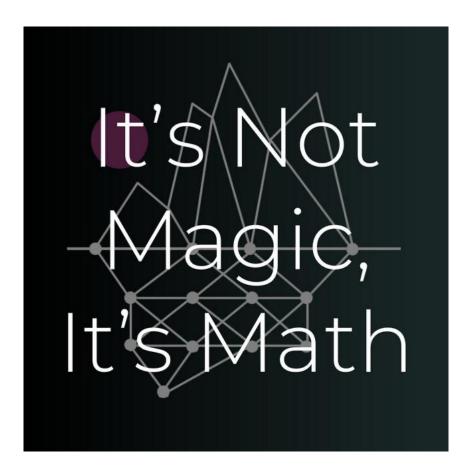
Recommendation engines – ML Labeling -> ML Clustering -> Software



The recommendation engine is software that picks from the local cluster and adjusts to preferences based on thumbs up/down

In other words...





A.I. in music – it all starts with data and labeling

Step 1- Process, label, and link data

Supervised Learning

- Auto label and cluster for the "music genome"
- Generate "stems" –
 pulling out pieces of songs

Step 2- Assist the human in making decisions

Automation (and a little Unsupervised Learning)

- Pandora-style playlists
- Personalized music alerts and AI generated playlists
- Spot new trends
- Predict the next hit

Step 3- Create for the human

Generative Al

- Create original lyrics
- Create original compositions and melodies
- Mimic artists "Fake Drake"

How does Enterprise Al apply to Quality?

Processing, labeling, and linking leads to rocket ships



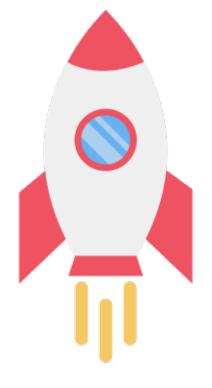




Use A.I.* to gather, process, label, and link data

*(and good old fashioned data engineering, ML, and automation)





A.I. in quality – it all starts with data and labeling

Wave 1- AI based data processing – extract, label, and link data

Supervised Learning (and good data engineering)

- a. Text extraction
- b. Entity and info extraction
- c. Text translation
- d. Text vectorization
- e. Text labeling
- f. Entity resolution
- g. ...

Wave 2- Assist the human in making decisions

Automation (and a little Unsupervised Learning)

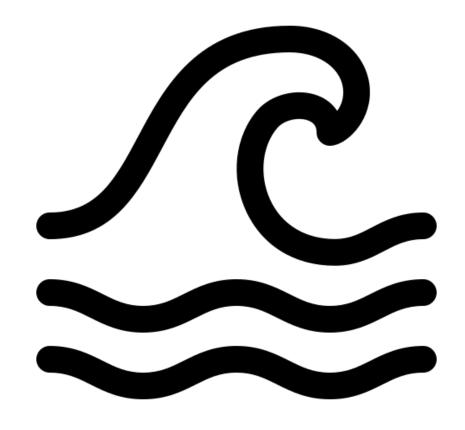
- a. Auto-notifications
- b. Auto-routing
- c. Auto-impact assessments
- d. Auto-diagnostics
- e. Trend detection
- f. Predictive alerts
- g. Natural language search

h.

Wave 3- Create for the human

Generative Al

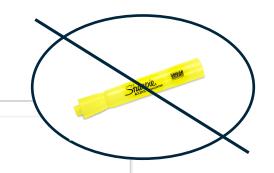
- a. Suggest edits to SOPs
- b. Suggest agency response
- C. ...



Wave 1

Structuring the unstructured:

Al based text extraction (aka IDP)



WARNING LETTER

Episciences Inc.

MARCS-CMS 631902 - AUGUST 25, 2022

Delivery Method:

VIA EMAIL CONFIRMED DELIVERY

Product:

Drugs

Recipient:

Mr. Chad M. Thayer

Chief Operating Officer

Episciences Inc.

