



Office of Regulation Management  
Office of Government-wide Policy  
U.S. General Services Administration



# Regulatory Data

April 30, 2020  
2:00 pm  
**U.S. General Services Administration**

# Welcome

- The meeting will start shortly
- All attendees are muted. You will hear audio as soon as the meeting starts
- Pose questions for the audience through the chat function
- If possible, please use your computer and not the telephone line
- Closed Captioning is provided in the window directly below the screen share. It is also available at  
<https://www.streamtext.net/player?event=NRGCO>.
- Please remember to send suggestions on modernizing regulatory services using [Regulations.gov](#)  
Go to: <https://www.regulations.gov/document?D=GSA-GSA-2020-0002-0012>  
or search for “Office of Regulation Management” in [Regulations.gov](#)

# Introductions



**Virginia Huth**  
Deputy Associate Administrator  
Office of Regulation Management  
General Services Administration

# Opening Remarks



Dominic Mancini is the Deputy Administrator of the Office of Information and Regulatory Affairs at OMB. An economist by training with a PhD from the University of North Carolina at Chapel Hill, he began his Federal career at the Food and Drug Administration.

He has spent the majority of his 20 year career in public service at the Office of Information and Regulatory Affairs.

## Dominic Mancini

# Opening Remarks



**Jessica Salmoiraghi**

Jessica Salmoiraghi is the Associate Administrator for the Office of Government-wide Policy at GSA. In addition to oversight for the Office of Regulation Management, she oversees seven other offices that provide government-wide policy and guidance for programs such as real and personal property, travel, regulatory policy development, government-wide programmatic support and information technology.

Ms. Salmoiraghi also serves as the Agency's Chief Acquisition Officer.

Prior to her public service at GSA, she served in various leadership positions in architecture, engineering and construction organizations focused on national and international procurement.

# Logistics

- Questions from the audience
- Accessing a recording of today's meeting
- Submitting comments

# Panelist Introductions – Topic 1

## Jarilyn Dupont

Jarilyn Dupont is the Director of Regulatory Policy within the U.S. Food and Drug Administration's Office of Policy, Legislation and International Affairs.

Ms. Dupont has been with the FDA since 1996, including positions with FDA's Europe Office, Office of Legislation, Center for Drug Evaluation and Research and the Office of Crisis Management.

Prior to her work with the FDA, she was Counsel for the U.S. House Judiciary Committee. She also has legal experience in private practice, state and local legislative and administrative agencies, and non-profit legal services organizations.

# Panelist Introductions – Topic 1



**Tom Sabo**

Tom Sabo is a principal solutions architect at SAS.

Since 2005, he has been immersed in the field of text analytics and Artificial Intelligence as it applies to federal government challenges.

Mr. Sabo presents work internationally on diverse topics including modeling applied to government procurement, counter human trafficking, and using analytics to leverage and predict research trends.

He holds a bachelor's degree in cognitive science and a master's in computer science, both from the University of Virginia.

# Panelist Introductions – Topic 2



**Carey Johnston**

Carey Johnston is an Environmental Engineer with the U.S. Environmental Protection Agency's Office of Policy.

Mr. Johnson started with the U.S. EPA in 1995. His work includes a rulemaking that will require certain permittees and regulators to electronically report information in lieu of filing written reports.

He has an undergraduate degree in Mechanical Engineering from the University of Virginia and a graduate degree in Environmental Engineering from Virginia Tech. He is also a licensed Professional Engineer in the Commonwealth of Virginia.

# Panelist Introductions – Topic 2



**Michelle Cosby**

Michelle Cosby is president of the American Association of Law Libraries, the only national association dedicated to the legal information profession and its professionals.

Ms. Cosby is currently the director of the law library at Temple University, where she leads the law library in the support of research for law faculty and students.

She received her B.A. in Sociology from Butler University and earned a dual master of library science degree and juris doctorate degree at Indiana University Bloomington.

# Panelist Introductions – Topic 3



**Elizabeth Kowalewski**

Elizabeth Kowalewski is a Senior Analyst at the U.S. Government Accountability Office.

Since 2016, she has served as an Analyst-in-Charge in GAO's Forensic Audits and Investigative Service team.

Ms. Kowalewski received her Master of Public Affairs in Sustainable Development from Indiana University's School of Public and Environmental Affairs in 2008, and BA in Environmental Studies and Politics from Ursinus College in 2006.

## Panel 2



**Hudson Hollister**

Hudson Hollister is the Founder and Principal at HData.

He is exploring opportunities to make regulations machine-readable and interoperable through the adoption of an open data structure.

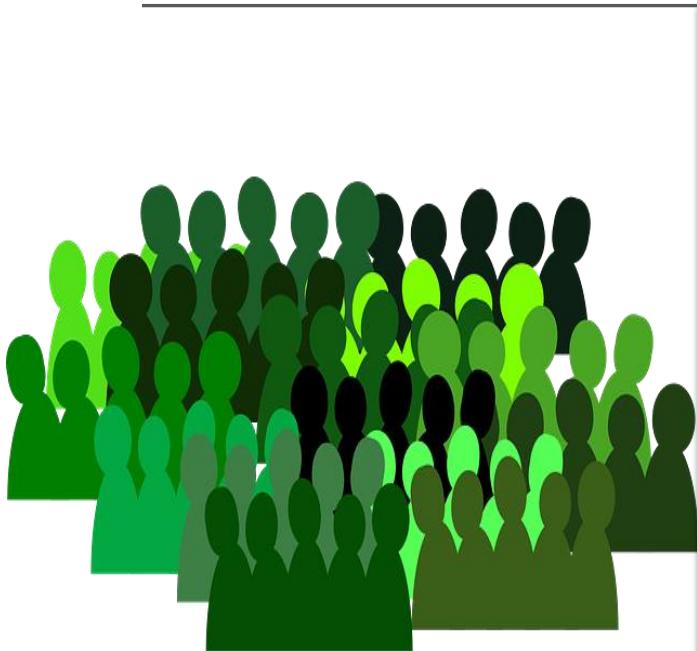
Mr. Hollister previously founded the Data Coalition, which advocates for open data reforms in Congress, including the DATA Act of 2014, the OPEN Government Data Act and GREAT Act of 2019, and the currently-pending Financial Transparency Act.

# Topic 1

## Data Analysis in the Management of Mass and Fake Comments

- **Jarilyn Dupont, J.D.**: Director of Regulatory Policy, U.S. Food and Drug Administration's Office of Policy, Legislation, and International Affairs
- **Tom Sabo**: Principal Solutions Architect at SAS Institute, Inc.

# FDA Interactions with External Participants In Rulemaking Process



- ⌚ **Comments to the dockets on rulemaking**
- ⌚ Public Meetings and Hearings
- ⌚ Requests for Data/Information
- ⌚ Advisory Committee Meetings
- ⌚ Focus Groups/Surveys
- ⌚ Citizen petitions

# FDA Rulemaking and Comment Review

- FDA does not consider the number of comments in determining a particular issue – the final rule is not determined by “popular vote.”
  - FDA reviews the comments, data, and research submitted to the docket as it considers appropriate next steps.
  - The Agency makes decisions on the provisions for a final rule based on sound reasoning, data, and scientific evidence.
  - FDA’s review is in accordance with the Administrative Procedure Act and court decision requirements that relevant and substantive comments are considered and responded to in the final rule.
- 
- FDA’s Docket Office has software tools to help identify duplicate comments with a high degree of confidence.
  - FDA Centers/Offices organize and review comments typically by subject matter and issues raised in the comment. Subject matter experts participate in developing responses.
  - In some instances we have used text analytics analysis to break down similar comments by content and theme.
  - The Agency can identify likely BOT comments.
  - Outside contractors are used in some instances to categorize and identify similar comment submissions.

