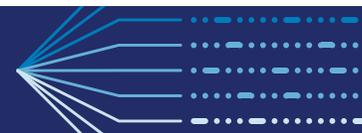




Use of Machine Learning to Reduce State IT Burden of Regulatory Partner Submissions

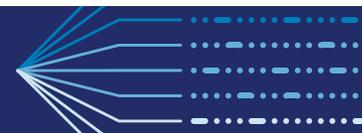
FDA Digital Transformation Symposium



If we want to make submitting data easier...

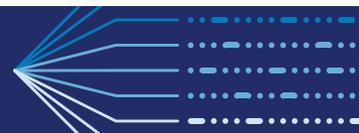
- ORA has a State, local, tribal, and territorial (SLTT) submission IT solution: ORA's Data Exchange
 - System-to-System communication allows for the transfer of data directly from State IT systems to FDA. Several States leverage this solution and generally have vendor support.
 - The ORA Partners Portal (ORAPP) is a web-based portal that allows for sending structured data, typically using a template. ORAPP also allows sending of FDA data to States.
 - Information submitted to the ORA DX as documents cannot be parsed into structured data and is typically manually reviewed.

- Challenges to State IT Implementations of Data Submission
 - There may be limited IT personnel to be able to commit to FDA submission solutions.
 - There may be limited funding to commit to FDA submission solutions.
 - States have many competing priorities and FDA submission may not be top priority.



...Let's look at MLaaS

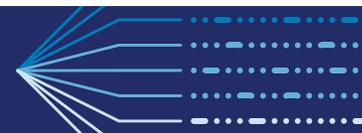
- ...and test with FDA Form 4056: Produce Farm Inspection Observations
 - Can be a fillable pdf.
 - Can be filled in by hand or scanned (so lots of hand-writing or images of the form might be expected).
- ORA DX would be able to provide end-to-end solution of uploading documents, extracting data, and saving data through ORAPP or future platform.
- Like SLTT's, ORA also has resource constraints and any MLaaS implementation requires a determination of the benefit to both SLTT, FDA, and public health.



Description of the Test Sample Data Set

- The sample contains 166 forms to demonstrate different cases.
- Font:
 - Regular
 - Script (*to simulate handwriting*)
 - Handwriting
- Rotation:
 - Aligned
 - Rotated (rotated and/or shifted when scanned)

Rotation	Font	Count
aligned	handwriting	6
aligned	regular	40
aligned	script	40
rotated	regular	40
rotated	script	40



Entire PDF with Multiple Pages can be Parsed Preserving the Hierarchy



Example output with hierarchical relationships between sections, questions and their sub-components:
 (Only a selection of rows included to show the preserved hierarchy.)

Page	Section	Key	Subkey
6	Sprouts (21 CFR Part 112, Subpart M)	39. § 112.142: Seeds or beans used to grow sprouts	Observation
6	Sprouts (21 CFR Part 112, Subpart M)	39. § 112.142: Seeds or beans used to grow sprouts	Corrective action taken
6	Sprouts (21 CFR Part 112, Subpart M)	39. § 112.142: Seeds or beans used to grow sprouts	Description
6	Sprouts (21 CFR Part 112, Subpart M)	40. § 112.143(a): Fully-enclosed buildings	Observation
6	Sprouts (21 CFR Part 112, Subpart M)	41. § 112.143(b): Cleaning and sanitizing food-contact surfaces	Corrective action taken
6	Sprouts (21 CFR Part 112, Subpart M)	41. § 112.143(b): Cleaning and sanitizing food-contact surfaces	Description
7	Sprouts (21 CFR Part 112, Subpart M)	42. §§ 112.144(a), 112.145, and 112.146: Environmental monitoring for <i>Listeria</i> species or <i>L. monocytogenes</i>	Description
7	Sprouts (21 CFR Part 112, Subpart M)	43. §§ 112.144(b) and (c), 112.147 and 112.148: Testing spent irrigation water or in-process sprouts for pathogens (written	Observation
7	Sprouts (21 CFR Part 112, Subpart M)	43. §§ 112.144(b) and (c), 112.147 and 112.148: Testing spent irrigation water or in-process sprouts for pathogens (written	Corrective action taken

Sprouts (21 CFR Part 112, Subpart M)