10211 West Emerald Street

Boise, ID 83704-8987

United States

$\mathbf{\underline{}}_{(b)(4)}$ (mailto:(b)(4))

Issuing Office:

Division of Pharmaceutical Quality Operations IV

United States

WARNING LETTER

August 25, 2022

Dear Mr. Thayer:

The U.S. Food and Drug Administration inspected your drug manufacturing facility, Episciences Inc., FEI 3006435842, at 10211 West Emerald Street, Boise, from March 7 to 11, 2022.

This warning letter summarizes significant violations of Current Good Manufacturing Practice (CGMP) regulations for finished pharmaceuticals. See Title 21 Code of Federal Regulations (CFR), parts 210 and 211 (21 CFR parts 210 and 211).



Word-level segmentation and OCR

WARNING LETTER

Episciences Inc.

MARCS-CMS 631902 - AUGUST 25, 2022

elivery Metho

VIA EMAIL CONFIRMED DELIVERY

Product:

Drugs

Recipient

Mr. Chad M. Thayer

Chief Operating Officer

Episciences Inc.

10211 West Emerald Street

5013E, ID 05704-090

United States

<u>(b)(4) (mailto:(b)(4))</u>

Issuing Office:

Division of Pharmaceutical Quality Operations IV

United States

WARNING LETTER

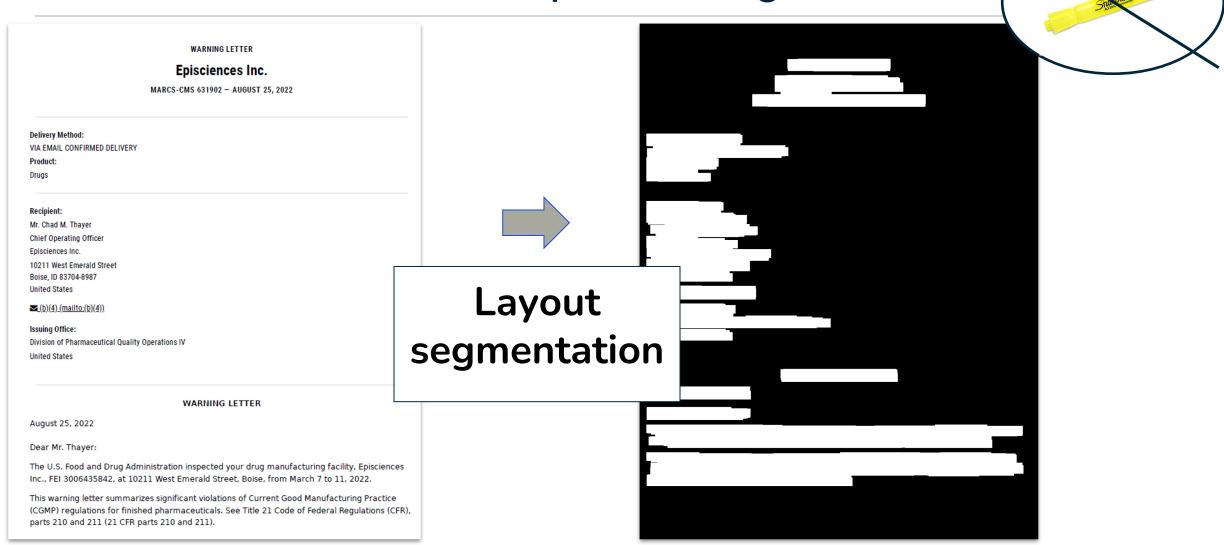
August 25, 2022

Dear Mr. Thayer:

The U.S. Food and Drug Administration inspected your drug manufacturing facility, Episciences Inc., FEI 3006435842, at 10211 West Emerald Street, Boise, from March 7 to 11, 2022.