# FDA Selected Dockets – Comment Submissions

Docket ID	Total comments	FDA Center	Docket Title
<b>FDA-2017-N-6565</b>	525,316	CTP	Regulation of Flavors in Tobacco Products
<b>FDA-1995-N-0259</b>	99,120	OC	Regulations Restricting the Sale & Distribution of Cigarettes & Smokeless Tobacco to Protect Children & Adolescents
<b>FDA-2014-N-0189</b>	79,668	CTP	Deeming Tobacco Products To Be Subject to the Food, Drug, and Cosmetic Act, as Amended by the Family Smoking Prevention and Tobacco Control Act; Regulations on the Sale and Distribution of Tobacco Products and Required Warning Statements for Tobacco Products
<b>FDA-1977-N-0220</b>	68,942	OC	Saccharin & Its Salts*
<b>FDA-2017-D-6580</b>	53,672	CDER	Drug Products Labeled as Homeopathic
<b>FDA-2018-P-2962</b>	16,587	CDER	Request that the FDA issue a declaratory order finding that Petitioners have standing, and that the FDA, under the FDCA, establish an expert advisory committee in accordance with FDA policy on Advisory Committees, Licensed and/or properly certified homeopathic, adopt the long standing effective guidance existing in Compliance Policy Guidance (CPG) 400.400 (with minor revisions) as a regulation
<b>FDA-2019-D-0661</b>	15,414	CDER	Modifications to Compliance Policy for Certain Deemed Tobacco Products; Draft Guidance for Industry; Availability
<b>FDA-2018-N-3522</b>	11,908	CFSAN	Use of Dairy Terms in the Labeling of Plant-Based Products
<b>FDA-2018-N-3685</b>	10,664	CDER	International Drug Scheduling; Convention on Psychotropic Substances; Single Convention on Narcotic Drugs; ADB-FUBINACA; ADB-CHMINACA; Cyclopropyl Fentanyl; Methoxyacetyl Fentanyl; para-Fluoro Butyrfentanyl; Tramadol; Pregabalin; Cannabis Plant and Resin; and 8 additional substances; Request for Comments
<b>FDA-2015-N-0540</b>	9,451	CDER	Homeopathic Product Regulation: Evaluating FDA's Regulatory Framework After a Quarter-Century

\* Older Docket – Not available in Regulations.gov

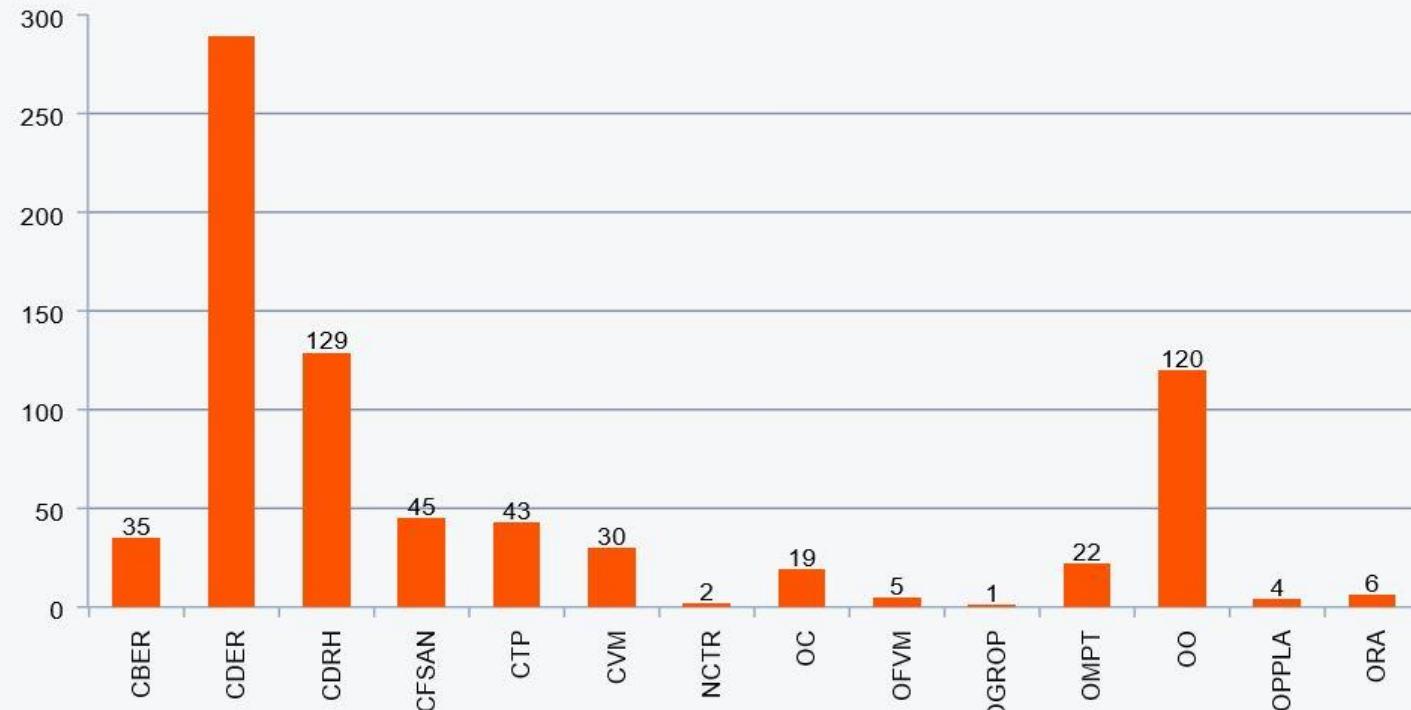
# 2019 Federal Register Documents

Jan 1, 2019 

End Date Published 

Jan 1, 2020 

**Published FRDTS Records Count by Center**



# 2019 Federal Register Documents

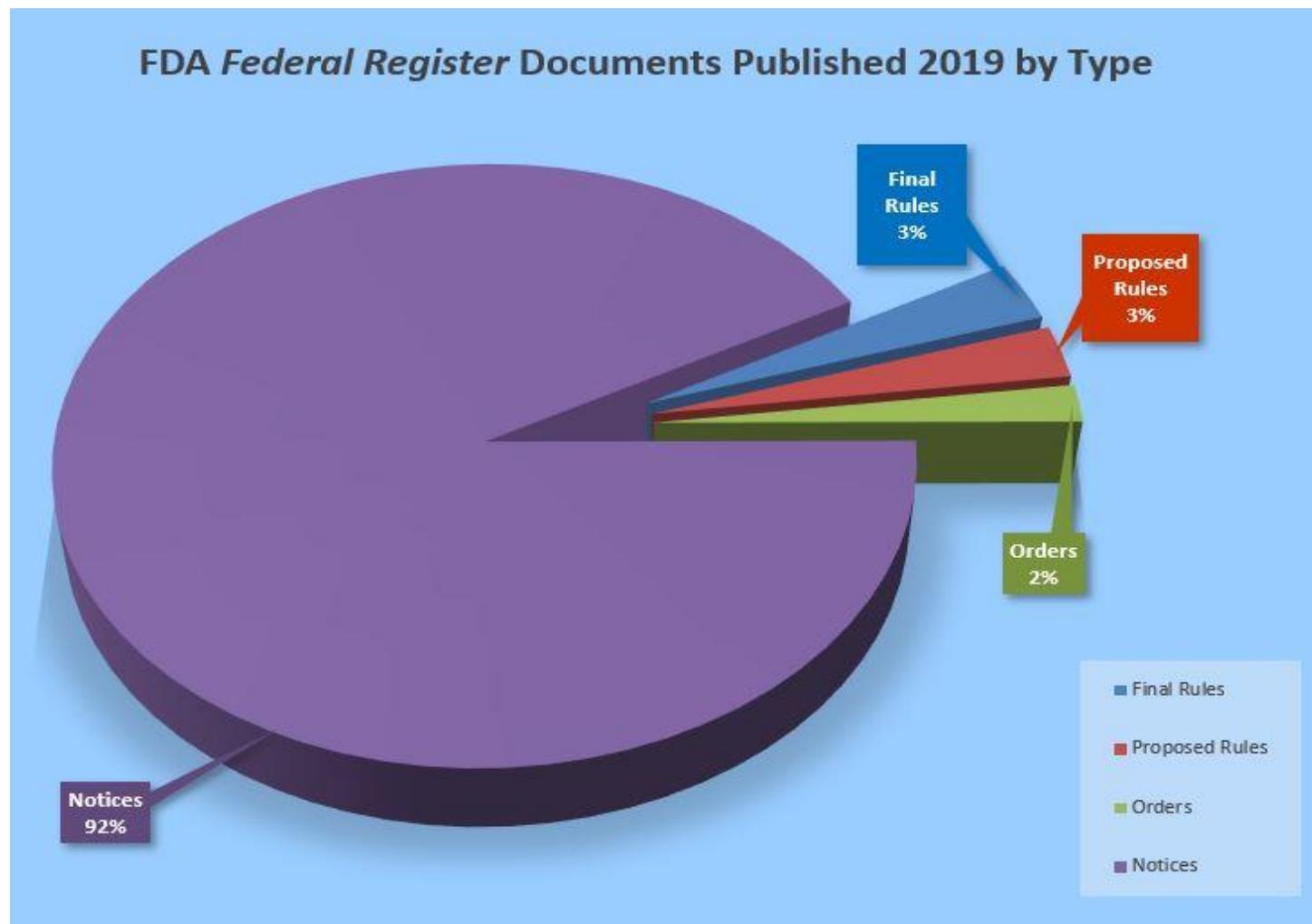
## FDA Regulations and Other *Federal Register* Documents Published 2019