Check here if entity does not engage in growing, harvesting, packing, and/or holding of sprouts

39. § 112.142: Seeds or beans used to grow sprouts
 Observation Corrective action taken
 Description: *Sprout seeds, beans*

40. § 112.143(a): Fully-enclosed buildings
 Observation Corrective action taken
 Description: *Refrigerated warehouse*

41. § 112.143(b): Cleaning and sanitizing food-contact surfaces
 Observation Corrective action taken
 Description: *Sanitizing surfaces of food*

FORM FDA 4056 (01/19) Page of

FARM NAME (include business name, if different)

DATE(S) OF INSPECTION FEI NUMBER

42. §§ 112.144(a), 112.145, and 112.146: Environmental monitoring for *Listeria* species or *L. monocytogenes* (written environmental monitoring plan, collection and testing, corrective actions)
 Observation Corrective action taken
 Description: *Sample farm enviro. monitoring*

43. §§ 112.144(b) and (c), 112.147 and 112.148: Testing spent irrigation water or in-process sprouts for pathogens (written sampling plan, collection and testing, corrective actions)

Performance Metrics



To evaluate the model, we used the performance metrics below:

Overall accuracy:

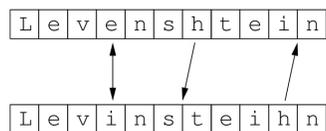
- Overall accuracy (combined accuracy on text fields and checkboxes): computed as the weighted average of the Levenshtein score ($1 - \text{edit distance}/\text{text length}$) for text fields, and the checkbox accuracy.

Accuracy on checkboxes and text fields separately:

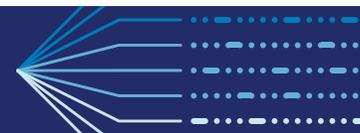
- Percentage of checkboxes that are correctly detected as checked or unchecked.
- Percentage perfectly captured text fields: The percentage of text fields with the predicted text that are 100% correct –with no character mistakes. We use the Levenshtein score.

Levenshtein distance, aka edit distance, between the extracted data from forms and their true values is computed as follows:

Ex:



- Edit distance = 3 (substitution, insertion, deletion)
- Levenshtein score = $1 - \frac{3}{11} = 0.727$