This warning letter summarizes significant violations of Current Good Manufacturing Practice (CGMP) regulations for finished pharmaceuticals. See Title 21 Code of Federal Regulations (CFR), parts 210 and 211 (21 CFR parts 210 and 211).









Because your methods, facilities, or controls for manufacturing, processing, packing, or holding do not conform to CGMP, your drug products are adulterated within the meaning of section 501(a)(2)(B) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), 21 U.S.C. 351(a)(2)(B).

We reviewed your March 17, 2022, response to our Form FDA 483 in detail.

During our inspection, our investigator observed specific violations including, but not limited to, the following.

 Your firm failed to thoroughly investigate any unexplained discrepancy or failure of a batch or any of its components to meet any of its specifications, whether the batch has already been distributed (21 CFR 211.192).

Your investigations into out-of-specification (OQS) laboratory results are inadequate because they do not include scientifically supported gonclusions and lack corrective actions and preventive actions (CAPA). For example:

A. Your investigation into bulk drug product assay OOS result of active ingredient zinc oxide reported by your contract testing laboratory for "(b)(4)" Sunscreen, Lot # (b)(4), was inadequate. You determined the root cause to be sampling error and attributed it to a non-homogenous blending sample collected from the "top of the mixer scrapings." The bulk batch was subsequently resampled, retested for assay, and released after obtaining passing results. Your investigation lacked hypothesis testing and adequate evidence to identify the root

and you did not document whether other batehes and drug products manufactured in y facility were affected. In addition, your investigation failed to address the non-homogen your products based on where test samples are collected.

B. Your investigation into OOS bulk test results for appearance and viscosity for "Purifyit Lot #\$ 21A044, 21A045, 21A046, and 21A085, reported by your contract testing laboral inadequate. You failed to adequately determine the root cause for these failures. Instea inspected sample tubes from each lot for appearance, retested only one of the four lot assay and subsequently released all four lots on the basis of this limited testing.

You failed to identify appropriate CAPA for these investigations to prevent recurrence or events. There is no assurance that all batches produced under inadequate conditions he thoroughly evaluated, and that your firm has identified all significant variables associate your manufacturing process.

Moreover, your standard operating procedure (SOP) titled "Out of Specification Investigations," states that batches may be released with OOS results at the discretion of the Quality Assurance (OA) manager.

For more information about handling failing, OOS, out-of-trend, or other unexpected results and documentation of your investigations, see FDA's guidance document, *Investigating Out-of-Specification (OOS)* Test Results-for Pharmaceutical Production at: https://www.fda.gov/media/158416/download/. https://www.fda.gov/media/158416/download/.

page_num segNum

	2	10	Because your methods, facilities, or controls for manufacturing, processing, packing, or holding do not conform to CGMP, your drug products are adulterated within the meaning of section 501(a)(2)(B) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), 21 U.S.C. 351(a)(2)(B).
	2	9	We reviewed your March 17, 2022, response to our Form FDA 483 in detail.
	2		During our inspection, our investigator observed specific violations including, but not limited to, the following.
	2	7	1. Your firm failed to thoroughly investigate any unexplained discrepancy or failure of a batch or any of its components to meet any of its specifications, whether the batch has already been distributed (21 CFR 211.192).
	2	6	Your investigations into out-or-specification (OOS) laboratory results are inadequate because they do not include scientifically supported conclusions and lack corrective actions and preventive actions (CAPA). For example:
			A. Your investigation into bulk drug product assay OOS result of active ingredient zinc oxide reported by your contract testing laboratory for "(b)(4)"

Segment Extraction

other batches and drug products manufactured in your facility were affected. In addition, your investigation failed to address the non-homogeneity of your products based on where test samples are collected.

B. Your investigation into OOS bulk test results for appearance and viscosity for "Purifying Wash, Lot #s 21A044, 21A045, 21A046, and 21A085, reported by your contract testing laboratory was inadequate. You failed to adequately determine the root cause for these failures. Instead, you inspected sample tubes from each lot for appearance, retested only one of the four lots for assay and subsequently released all four lots on the basis of this limited testing.