	CBER	CDER	CDRH	CFSAN	CTP	CVM	NCTR	OC	OFVM	OMPT	OO	OPPLA	ORA & OGROP	Total
Final Rules*	1	4	5	6	0	7	0	0	1	0	0	0	0	24
Proposed Rules	2	6	2	3	6	0	0	0	1	1	0	0	0	21
ANPRMs:	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Orders	0	0	14	0	0	0	0	0	0	0	0	1	0	15
Notices**	32	279	108	36	37	23	2	19	3	21	120	3	7	690
Number of Published Federal Register Documents:	35	289	129	45	43	30	2	19	5	22	120	4	7	750

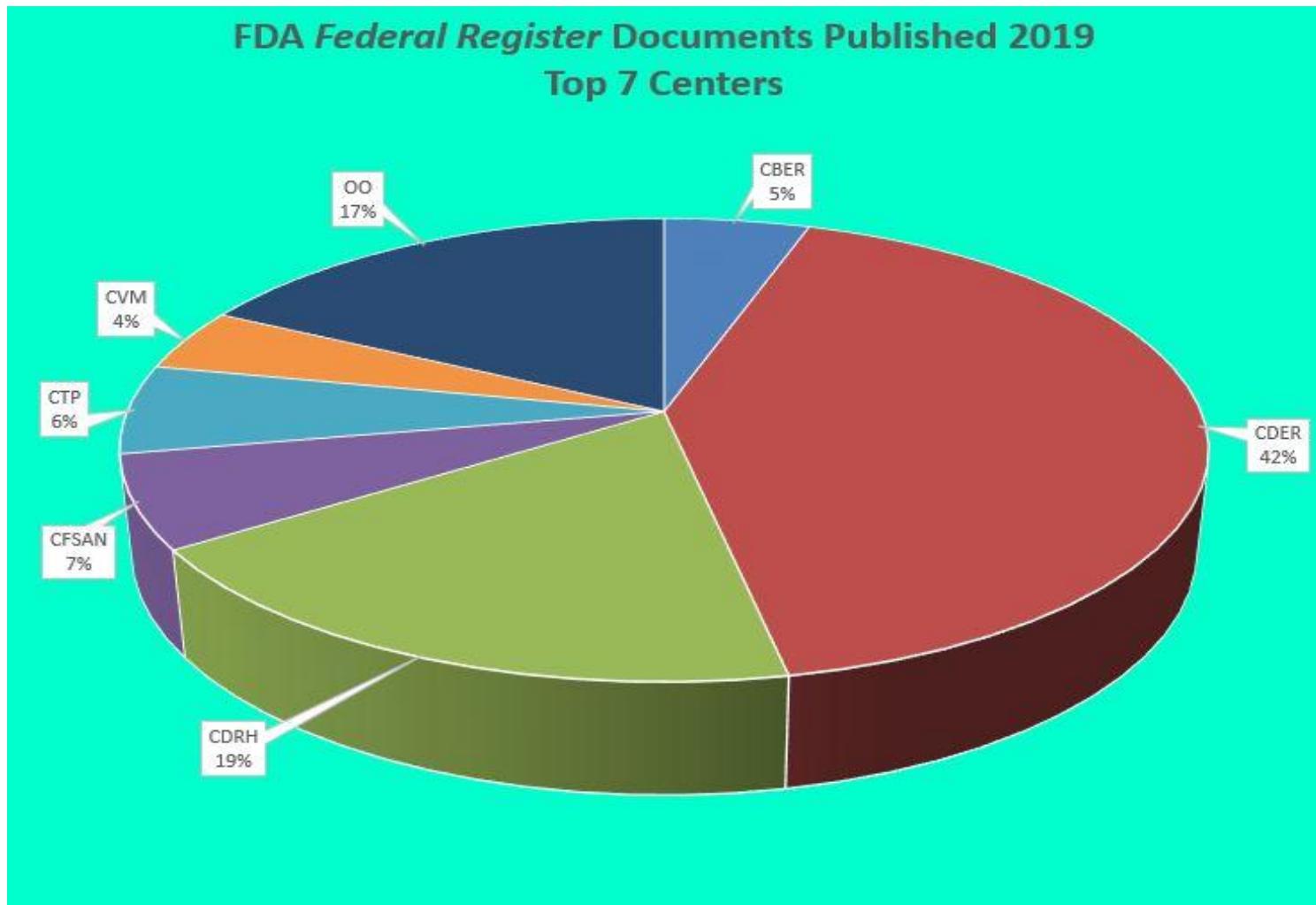
\* Includes New Animal Drug Application approvals, technical amendments, etc.

\*\* Includes all types of FR Notices (i.e., advisory committee meeting notices, public workshops, paperwork notices, notices of availability, etc...)

# 2019 Regulations and Notices



# 2019 Regulations and Notices





# Hearing Every Voice: Text Analytics for Regulations.gov Public Commentary





# Regulations.gov

FDA-2017-N-6565



## Regulation of Flavor in Tobacco Products

'....is seeking comments, data, research results, or other information about, among other things, **how flavors attract youth to initiate tobacco product use** and about whether and how certain flavors may help adult cigarette smokers reduce **cigarette use** and switch to potentially less harmful products...'



# Goals of Analysis

- **What are the main themes/concerns mentioned in comments?**
- **Who are the commenters?**
- **Can we identify form letters?**

The screenshot shows the regulations.gov homepage with a specific docket page for "Regulation of Flavors in Tobacco Products".

**Header:** regulations.gov - Your Voice in Federal Decision-Making

**Navigation:** Home | Help ▾ | Resources ▾ | Contact | Advanced Search

**Docket Information:**

- Docket ID: FDA-2017-N-6565
- Agency: Food and Drug Administration (FDA)
- Parent Agency: Department of Health and Human Services (HHS)
- + View More Docket Details
- Take a Tour!

**Primary Documents:** View All (26)

**Document Summary:**

- Regulation of Flavors in Tobacco Products
- Proposed Rule
- Posted: 03/21/2018
- ID: FDA-2017-N-6565-0001

**Comments:**

- Comment Period Closed Jul 19, 2018 11:59 PM ET
- 525,438 Comments Received\*
- Sign up for Email Alerts
- Tweet | Share | Email

**GSA Logo:**

# Challenge of Manual Analysis

 Your Voice in Federal Decision-Making

Home Help ▾ Resources ▾ Contact Advanced Search

## Regulation of Flavors in Tobacco Products

Docket Folder Summary [View all documents and comments in this Docket](#)

Docket ID: FDA-2017-N-6565 Agency: Food and Drug Administration (FDA) Parent Agency: Department of Health and Human Services (HHS)

+ View More Docket Details [Take a Tour!](#)

Primary Documents [View All \(26\)](#)

 Regulation of Flavors in Tobacco Products

Proposed Rule Posted: 03/21/2018 ID: FDA-2017-N-6565-0001

Comment Period Closed Jul 19, 2018 11:59 PM ET

 Sign up for Email Alerts

525,438 Comments Received\*

[Tweet](#) [Share](#) [Email](#)

# Challenge of Manual Analysis

**regulations.gov**  
Your Voice in Federal Decision-Making

Home Help ▾ Resources ▾ Contact Advanced Search

Regulation of Flavors in Tobacco Products

Docket Folder Summary View all documents and comments in this Docket

Docket ID: FDA-2017-N-6565 Agency: Food and Drug Administration (FDA) Parent Agency: Department of Health and Human Services (HHS)

+ View More Docket Details

Primary Documents [View All \(26\)](#)