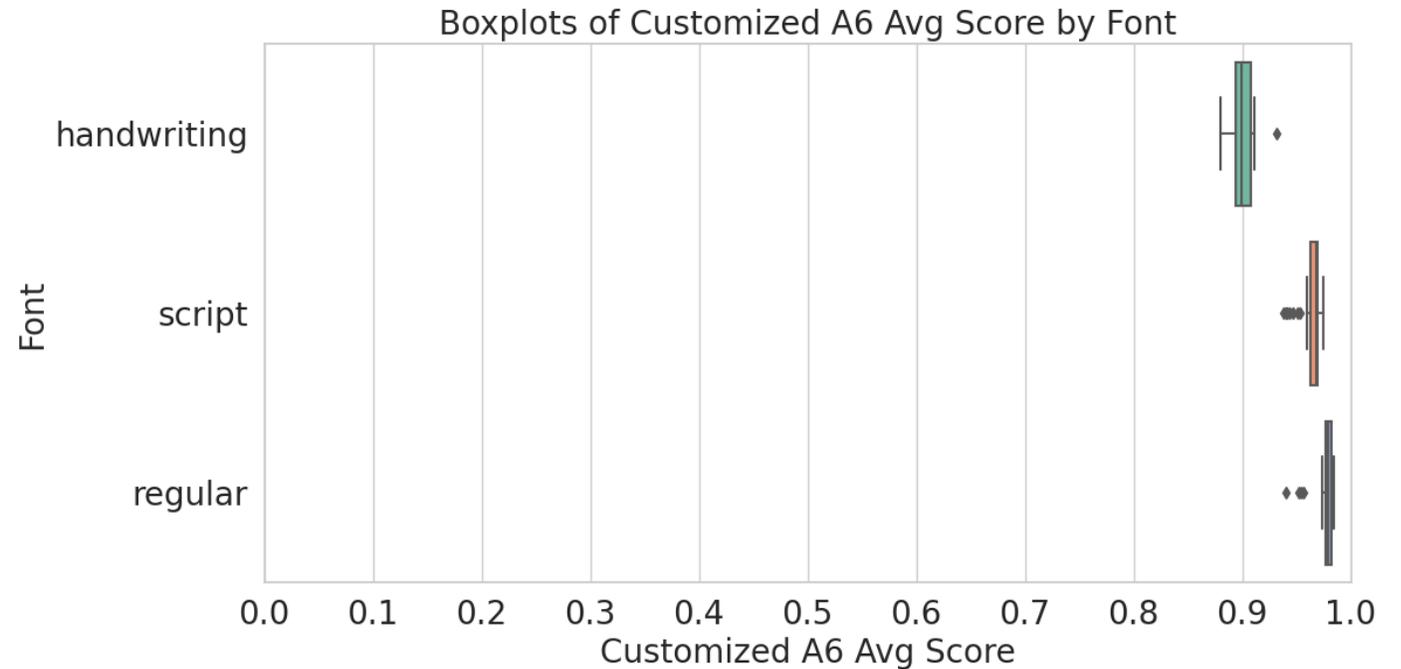


Statistics of Processing Entire PDF (Multiple Pages) - Aligned

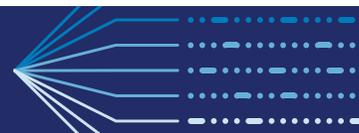


Overall accuracy scores in average are:

- 98% for forms with regular font
- 96% for script font
- 90% for handwriting



font	count	mean	std	min	25%	50%	75%	max
regular	40	0.977	0.009	0.94	0.977	0.979	0.982	0.984
script	40	0.964	0.01	0.938	0.963	0.968	0.97	0.975
handwriting	6	0.902	0.018	0.88	0.894	0.899	0.908	0.932



Checkboxes can be Parsed with a High Accuracy



- Checkboxes in forms can be “checked” or “unchecked”
- Checkboxes can be filled out variety of ways such as ‘x’ or checkmarks

40. § 112.143(a): Fully-enclosed buildings

Observation Corrective action taken

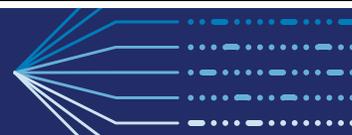
Description:
Large structures with walls & roof

Checkbox detection accuracy over the aligned documents:

rotation	font	count	mean	std	min	25%	50%	75%	max
aligned	regular	40	0.996	0.005	0.99	0.99	1	1	1
aligned	script	40	0.996	0.005	0.99	0.99	1	1	1
aligned	handwriting	6	0.992	0.008	0.98	0.99	0.99	0.998	1

- Our system can detect checkbox status robustly including handwriting

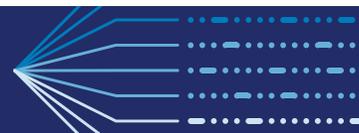
Page	Section	Key	Subkey	CheckBox	Value predicted
6	Sprouts (21 CFR Part 112, Subpart M)	40. § 112.143(a): Fully-enclosed buildings	Observation	TRUE	selected
6	Sprouts (21 CFR Part 112, Subpart M)	40. § 112.143(a): Fully-enclosed buildings	Corrective action taken	TRUE	selected
6	Sprouts (21 CFR Part 112, Subpart M)	40. § 112.143(a): Fully-enclosed buildings	Description:	FALSE	Large structures with walls & roof



Challenges with Document Alignment



- Potential challenges with the layout of the form when it is scanned:
 - Shift
 - Zoom in/out
 - Rotation
 - Distortion



Robust Detection for Different Shifts



DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration

PRODUCE FARM INSPECTION OBSERVATIONS



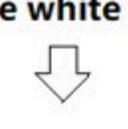
Name of State and Department (if acting under commission with FDA)		DISTRICT OFFICE ADDRESS	
Conducting Inspection Agency		2, Kavi Guru Road, Garia.	
DISTRICT OFFICE PHONE NUMBER	DATE(S) OF INSPECTION	FEI NUMBER	
919.856.2131	May 13th, 2019	22213121	
LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO: Smith, John Q., D., Doctor			
FARM NAME (include business name, if different)			
SMITH'S FARM			
OWNER/OPERATOR			
John Smith			
FARM MAILING ADDRESS		FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates)	
1051 E Valley Rd, Santa Y		N/A	
TYPE OF INSPECTION:		CROPS OBSERVED DURING INSPECTION	
<input checked="" type="checkbox"/> Initial <input type="checkbox"/> Routine <input type="checkbox"/> Follow-up <input type="checkbox"/> For-cause <input type="checkbox"/> Other (please specify)		Corn, Wheat, Soybean	

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.