Sunscreen, Lot # (b)(4), was inadequate. You determined the root cause to be sampling error and attributed it to a non-homogenous blending

sample collected from the "top of the mixer scrapings. The bulk batch was subsequently resampled, retested for assay, and released after obtaining

passing results. Your investigation lacked hypothesis testing and adequate evidence to identify the root cause and you did not document whether

Line Text

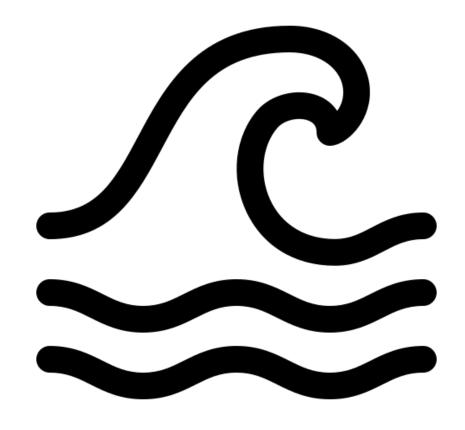
- You failed to identify appropriate CAPA for these investigations to prevent recurrence of such events. There is no assurance that all batches produced under inadequate conditions have been thoroughly evaluated, and that your firm has identified all significant variables associated with your manufacturing process.
- Moreover, your standard operating procedure (SOP) titled "Out of Specification Investigations, states that batches may be released with OOS results at the discretion of the Quality Assurance (QA) manager.
- For more information about handling failing, OOS, out-of-trend, or other unexpected results and documentation of your investigations, see FDA's guidance document, Investigating Out-of- Specification (OOS) Test Results for Pharmaceutical Production at: https://www.fda.gov/media/158416/download(https://www.fda.gov/media/158416/download)





segNum	relavence Tag	relavenceTag_prob	LineText	page_num
10	introduction	0.98	ecause your methods, facilities, or controls for manufacturing, processing, packing, or holding do not conform to CGMP, your drug products are adulterated vithin the meaning of section 501(a)(2)(B) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), 21 U.S.C. 351(a)(2)(B).	
9	introduction	0.98	Ne reviewed your March 17, 2022, response to our Form FDA 483 in detail.	
8	introduction	0.98	During our inspection, our investigator observed specific violations including, but not limited to, the following.	2
7	deficiencytitle	0.99	1. Your firm failed to thoroughly investigate any unexplained discrepancy or failure of a batch or any of its components to meet any of its specifications, whether he batch has already been distributed (21 CFR 211.192).	
6	deficiency	0.98	Your invest they do not include scientifically supported conclusions and lack corrective a	2
5	deficiency	0.98	A. Your involute ## A. Your involute ## Segment Lot # (b)(4) "top of the lacked hype in your facil Categorization Segment eported by your contract testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscreen, ed it to a non- homogenous blending sample collected from the particular testing laboratory for "(b)(4)" Sunscr	2
4	deficiency	0.99	B. Your investigation into OOS bulk test results for appearance and viscosity for "Purifying Wash, Lot #s 21A044, 21A045, 21A046, and 21A085, reported by your contract testing laboratory was inadequate. You failed to adequately determine the root cause for these failures. Instead, you inspected sample tubes from each lot for appearance, retested only one of the four lots for assay and subsequently released all four lots on the basis of this limited testing.	2
3	deficiency	0.96	You failed to identify appropriate CAPA for these investigations to prevent recurrence of such events. There is no assurance that all batches produced under inadequate conditions have been thoroughly evaluated, and that your firm has identified all significant variables associated with your manufacturing process.	
2	deficiency	0.98	Moreover, your standard operating procedure (SOP) titled "Out of Specification Investigations, states that batches may be released with OOS results at the discretion of the Quality Assurance (QA) manager.	2
1	feedback,fdarequestforinformation	0.44,0.32	For more information about handling failing, OOS, out-of-trend, or other unexpected results and documentation of your investigations, see FDA's guidance document, Investigating Out-of- Specification (OOS) Test Results for Pharmaceutical Production at: https://www.fda.gov/media/158416/downlownload(https://www.fda.gov/media/158416/download)	2