PR Regulation of Flavors in Tobacco Products

Proposed Rule Posted: 03/21/2018 ID: FDA-2017-N-6565-0001

Comment Period Closed Jul 19, 2018 11:59 PM ET

Sign up for Email Alerts

**525,438** Comments Received

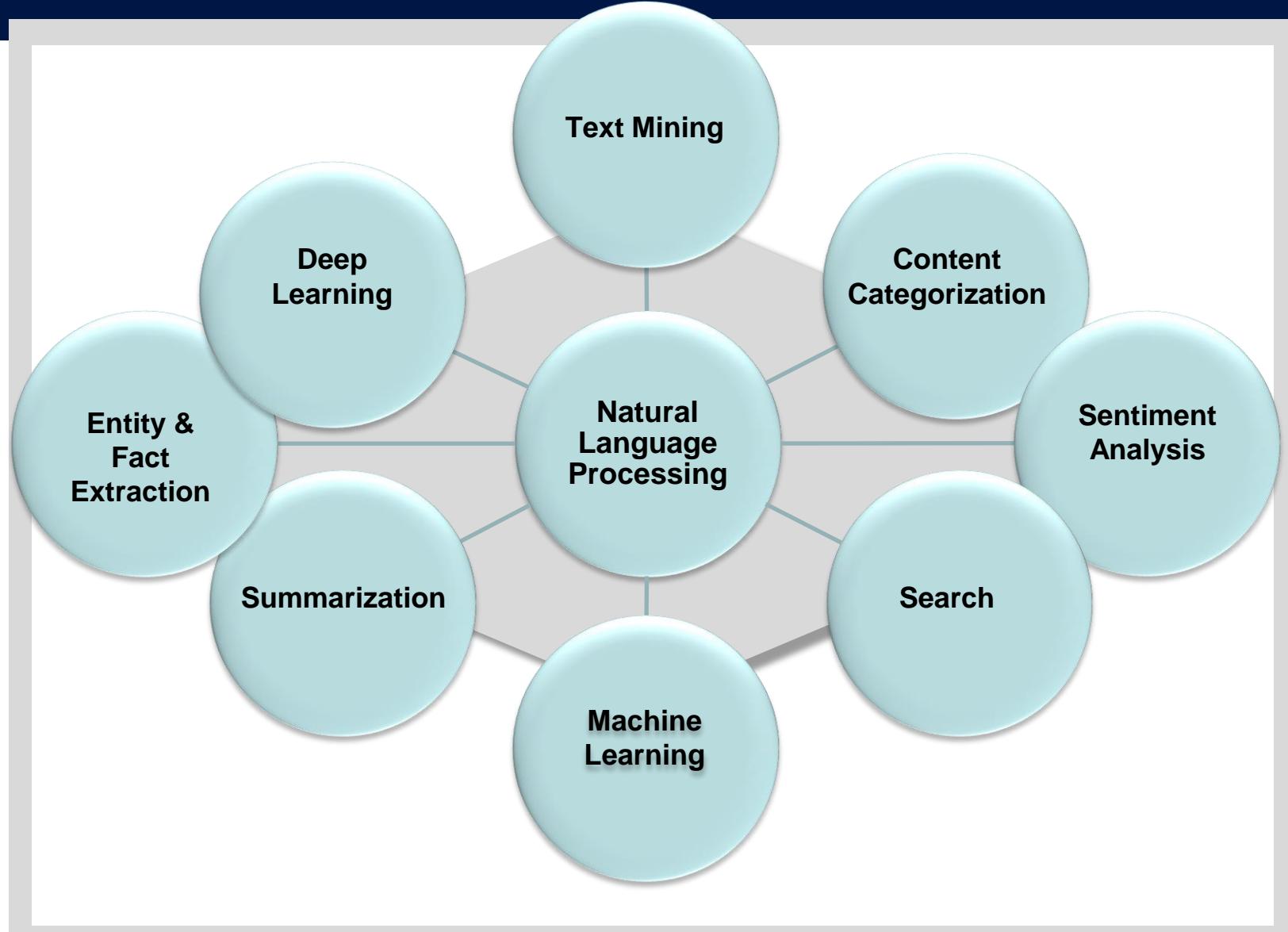
Tweet Share Email

# What is Text Analytics?

***“Using technology to scale the human acts of reading, organizing, and quantifying freeform text in meaningful ways.”***



# Text Analytics/AI Capabilities



# Text Analytics

## API use and Data Preprocessing

## Comment

At the age of 18 I had been smoking a pack a day and sometimes 2 packs a day on the weekend for 2 years with chewing tobacco being mixed in intermittently. I smoked Marlboro Black 100's and chewed Grizzly Straight. Very harsh, dark flavored products. When I quit using tobacco I hadn't really planned on it because I enjoyed the taste of cigarettes and dip. A close friend introduced me to vaping after he quit smoking using a Hawaiian punch flavored liquid. He previously smoked Newport Reds and Marlboro Reds. I wasn't very interested at first but he later showed me two different flavored eliquids: one being milk and honey flavored while the other was ~~was number one and mint flavored. At that mint brought~~ device and quit smoking. Since then I have been tobacco free for about 3 years now. I prefer eliquid opposed to a tobacco or nicotine product. I believe that those in the age range of 40-70 years old when they initially think it is too weird for them to try it is a tactic for targeting youth.

	document	sid	sentences
1		1	1 You have no right as adult use 21+ in my state California to take away our flavors.
2		1	2 Out lawing flavors will do nothing but make us the "adult" consumers angry but "children" under 18 will find a way to get there hands on tobacco products regardless of "flavor regulation" wheather it be cigarettes or e-liquid you just can't win so be realistic and let us adults have our flavors.
3		1	3 These flavors has helped so many people quit smoking, I for one vape non tobacco flavors so why should my rights be cut?
4		1	4 secondly, I hate tobacco cigarettes the smell irritates my lungs when someone decides to light up around me but the e-liquid flavors don't bother with my lungs when someone is vaping around me unlike cigarettes.
5		1	5 I rather smell chemy flavor then the traditional second hand cancer stick.
6		1	6 We want our flavors FDA!
7		2	1 Flavored e liquids have help many people quit smoking
8		3	1 I used to smoke cigarettes for years, and after I switch to vaping with flavors it has helped me stay away from them for quite a while.
9		3	2 I believe if the flavor is taken away, it will make more people like me a lot harder to have other alternatives.
10		4	1 I do not smoke and I do not vape, but I have friends and family who have stopped smoking by substituting vape products for cigarettes.

# Text Analytics

## Concerns and Themes

Category	Example
TOBACCO_USA GE	I was a 2 pack a day smoker for over 10 years and the day I picked up a vape is the day I was able to quit cigarettes.
WHY_FLAVORS	Most of the people I know quit smoking (who quit smoking rather) stopped by using flavors OTHER than Menthol and Tobacco flavors;
RESUME_SMOKI NG	But lets just say if you took all flavors away from us I would more than likely go back to smoking
TRY_QUIT	I have tried several times to quit smoking tried every method none of them work until I started vaping
STOP_SMOKING	many people have benefited from these flavors and have quit smoking
SAVE_LIVES	Vaping can save lives if I can stop smoking cigarettes after 20 years of a pack and a half a day and now I'm down to 3 mg of nicotine
CHILD_USAGE	would you rather have kids vaping or smoking a cig I would rather hear a kid vaping than smoking to be honest.

**Categories:** Unsupervised learning generates combinations of terms connected with Boolean logic to identify groups of concerns and themes in comments.

# Text Analysis – Form Letters

## Cluster X

### FDA-2017-N-6565-18305 (Response 1)

Declaration of [@advFirst][@advLast]

I, [@advFirst][@advLast], declare and state as follows:...

2. I am submitting this declaration in response to the FDAs request for public comments on the above-referenced docket regarding a proposed rule on the regulation of flavors in tobacco products, including electronic nicotine delivery system (ENDS), or vapor products....

5. I have been using vapor products for 5 years and 11 months.

6. The categories of flavors of nicotine-containing e-liquid that I have used include Menthol/Mint, Fruit, Desserts. Of these, the flavor category that I use most often is Fruit.

7. Before I started using flavored e-liquid products, I typically smoked one pack of cigarettes per day.

8. E-cigarettes have helped my smoking cessation more than any other product or medication on the market to date. Having tried nicotine patches, gum, as well as [...] abbreviated for length...]

9. Since I began using flavored e-liquid products, I have been able to quit smoking cigarettes....

### FDA-2017-N-6565-18360 (Response 2)

Declaration of [@advFirst][@advLast]

I, [@advFirst][@advLast], declare and state as follows:...

2. I am submitting this declaration in response to the FDAs request for public comments on the above-referenced docket regarding a proposed rule on the regulation of flavors in tobacco products, including electronic nicotine delivery system (ENDS), or vapor products....

5. I have been using vapor products for 3 years and 5 months.

6. The categories of flavors of nicotine-containing e-liquid that I have used include Tobacco, Menthol/Mint, Fruit, Desserts, Other Sweets, Other Flavors. Of these, the flavor category that I use most often is Desserts.

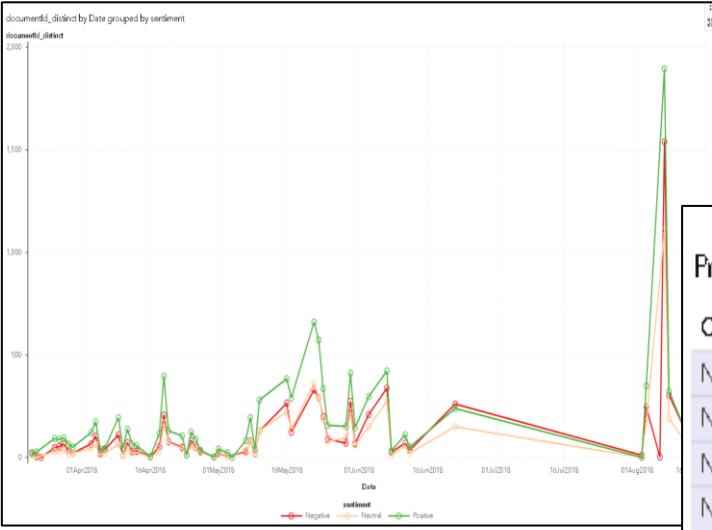
7. Before I started using flavored e-liquid products, I typically smoked one pack of cigarettes per day.