This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112) or any other legal requirement.

Representatives of the regulatory agency should record their observations on this form as clearly and specifically as possible and should order their observations by significance within each area (most important first). In some cases, an observation may relate to more than one topic area. Representatives of the regulatory agency should record observations in the topic area listed below that, in the representatives' judgment, is the most appropriate topic. Not all topic areas may be applicable in every situation. In

more white space



DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration

PRODUCE FARM INSPECTION OBSERVATIONS



Name of State and Department (if acting under commission with FDA)		DISTRICT OFFICE ADDRESS	
Conducting Inspection Agency		2, Kavi Guru Road, Garia.	
DISTRICT OFFICE PHONE NUMBER	DATE(S) OF INSPECTION	FEI NUMBER	
919.856.2131	May 13th, 2019	22213121	
LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO: Smith, John Q., D., Doctor			
FARM NAME (include business name, if different)			
SMITH'S FARM			
OWNER/OPERATOR			
John Smith			
FARM MAILING ADDRESS		FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates)	
1051 E Valley Rd, Santa Y		N/A	
TYPE OF INSPECTION:		CROPS OBSERVED DURING INSPECTION	
<input checked="" type="checkbox"/> Initial <input type="checkbox"/> Routine <input type="checkbox"/> Follow-up <input type="checkbox"/> For-cause <input type="checkbox"/> Other (please specify)		Corn, Wheat, Soybean	

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.

This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112)

Robust Detection for Different Zoom Scales



Original Scale

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration

PRODUCE FARM INSPECTION OBSERVATIONS

Name of State and Department (if acting under commission with FDA) DISTRICT OFFICE ADDRESS

Conducting Inspection Agency

DISTRICT OFFICE PHONE NUMBER DATE(S) OF INSPECTION FEI NUMBER

919.856.2133 May 13th, 2019 PARRISIDE

LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO:
Smith, John Q., D., Doctor

FARM NAME (include business name, if different)

SMITH'S FARM

OWNER/OPERATOR

John Smith

FARM MAILING ADDRESS FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates)

1051 E Valley Rd, Santa Y N/A

TYPE OF INSPECTION: Initial Routine Follow-up For-cause Other (please specify)

CROPS OBSERVED DURING INSPECTION

Corn, Wheat, Soybean

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.

This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112) or any other legal requirement.

Representatives of the regulatory agency should record their observations on this form as clearly and specifically as possible and should order their observations by significance within each area (most important first). In some cases, an observation may relate to more than one topic area. Representatives of the regulatory agency should record observations in the topic area listed below that, in the representatives' judgment, is the most appropriate topic. Not all topic areas may be applicable in every situation. In

Detection for Smaller Scale

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration

PRODUCE FARM INSPECTION OBSERVATIONS

Name of State and Department (if acting under commission with FDA) DISTRICT OFFICE ADDRESS

Conducting Inspection Agency

DISTRICT OFFICE PHONE NUMBER DATE(S) OF INSPECTION FEI NUMBER

919.856.2133 May 13th, 2019 PARRISIDE

LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO:
Smith, John Q., D., Doctor

FARM NAME (include business name, if different)

SMITH'S FARM

OWNER/OPERATOR

John Smith

FARM MAILING ADDRESS FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates)

1051 E Valley Rd, Santa Y N/A

TYPE OF INSPECTION: Initial Routine Follow-up For-cause Other (please specify)

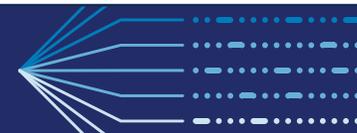
CROPS OBSERVED DURING INSPECTION

Corn, Wheat, Soybean

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.

This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112) or any other legal requirement.