Wave 1

Structuring the unstructured:

Understanding meaning (NLP)

Understanding meaning – how do we do it?

segNum	rel	avenceTag re	elavenceTag_prob	LineText		page_num
	introduction	-		Because your methods, facilities, or controls for manufacturing, processing, packing, or holding do not conform to CGMP, your drug products are adulterated within the meaning of section 501(a)(2)(B) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), 21 U.S.C. 351(a)(2)(B).		2
9	introduction	on 0.98		We reviewed your March 17, 2022, response to our Form FDA 483 in detail.		2
8	introduction	ction 0.98 During our inspection, our investigator observed specific violations including, but not limited to, the following.			2	
7	deficiencytitle	1. Your firm failed to thoroughly investigate any unexplained discrepancy or failure of a batch or any of its components to meet any of its specification the batch has already been distributed (21 CFR 211.192).		ions, whether	2	
6	deficiency	Your investigations into out-or-specification (OOS) laboratory results are inadequate because they do not include scientifically supported conclusions corrective actions and preventive actions (CAPA). For example:			ons and lack	2
5	deficiency	Text:1. Your firm failed to thoroughly investigate any unexplained discrepancy or failure of a batch or any of its components to meet any of its specifications, whether the batch has already been distributed (21 CFR 211.192).				2
4	deficiency	L1/L2: Reviews and Approvals->Batch Release, 0.74 Common Object: Specifications; 0.89			ported by tubes from	2
3	deficiency				ed under process.	2
2	deficiency	batch or any of its components to meet any of its specifications				2
1	feedback,fda	Summary: The CFR Title, Part, and Section that corresponds to the violation mentioned in the text snippet is 21 CFR 211.192.			uidance	2

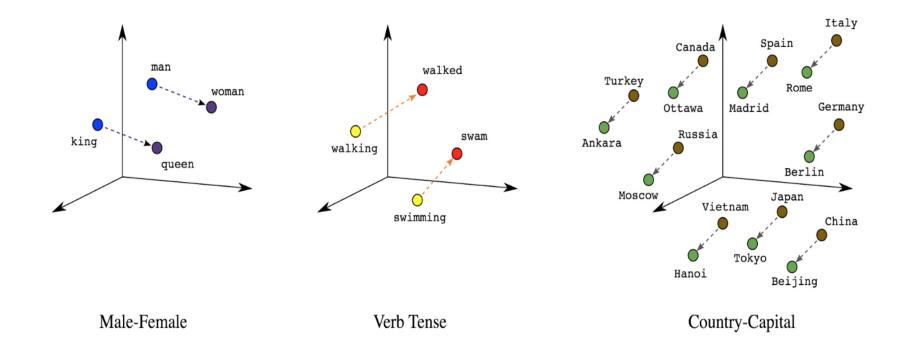


Understanding meaning: vectors

Quiz: What is the **Berlin**of **France**? (fill in the blank)