8. Over my thirty plus years of smoking (Marlboro Red, Camel Light, Marlboro Medium/27, Dunhill, and Sobrane) [...] abbreviated for length...]

9. Since I began using flavored e-liquid products, I have been able to quit smoking cigarettes....

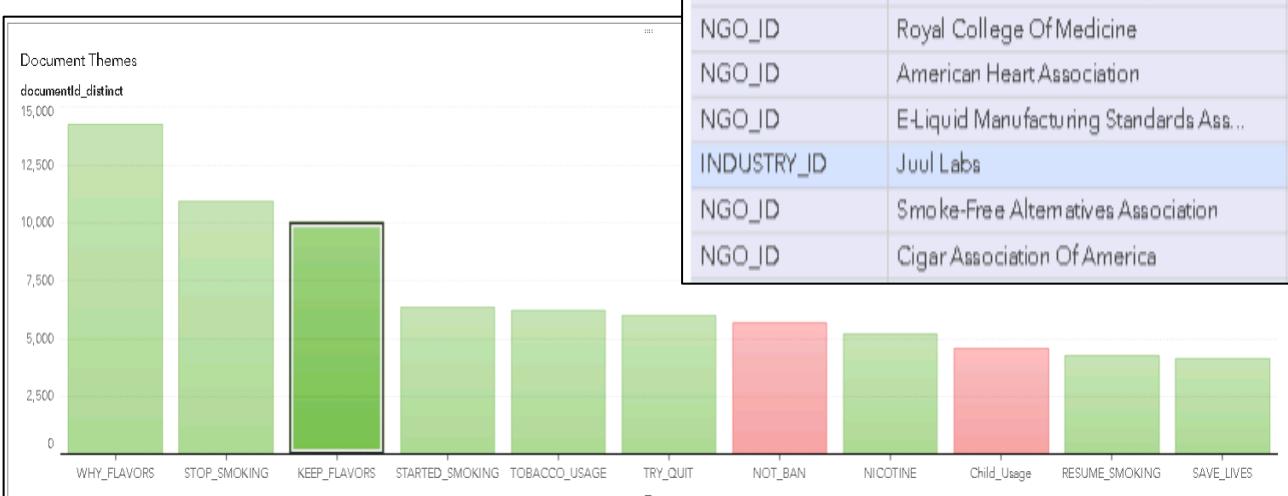
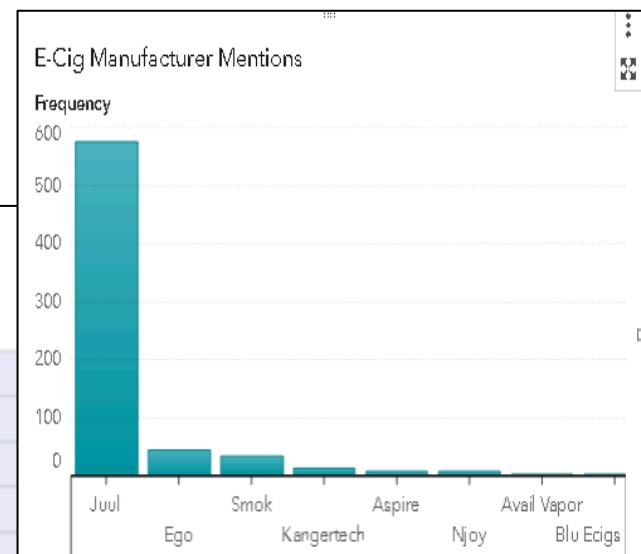
# FDA-2017-N-6565

## Quantify and Report



### Professional Organization

Concept	Extracted Term
NGO_ID	Royal College Of Physicians
NGO_ID	The Royal College
NGO_ID	American Cancer Society
NGO_ID	American Lung Association
NGO_ID	The American Cancer Society
NGO_ID	Royal College Of Medicine
NGO_ID	American Heart Association
INDUSTRY_ID	E-Liquid Manufacturing Standards Ass...
INDUSTRY_ID	Juul Labs
NGO_ID	Smoke-Free Alternatives Association
NGO_ID	Cigar Association Of America



# Text Analysis of *Regulations.gov* Comments

## Solution Benefits and ROI

10,000  
Comments



5 mins  
Manual  
Review



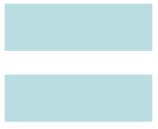
833 Hours  
or  
20 FTE  
Weeks

- Consistency
- Transparency
- Scalability

Scalable  
Score Code



1 week  
Model  
Development



Repeatable  
Solution

# Topic 1

# Audience Q&A

# Topic 2

## Role of Information in the Rulemaking Process

- **Carey Johnston, P.E.**: Environmental Engineer, U.S. Environmental Protection Agency's Office of Compliance
- **Michelle Cosby, J.D.**: President, American Association of Law Libraries

# Analyses Supporting Regulatory Development and Public Access

Carey A. Johnston, P.E.

U.S. EPA, Office of Compliance

GSA Public Meeting on Regulatory Data

Topic 2: Role of Information in the Rulemaking Process

30 April 2020

# Developing Rulemakings

## *Administrative Procedure Act of 1946*

- The Administrative Procedure Act (APA) is a Federal statute that governs how EPA (and other Federal administrative agencies) may propose and establish regulations and grants Federal courts oversight over all agency actions.
- In accordance with the APA, EPA makes most of the data and analyses that support regulatory development available for public review and comment.
  - Federal Government's on-line docket system (<https://www.regulations.gov>)
  - Program websites (e.g., <https://www.epa.gov/eg/>)
- EPA may need to redact information from the on-line docket system. For example, documents that are copyrighted or claimed as confidential business information
- Industry and public interest groups can challenge final Agency actions.

# Public Comments and Other Data Sources

- For most rulemakings, the public has an opportunity to review and comment on the proposed rule and supporting record prior to final Agency action.
  - The record is the official repository of the information and decisions that can be used in any subsequent litigation.
- This provides the public with the ability to:
  - correct errors in the Agency's record;
  - submit new information to supplement the existing record; and
  - provide comment on the proposed action and alternative actions.
- Public comments are also collected and made publicly available through the Federal Government's on-line docket system.

# Public Access to the Rulemaking Record

Typically, the record will contain supporting documents that are collected or developed during EPA's rulemakings. These items may include:

- Final site visit and sampling reports;
- Industry survey data;
- Summaries and transcripts of public meetings and hearings;
- Meeting minutes;
- Records of communications with parties outside of EPA, including telephone calls, memoranda, and letters;
- Public statements made by EPA employees in their official capacities; and
- Public comments, including written opinions, data, and other supporting information.



# Example of Data Collections & Analyses

## *National Industrial Wastewater Regulations*

- Data Collection
  - Surveys
  - Inspect production processes & wastewater treatment systems
  - Wastewater sampling
  - Industry Subcategorization
- Technology Assessment
  - Wastewater characterization; technology performance
  - Costs to install new technologies & process changes
  - Statistical analyses to derive effluent limit
- Economic Analysis
  - Economic achievability; analysis of market effects; cost-effectiveness
  - Value of environmental & human health benefits
- Environmental Assessment
  - Pollutant transport & exposure pathways; hazards
  - National & local impacts; benefits



# Example - Industrial Wastewater Regulations

## *Site Visits*

- Site visits help regulators meet industry representatives and discuss current wastewater treatment and potential alternatives.
- EPA identifies its objectives and data needs and prepares check-list to ensure all the necessary data is collected during the site visit.
- Information is documented in a site visit report and can include:
  - Description of the facility and its operations;
  - Overview of wastewater treatment;
  - Previously collected discharge data; and
  - Photos.
- All or parts of the site visit report can be claimed as confidential business information (not releasable to the public).
  - EPA's procedures for handling confidential business information (CBI) are documented in its regulations (40 CFR part 2, Subpart B).

# Example - Industrial Wastewater Regulations

## *Wastewater Sampling*

- After conducting a number of site visits, EPA may identify facilities for sampling.
- Significant work is needed in advance of the sampling episode in order to:
  - Select sampling locations based on knowledge of process and treatment operations;
  - Select analytes and appropriate methods to capture needed data; and
  - Prioritize sample collection to maximize data gathered without sacrificing quality
- Information is documented in a sampling episode report and can include:
  - Description of the facility and its operations;
  - Overview of wastewater treatment;
  - Sampling data; and
  - Photos
- All or parts of the sampling episode report can be claimed as confidential business information (not releasable to the public).