Representatives of the regulatory agency should record their observations on this form as clearly and specifically as possible and should order their observations by significance within each area (most important first). In some cases, an observation may relate to more than one topic area. Representatives of the regulatory agency should record observations in the topic area listed below that, in the representatives' judgment, is the most appropriate topic. Not all topic areas may be applicable in every situation. In



Automatic Correction for Rotation

- Developed automated way to detect rotation of an image during pre-processing
- Usable for correction of rotational tilt of scanned images. More work needed

Rotated

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration
PRODUCE FARM INSPECTION OBSERVATIONS

Name of State and Department (if acting under commission with FDA) Virginia, FDA
DISTRICT OFFICE ADDRESS 123 Main Street, Anytown, USA

DISTRICT OFFICE PHONE NUMBER 555-867-5309
DATE(S) OF INSPECTION 10/13/19 & 10/20/19
FEI NUMBER 3040090598

LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO: John Doe, B.S., Manager

FARM NAME (include business name, if different) Green Acres Farms LLC

OWNER/OPERATOR John Smith

FARM MAILING ADDRESS 67 Inn Lane, Anytown FL, 12345 USA
FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates) 28.7336,-81.2456

TYPE OF INSPECTION:
 Initial Routine Follow-up For-cause
 Other (please specify)

CROPS OBSERVED DURING INSPECTION Apples, onions, corn, cherries

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.
This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112) or any other legal requirement.

Representatives of the regulatory agency should record their observations on this form as clearly and specifically as possible and should order their observations by significance within each area (most important first). In some cases, an observation may relate to more than one topic area. Representatives of the regulatory agency should record observations in the topic area listed below that, in the representatives' judgment, is the most appropriate topic. Not all topic areas may be applicable in every situation. In addition, representatives of the regulatory agency may not examine every aspect of the farm's operation during an inspection, so a topic area left blank should not be interpreted to mean the farm is in compliance, or not in compliance, with requirements related to that topic area.

Representatives of the regulatory agency should discuss all observations with the management of the farm or their representative as they are observed, or on a daily basis, to minimize surprises, errors, and misunderstandings when this form is issued. Discussion should include those observations which may be written on the form and those that will only be discussed with management during the closeout meeting. This form should be issued during the exit conference of all produce inspections, including when no observations have been recorded.

The farm may use this opportunity to ask questions about the observations or to request clarification. If the farm has implemented, or plans to implement, corrective action in response to an observation, this may be discussed with the representatives of the regulatory agency during the inspection. Representatives of the regulatory agency should annotate the form, as applicable, with any completed or promised corrections discussed during the inspection. FDA representatives are encouraged to verify the farm's completed corrective actions during the inspection as long as the verification does not unreasonably extend the duration of the inspection. Inclusion of annotations regarding corrective actions does not signify any conclusion by the regulatory agency regarding the sufficiency of the actions.

FORM FDA 4056 (01/19) Page of PSC Publishing Services (301) 443-6740 EF

Corrected

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration
PRODUCE FARM INSPECTION OBSERVATIONS

Name of State and Department (if acting under commission with FDA) Virginia, FDA
DISTRICT OFFICE ADDRESS 123 Main Street, Anytown, USA

DISTRICT OFFICE PHONE NUMBER 555-867-5309
DATE(S) OF INSPECTION 10/13/19 & 10/20/19
FEI NUMBER 3040090598

LAST NAME, FIRST NAME, MIDDLE INITIAL AND TITLE OF INDIVIDUAL TO WHOM REPORT IS ISSUED (Most responsible individual present) TO: John Doe, B.S., Manager

FARM NAME (include business name, if different) Green Acres Farms LLC

OWNER/OPERATOR John Smith

FARM MAILING ADDRESS 67 Inn Lane, Anytown FL, 12345 USA
FARM PHYSICAL LOCATION, IF DIFFERENT FROM MAILING ADDRESS (e.g., location identifiers such as GPS coordinates) 28.7336,-81.2456

TYPE OF INSPECTION:
 Initial Routine Follow-up For-cause
 Other (please specify)

CROPS OBSERVED DURING INSPECTION Apples, onions, corn, cherries

This form lists factual observations made by the FDA representative(s) during the inspection of the farm's operation.
This is not a final FDA determination of compliance, or non-compliance, with the Produce Safety Rule (21 CFR Part 112) or any other legal requirement.