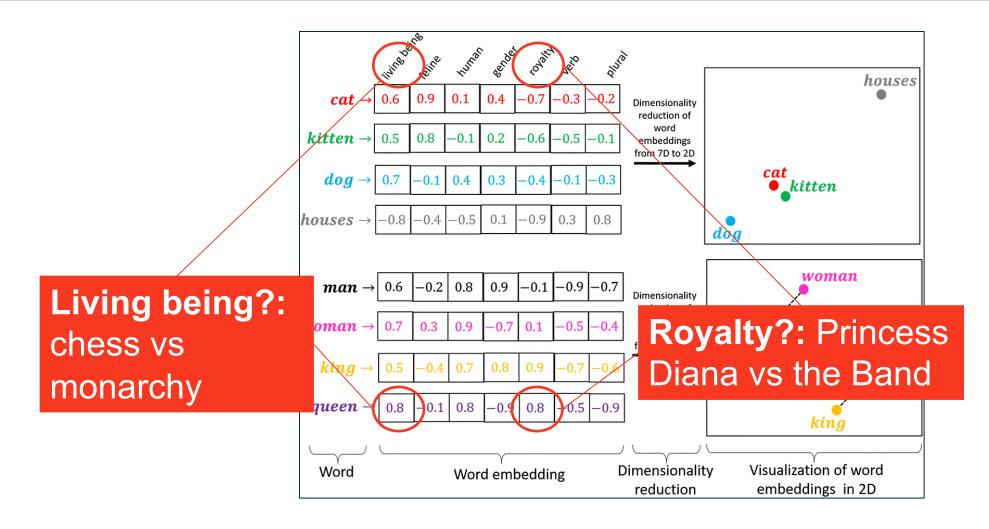


Understanding meaning: many types of vectors





Understanding meaning: embeddings





1000's of examples to "teach" the model





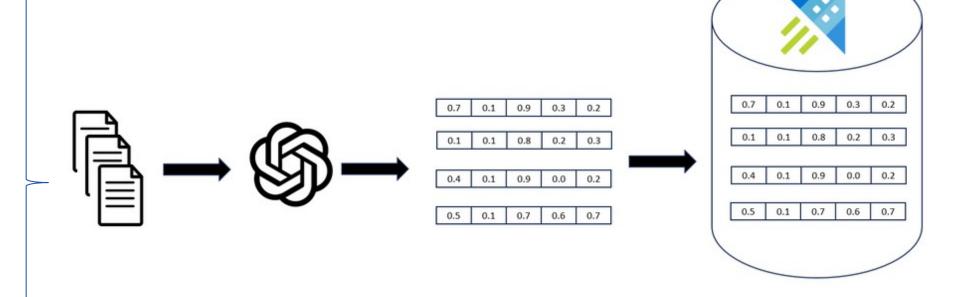








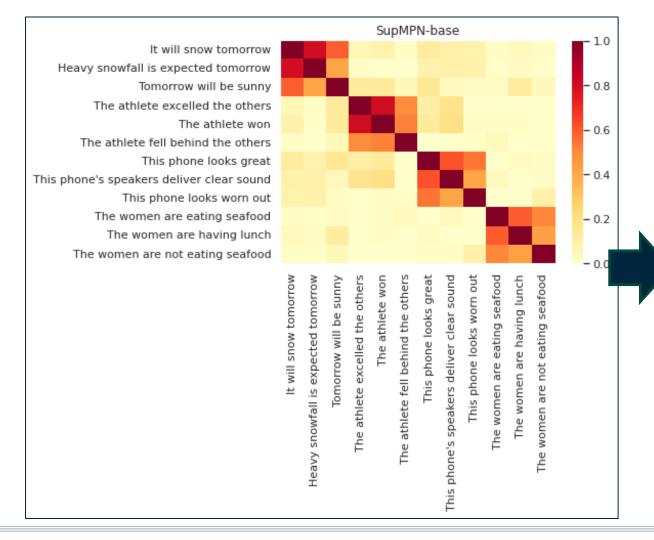




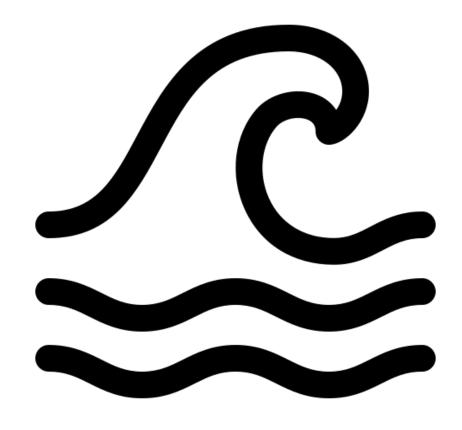
Text Embeddings

Vector Database

Sentence similarity – Just like music!