# Example - Industrial Wastewater Regulations

## *Industrial Surveys*

- EPA often collects engineering and financial data to support its assessment of available and affordable technology options
- This data collection provide a statistical basis for national estimates of number of facilities, incremental compliance costs and pollutant reductions, economic impacts and environmental benefits.
- The Paperwork Reduction Act (PRA) stipulates that every Federal agency must obtain approval from the Office of Management and Budget (another part of Federal Government) before collecting the same or similar information from 10 or more entities.
- In order to comply with the PRA, Federal agencies must prepare an Information Collection Request (ICR) in order to document that the data collection is required, minimized, and not duplicative.
  - Process often takes 12 to 18 months to complete ICR as there can be multiple rounds of public review for the survey instruments.
  - Some or all of the survey response can be claimed as CBI

# Example - Industrial Wastewater Regulations

## *Meetings with Stakeholders*

- During the development of an effluent guideline or study EPA often meets with industry and other public stakeholders.
- These meetings can help identify, collect, and clarify information and analyses.
- EPA also can conduct public meetings with an announcement in the Federal Register.
- EPA records attendance and the minutes for the supporting record.



# Example - Industrial Wastewater Regulations

## *Pulling it all together*

- After collecting data from industry EPA analyzes the data to identify or refine:
  - Technology options and the related pollutant reductions and incremental compliance costs
  - Economic impacts
  - Industry subsectors
  - Environmental benefits
- EPA conducts an internal decision-making process and documents its decision in a proposed rule for public comment.
- Proposed rulemaking are published in the Federal Register.
  - Public comment periods are typically 60 – 90 days or more.
  - Public comments are collected through Federal Government's on-line docket system.

# Example - Industrial Wastewater Regulations

## *Example Documentation*

- EPA typically summarizes the engineering analyses for its effluent guidelines rulemakings in “Development Documents.”
- These documents provide summaries of the data collection and the analyses supporting the engineering analysis for the proposed and final rules.
- These documents provide expand on the discussions in the preambles to the proposed and final rules and also highlight other key documents in the docket.



**Development Document For the Final Effluent Limitations Guidelines and Standards for the Metal Products and Machinery Point Source Category**

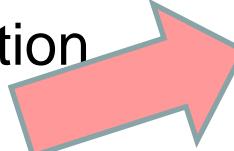


Printed on paper containing at least 30% postconsumer recovered fiber.

# Example - Industrial Wastewater Regulations

## *Example Documentation*

- These supporting documents are also used to:
  - help answer questions during implementation of the final rule; and
  - address public comments on the proposed rule.
- This documents can also help answer issues that arise if there is any litigation on the final action



From 1997 to 1999, EPA conducted 67 visits to iron and steel facilities in the United States and Canada in order to collect information on each site's manufacturing operations, wastewater generation, and wastewater treatment systems.<sup>5</sup> EPA in 1998 also solicited technical and economic information relevant to promulgation of a revised rule from various participants in the cokemaking industry through four surveys.<sup>6</sup> On the basis of the site visits and surveys, EPA selected 16 sites at which to perform wastewater sampling in order to characterize the effectiveness of the treatment processes. *See* Proposed Development Document at § 3.3. During this period, EPA conducted a variety of outreach efforts, in which the Institute participated, *id.* at § 3.5, including five stakeholders' meetings between 1998 and 2000 at which EPA described its preliminary position on the model technology options and data quality protocols, and solicited further comment and relevant data. *Id.*<sup>7</sup> EPA obtained additional information from secondary sources, including trade journals, industry databases, and studies by

---

dissolved form of nitrogen. Depending on other variables, high levels of ammonia-N are often toxic for aquatic life.

<sup>5</sup> See EPA Development Document for the Proposed Effluent Limitations Guidelines and Standards for the Iron and Steel Manufacturing Point Source Category ("Proposed Development Document") at § 3.2, No. 821-B-00-011 (December 2000).

U.S. District of Columbia Appeals Court, Iron and Steel ELGs (2006)

# Summary

- EPA makes most of the data and analyses that support regulatory development available for public review and comment.
- These documents and data can include site visits, survey, environmental or discharge/emissions sampling, meeting notes.
- EPA may conduct additional data collection in response to public comment and document its comments responses in the record.
- EPA uses the Federal Government's on-line docket system for regulatory data collection and analyses as well as for the collection of public comments.
- The on-line docket system generally contains all publicly available data, documents, and information. Information that is not made available through the online docket include information that is under copyright or claimed as CBI.



# GSA Public Meeting on Regulatory Data

## Role of Information in the Rulemaking Process

Michelle Cosby  
President  
American Association of Law Libraries  
April 30, 2020



# Recommendation 1

## **Integrate Systems and Improve Collaboration**

- Increase functionality between FederalRegister.gov and Regulations.gov
- Integrate Regulations.gov data with GPO's govinfo.gov

# Recommendation 2

## **Allow Users to Move Seamlessly Through Rulemaking Stages**

- Develop a revised classification scheme for document and docket types
- Enlist government and private sector experts in rulemaking, IT, and data management—as well as librarians—to help improve the “Advanced Search” feature

# Recommendation 3

## Link to Administrative Law Guides

- Include links on Regulations.gov to authoritative guides about administrative law, such as those produced by many academic and government law libraries

# Thank You

[mcosby@aall.org](mailto:mcosby@aall.org)

312.939.4764



## Topic 2

# Audience Q&A

# Topic 3

## Other Uses for Regulatory Information

- **Elizabeth Kowalewski:** Senior Analyst, U.S. Government Accountability Office
- **Hudson Hollister:** Founder and Principal at HData



# GAO's Analysis of Public Comments

Elizabeth  
Kowalewski, Senior  
Analyst



# Agenda

- GAO's Scope and Methodology
- Data Acquisition Challenges
- Data Analysis Challenges

## GAO's Scope and Methodology

GAO began work on this topic in early 2018. This work was requested as a result of widespread interest in the integrity of the public comment process following claims that comments using false identity information had been submitted to several high-profile rulemakings.

## We are conducting a comprehensive analysis of millions of comments across dockets and agencies

- We selected 10 agencies as case studies that received a high volume of public comments during the course of rulemaking proceedings **that accepted comments** from January 1, 2013, through December 31, 2017 .
- The 8 participating agencies we selected represent more than 90% of all comments agencies reported to Regulations.gov during our time period.
- We plan to describe characteristics of the comment data, and are surveying a generalizable sample of comments with email addresses to verify the source of the comments.

# Data Acquisition Challenges

In accordance with the requirements of the Administrative Procedure Act, the openness of the public comment system fosters significant variation in the data

## Regulations.gov was built to accommodate wide variation across agencies

- After receiving permission from participating agencies, the PMO provided us with 100s of metadata files and almost 1 TB of XML files from a direct FDMS export
- As reported in GAO-19-483, however, there is little consistency in the way agencies use FDMS to store and post comments to Regulations.gov
- In addition to the FDMS exports, we had to obtain “archive” comments in other formats (.pst, .tiff) from 3 agencies that do not store all comments in the system.

# Data Analysis Challenges

Two years later, and we still do not yet have a reliable estimate of the number of comments submitted

## Sophisticated, unstructured text analytics are required for most comments

- Most comments contain attachments, with or without meaningful text in the Regulations.gov comment field, in every file format imaginable
- The analysis is done in R and Python. We used Apache Tika and Tesseract to pull text from the attachments. We then used spaCy for natural language processing
- We are using supervised machine learning to predict the type of attachment, to allow us to better connect names/emails with comment content



# Questions?

# Open Data Standards for Regulatory Materials

Better Transparency, Better Tools

# The Need for Structured Data in Regulation

Without structured data:

- Regulatory materials are drafted, reviewed, compared, and analyzed manually
- Costly transformations are necessary as regulatory materials move from agencies to the Federal Register, interagency systems, and codification
- Natural language processing and other text analysis tools are useful, but must be applied all over again at each stage of the process

With structured data:

- Regulatory materials will be drafted, reviewed, compared, and analyzed with modern tools
- Once first drafted, regulatory materials will require no transformation to move through publication, review, and codification
- With redlining and change management automatic, text analysis tools can be focused on substance

# U.S. Legislative Markup | USLM

Open-source data format:

- Expresses structures of legislation and regulation, allowing electronic drafting and comparison tools
- Nonproprietary and freely useable
- Maintained by U.S. House Clerk and Government Publishing Office
- Allows electronic management of codes of laws and codes of rules - like a codebase
- Based on international format adopted by OASIS, international standard-setting body

Adoption of USLM:

- Clerk of the House: automatic redlining to show changes between amendments and bills
- Office of the Law Revision Counsel: drafting codification bills
- H. Res. 756: future adoption for all legislative materials by entire U.S. House of Representatives
- GPO projects: U.S. Code; 2 full years of the Federal Register; CFR titles 5, 12, 27, 40
- California, Oregon, Nova Scotia legislatures



Vol. Vol. 82 No. No. 61

Friday, March 31, 2017

## Rules and Regulations

### DEPARTMENT OF HOMELAND SECURITY

#### Office of the Secretary

#### 6 CFR Part 5

Privacy Act of 1974: Implementation of Exemptions; Department of Homeland Security/United States Coast Guard-031 USCZ Law Enforcement (ULE) System of Records

AGENCY: Privacy Office, DHS.