Representatives of the regulatory agency should record their observations on this form as clearly and specifically as possible and should order their observations by significance within each area (most important first). In some cases, an observation may relate to more than one topic area. Representatives of the regulatory agency should record observations in the topic area listed below that, in the representatives' judgment, is the most appropriate topic. Not all topic areas may be applicable in every situation. In addition, representatives of the regulatory agency may not examine every aspect of the farm's operation during an inspection, so a topic area left blank should not be interpreted to mean the farm is in compliance, or not in compliance, with requirements related to that topic area.

Representatives of the regulatory agency should discuss all observations with the management of the farm or their representative as they are observed, or on a daily basis, to minimize surprises, errors, and misunderstandings when this form is issued. Discussion should include those observations which may be written on the form and those that will only be discussed with management during the closeout meeting. This form should be issued during the exit conference of all produce inspections, including when no observations have been recorded.

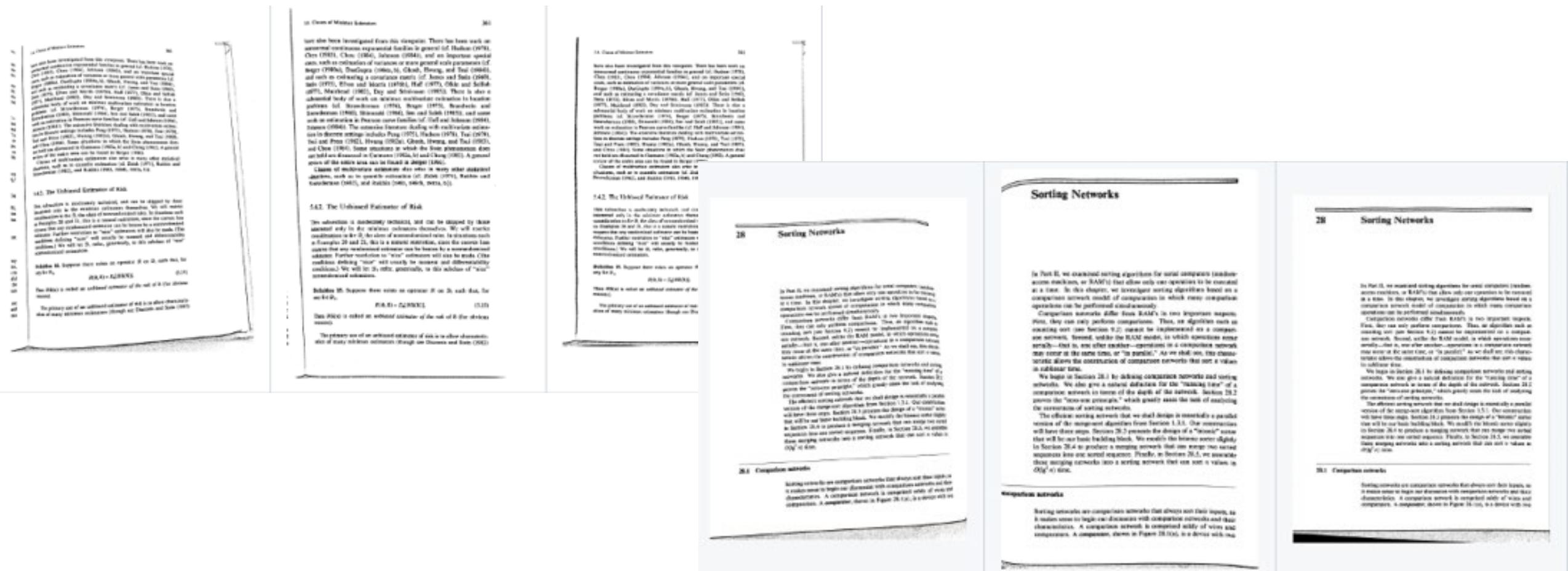
The farm may use this opportunity to ask questions about the observations or to request clarification. If the farm has implemented, or plans to implement, corrective action in response to an observation, this may be discussed with the representatives of the regulatory agency during the inspection. Representatives of the regulatory agency should annotate the form, as applicable, with any completed or promised corrections discussed during the inspection. FDA representatives are encouraged to verify the farm's completed corrective actions during the inspection as long as the verification does not unreasonably extend the duration of the inspection. Inclusion of annotations regarding corrective actions does not signify any conclusion by the regulatory agency regarding the sufficiency of the actions.