Hint: this is the underlying technology that powers natural language search

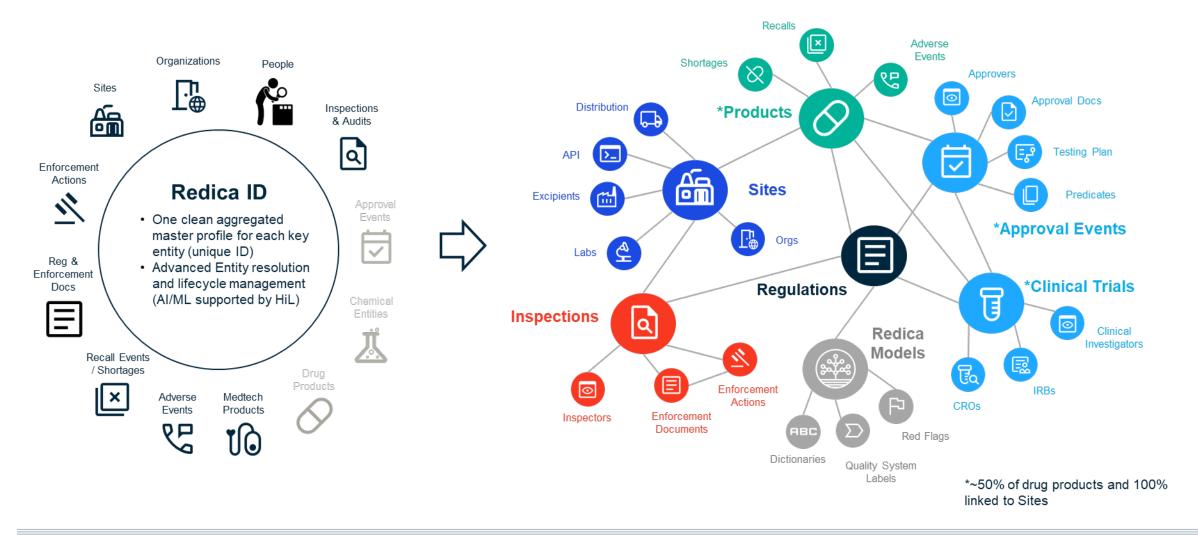


Wave 1

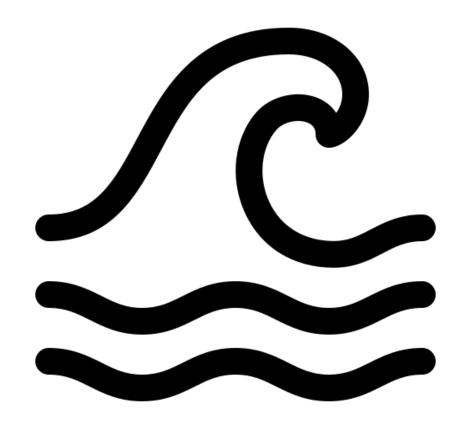
Structuring the unstructured:

Entity extraction and linking

Entity creation and linking







Amplify the intelligence of humans:

Wave 2



Al powered processing - simple visualizations





Al powered processing - simple visualizations



Al powered processing - simple visualizations



My dream - Investigation of an issue





Here is an issue of interest ...

1) How widespread is this? 2) Do we have this problem?"

1) "Looking at enformalis, adverse events – the issue firmultiple companies have been competitors"

note this issue. Here is the highlighted evidence and my confidence intervals for each item..."

"Would you like alerts on this topic as it evolves? check all that apply a) Guidance b) proposed regulations c) enforcement d) recalls e) AE's f) deviations ...

Click ->

Click ->

Artificial Intelligence Augmented