ACTION: Final rule.

#### SUMMARY:

The Department of Homeland Security is issuing a final rule to amend its regulations to exempt portions of a newly established system of records titled, "Department of Homeland Security/United States Coast Guard-031 USCZ Law Enforcement (ULE) System of Records" from certain provisions of the Privacy Act. Specifically, the Department exempts portions of the "Department of Homeland Security/United States Coast Guard-031 USCZ Law Enforcement (ULE) System of Records" from one or more provisions of the Privacy Act because of criminal, civil, and administrative enforcement requirements.

#### DATES:

This final rule is effective March 31, 2017.

#### FOR FURTHER INFORMATION CONTACT:

For general questions please contact: Marilyn Scott-Perez (202-475-3515), Privacy Officer, Commandant (CG-61), United States Coast Guard, Mail Stop 7710, Washington, DC 20593. For privacy issues please contact: Jonathan R. Cantor, (202-343-1717), Acting Chief Privacy Officer, Privacy Office, Department of Homeland Security, Washington, DC 20528.

#### SUPPLEMENTARY INFORMATION:

##### Background

The Department of Homeland Security (DHS) United States Coast Guard (USCG) published a Notice of Proposed Rulemaking (NPRM) in the *FEDERAL REGISTER*, 81 FR 88635, December 8, 2016, proposing to exempt portions of the system of records from one or more provisions of the Privacy Act because of criminal, civil, and administrative enforcement requirements. The system of records is the DHS/USCG-031 USCZ Law Enforcement (ULE) System of Records. The DHS/USCG-031 USCZ Law Enforcement (ULE) System of Records notice was published concurrently in the *FEDERAL REGISTER*, 81 FR 88697, December 8, 2016, and comments were invited on both the NPRM and System of Records Notice (SORN).

##### Public Comments

DHS received no comments on the NPRM and no comments on the SORN.

##### List of Subjects in 6 CFR Part 5

Freedom of information, Privacy.

For the reasons stated in the preamble, DHS amends chapter I of Title 6, Code of Federal Regulations, as follows:

- Revise the authority citation for part 5 to read as follows: Authority: 6 U.S.C. 101 *et seq.*; Pub. L. 107-296, 116 Stat. 2135; 5 U.S.C. 301.

Subpart A also issued under 5 U.S.C. 552. Subpart B also issued under 5 U.S.C. 552a.



```

<preface>
  <docPublicationName role="documentName">Federal Register</docPublicationName>
  <startingPage>15901</startingPage>
  <volume style="-uslm-sgm:VOL">Vol. 82.</volume>
  <issue style="-uslm-sgm:NO">No. 61</issue>
  <dc:date><!-- style="-uslm-sgm:DATE" -->Friday, March 31, 2017</dc:date>
  <docPublicationName role="unitName" style="-uslm-sgm:UNITNAME">Rules and Regulations</docPublicationName>
</preface>
<main>
  <rules style="-uslm-sgm:RULES">
    <collection role="agencyGroup">
      <component>
        <rule style="-uslm-sgm:RULE" identifier="us/fr/82/61/frdoc2017-06327">
          <meta> [10 lines]
          <preface><?page num="15901"?>
            <organization style="-uslm-sgm:AGENCY">DEPARTMENT OF HOMELAND SECURITY</organization>
            <organization style="-uslm-sgm:SUPERG">Office of the Secretary</organization>
            <affected style="-uslm-sgm:CFR"><ref href="/us/cfr/t6/p5">6 CFR Part 5</ref></affected>
            <subject style="-uslm-sgm:SUBJECT">Privacy Act of 1974: Implementation of Exemptions;
              Department of Homeland Security/United States Coast Guard-031 USCZ Law Enforcement
              (ULE) System of Records</subject>
            <organization style="-uslm-sgm:AGV">
              <headingText style="-uslm-sgm:HED">AGENCY:</headingText> Privacy Office, DHS.</organization>
            <action>
              <actionDescription>
                <headingText style="-uslm-sgm:HED">ACTION:</headingText> Final rule.</actionDescription>
              </action>
            </preface>
            <main>
              <rulePreamble><?page num="15901"?>
                <statement style="-uslm-sgm:SUM" role="summary">
                  <content>
                    <heading style="-uslm-sgm:HED">SUMMARY:</heading>
                    <p style="-uslm-sgm:F">The Department of Homeland Security is issuing a final rule to
                      amend its regulations to exempt portions of a newly established system of records titled,
                      "Department of Homeland Security/United States Coast Guard-031 USCZ Law Enforcement (ULE)
                      System of Records" from certain provisions of the Privacy Act. Specifically, the
                      Department exempts portions of the "Department of Homeland Security/United States Coast
                      Guard-031 USCZ Law Enforcement (ULE) System of Records" from one or more provisions of the
                      Privacy Act because of criminal, civil, and administrative enforcement requirements.</p>
                  </content>
                </statement>
                <statement style="-uslm-sgm:EFFDATE" role="effectiveDate">
                  <content>
                    <heading style="-uslm-sgm:HED">DATES:</heading>
                    <p style="-uslm-sgm:F">This final rule is effective
                      <date date="2017-03-31">March 31, 2017</date>.</p>
                  </content>
                </statement>
                <statement style="-uslm-sgm:FURINF" role="furtherInformation">
                  <content>
                    <heading style="-uslm-sgm:HED">FOR FURTHER INFORMATION CONTACT:</heading>
                    <p style="-uslm-sgm:F">For general questions please contact: Marilyn Scott-Perez (202-475-3515),
                      Privacy Officer, Commandant (CG-61), United States Coast Guard, Mail Stop 7710, Washington,
                      DC 20593. For privacy issues please contact: Jonathan R. Cantor, (202-343-1717), Acting Chief
                      Privacy Officer, Privacy Office, Department of Homeland Security, Washington, DC 20528.</p>
                  </content>
                </statement>
              </main>
            </rulePreamble>
          </rule>
        </component>
      </collection>
    </rules>
  </main>
</preface>

```

# USLM enables:

Auto-comparison

## Section 1. Short title

This Act may be cited as the "Endangered Salmon and Fisheries Predation Prevention Act".

[NOTE-- DELETED : Sec. 2. Findings]

## Sec. 32. Sense of Congress

It is the sense of the Congress that—

(1) preventing ~~on of~~ predation by sea lion, recovery of listed salmonid stocks *listed under the Endangered Species Act of 1973 (16 U.S.C. 1531 et seq.)*, and preventing ~~on of the~~ future listings of fish stocks in the Columbia River under the Endangered Species Act of 1973 (16 U.S.C. 1531 et seq.) is such *Act are* a vital priority; and

(2) the Federal Government should continue to fund lethal and nonlethal removal *of pinnipeds as well as deterrence* measures for preventing such predation.

## Sec. 43. Taking of sea lion on the Columbia River and its tributaries to protect endangered and threatened species of salmon and other nonlisted fish species

Section 120(f) of the Marine Mammal Protection Act of 1972 (16 U.S.C. 1389(f)) is amended to read as follows:

"(f) TEMPORARY MARINE MAMMAL REMOVAL AUTHORITY ON THE WATERS OF THE COLUMBIA RIVER ~~OR AND~~ ITS TRIBUTARIES.—

"(1) REMOVAL AUTHORITY.— Notwithstanding any other provision of this Act, the Secretary may issue a permit to an eligible entity to authorize the intentional lethal taking on the waters of the Columbia River and its tributaries of individually identifiable sea lion that are part of a population *or stock* that is not categorized under this Act as depleted *or strategic* for the purpose of protecting *the*

*"(A) species of salmon, steelhead, or eulachon* that are listed as endangered species or threatened species under the Endangered Species Act of 1973 (16 U.S.C. 1531 et seq.); and *other*