FORM FDA 4056 (01/19) Page of PSC Publishing Services (301) 443-6740 EF

Correction



Document Distortion – Additional Improvements



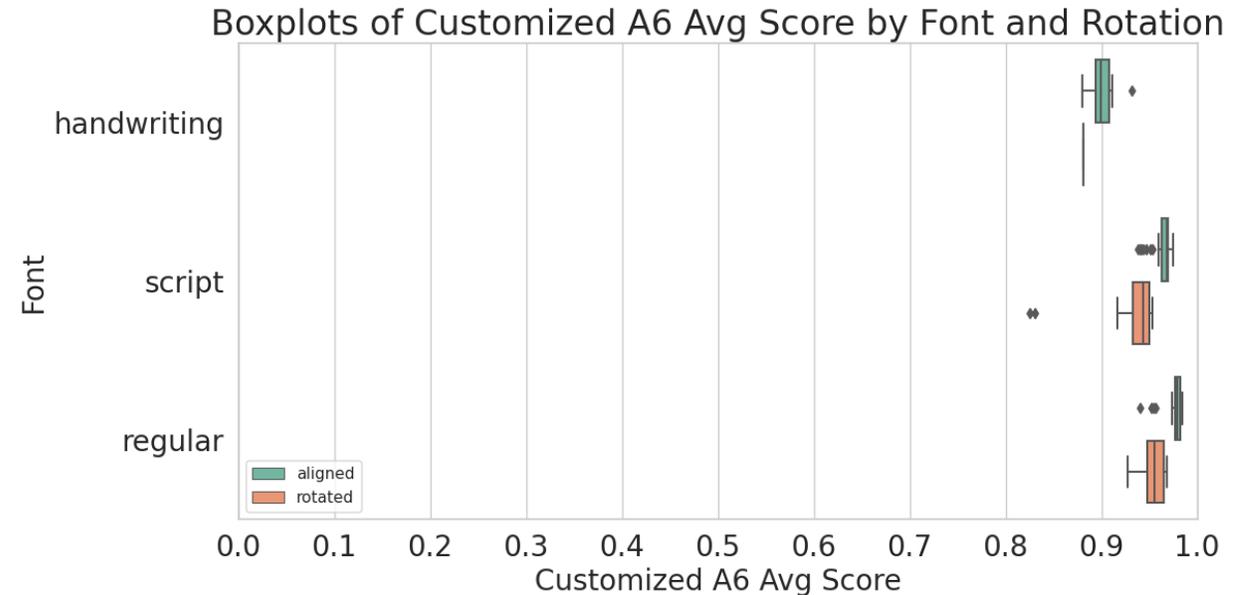
Statistics of Processing Entire PDF (Multiple Pages) – Rotated Forms



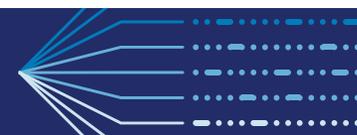
Rotation was +/- 10 degrees without additional correction

Average accuracy scores are 95% for rotated forms with regular font.

The average accuracy goes down to 94% for script font.



rotation	font	count	mean	std	min	25%	50%	75%	max
rotated	regular	40	0.954	0.011	0.927	0.948	0.955	0.965	0.968
rotated	script	40	0.936	0.027	0.826	0.933	0.944	0.95	0.953
rotated	handwriting	1	0.882		0.882	0.882	0.882	0.882	0.882



Custom Solution Improvements: Extended Bounding Boxes and Reading Order Correction



- Challenges for hand-writing related to text within the available writing space:
 - text can overflow over typical area and extend into margins
 - text that is not aligned across the rows
- Solved overflow issues by extending bounding boxes
- Reading order obtained from OCR (Optical Character Recognition) is corrected

Example : Question 45 – words from both lines were intermixed before correction

44. § 112.150: Record-keeping
 Observation Corrective action taken
 Description:
Farm owners must maintain records of food safety inspections, tests and other related activities.