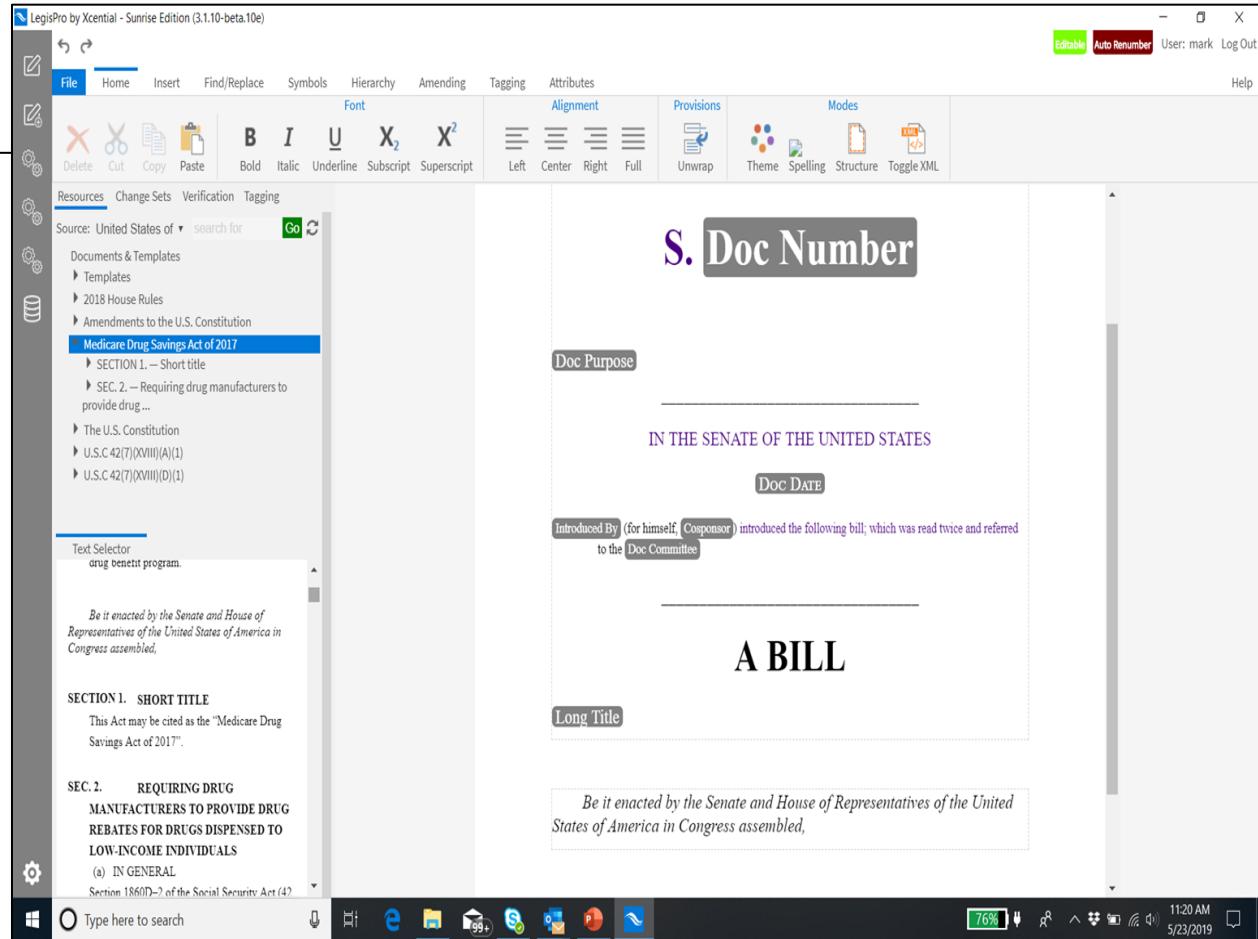
*"(B) species of lamprey or sturgeon that are not listed fish species as endangered or threatened but are listed as a species of concern.*

"(2) PERMIT PROCESS.—

"(A) IN GENERAL.— An eligible entity may apply to the Secretary for a permit under this subsection.

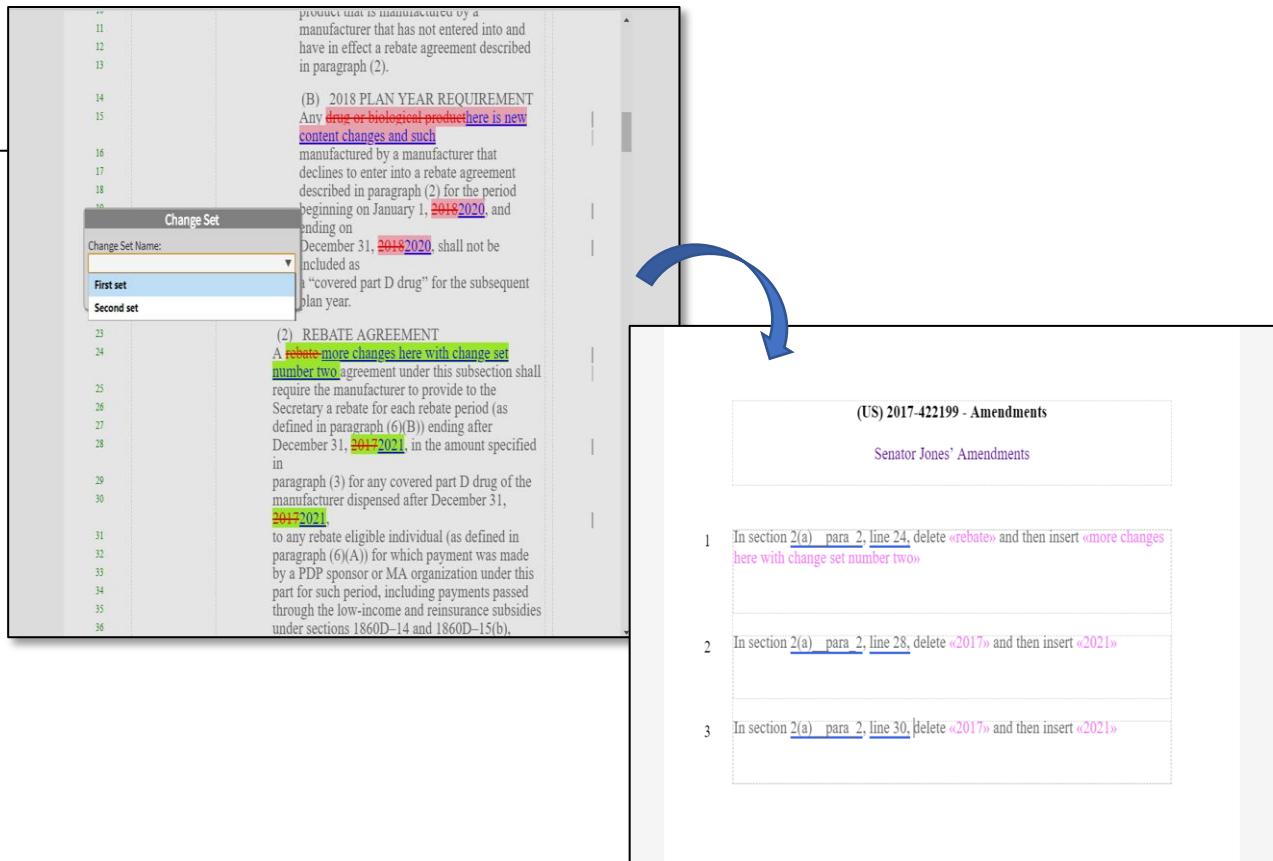
# USLM enables:

Drafting directly  
into existing codes



# USLM enables:

Amendment  
generations



# USLM enables:

Comment references  
to specific segments  
of a proposal

the comment process that follows, will turn on this fundamental question. ¶ 3. ¶ 4. Today, there are no legally enforceable rules by which the

A simple but powerful structure allows paragraph-level references.

from threatening Internet openness (as well as enhancing the transparency rule that is currently in effect). ¶ 4. The goal of this proceeding is to find the best approach to protecting and promoting Internet openness. Per the blueprint offered by the D.C. Circuit in its decision in [Verizon v. FCC](#) section 706 of the Telecommunications Act of 1996 [Title II of the Communications Act](#) section 706 [Title II](#) [Title III](#) section 706 [Title II](#) [text](#) It is important to always remember that the Internet is a collection of networks, not a single network. And that means that each broadband provider can either add to the benefits that the Internet delivers to Americans by

# Draft, Publish, Manage, Review, and Maintain in USLM

USLM throughout the regulatory process

Future: machine-readable regulation

1. Drafting (agencies)
2. Federal Register publication (NARA)
3. Regulations.gov comment management (GSA)
4. Reginfo.gov review (OIRA)
5. Code of Federal Regulations codification (GPO)

- ★ Agencies work with regulated industries to create USLM tags that express mandates
- ★ USLM versions of regulations include machine-readable mandates
- ★ Regulated enterprises' compliance software automatically ingests and executes mandates
- ★ For cost-benefit analysis, proposed rules can be tested on compliance software

# Topic 3

# Audience Q&A

# Concluding Remarks

- Today is another step in a broad conversation for modernizing electronic rulemaking
- Please remember to submit your suggestions

Go to: <https://www.regulations.gov/document?D=GSA-GSA-2020-0002-0012>  
or search for “Office of Regulation Management” in Regulations.gov

# Thank you for coming