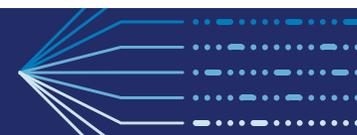
Records (21 CFR Part 112, Subpart O)

45. § 112.161 - 112.167: General record-keeping
 Observation Corrective action taken
 Description:
Maintaining records of food safety practices incl. location, dates and personnel involved.

Other Observations

46. Other
 Observation Corrective action taken
 Description:

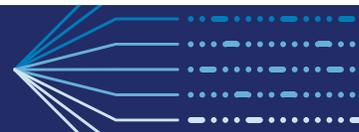
Page	Section	Key	Subkey	Checkbox	Value predicted
7	Records (21 CFR Part 112, Subpart O)	45. § 112.161 - 112.167: General record-keeping	Corrective action taken	TRUE	unselected
7	Records (21 CFR Part 112, Subpart O)	45. § 112.161 - 112.167: General record-keeping	Description	FALSE	Maintaining records of food safety practices incl. location, dates and personnel invol



Multiple Pages with Additional Observation and/or Comments



- Challenge: There can be multiple copies of “Additional Comments” pages added by user
- Conventional form parsing relies on fixed page count
- Solved this challenge using custom solution



Multiple Pages with Additional Comments Page – All Comments are Included in a Single Output



1st Page 8

2nd Page 8

Continuation Sheet

Continuation Sheet

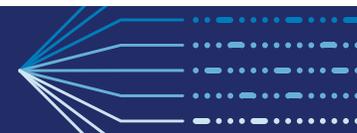
Additional Observations and/or Comments

Comments on Sunny Farms. It's very sunny here. They could put some shade, so the FDA Rep. doesn't bake in the sun. Some orange juice with ice would also help.

Additional Observations and/or Comments

Sunny Farms is very nice on a mild day.

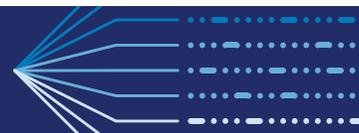
Page	Section	Key	Subkey	CheckBox	Value predicted	Value actual
8	Continuation Sheet	Additional Observations and/or Comments		FALSE	Comments on Sunny Farms, It's very sunny here. They could put some shade, so the FDA Rep. doesn't bake in the sun. Some orange juice with ice would also help. Sunny Farms is very nice on a mild day.	Comments on Sunny Farms. It's very sunny here. They could put some shade, so the FDA Rep. doesn't bake in the sun. Some orange juice with ice would also help. Sunny Farms is very nice on a mild day.



Text Spanning across Multiple Pages for the Same Question



- Challenge: There can be text spanning across two pages when the “Description” is at the end of a page
- Solved this challenge using custom solution



Text Spanning across Multiple Pages for the Same Question are Parsed Correctly



Health and Hygiene (21 CFR Part 112, Subpart D)

4. § 112.31: Measures to prevent ill or infected persons from contaminating covered produce with microorganisms of public health significance
 Observation Corrective action taken
 Description:

5. § 112.32: Hygienic practices of personnel
 Observation Corrective action taken
 Description:

6. § 112.33: Measures to prevent visitors from contaminating covered produce and food contact surfaces with microorganisms of public health significance
 Observation Corrective action taken
 Description: *Warning given.*

FORM FDA 4056 (01/19) Page of

← **Page Break!**

FARM NAME (include business name, if different)

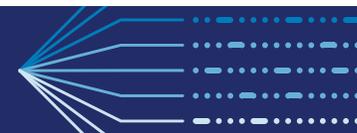
DATE(S) OF INSPECTION FEI NUMBER

It will be checked again in two days.

Agricultural Water (21 CFR Part 112, Subpart E)

7. § 112.41: Quality of agricultural water
 Observation Corrective action taken
 Description:

Page	Section	Key	Subkey	Check Box	Value predicted	Value actual
2	Health and Hygiene (21 CFR Part 112, Subpart D)	6. § 112.33: Measures to prevent visitors from contaminating covered produce and food contact surfaces with microorganisms of public health significance	Corrective action taken	TRUE	selected	X
2	Health and Hygiene (21 CFR Part 112, Subpart D)	6. § 112.33: Measures to prevent visitors from contaminating covered produce and food contact surfaces with microorganisms of public health significance	Description	FALSE	Warning given. It will be checked again in two days.	Warning given. It will be checked again in two days.



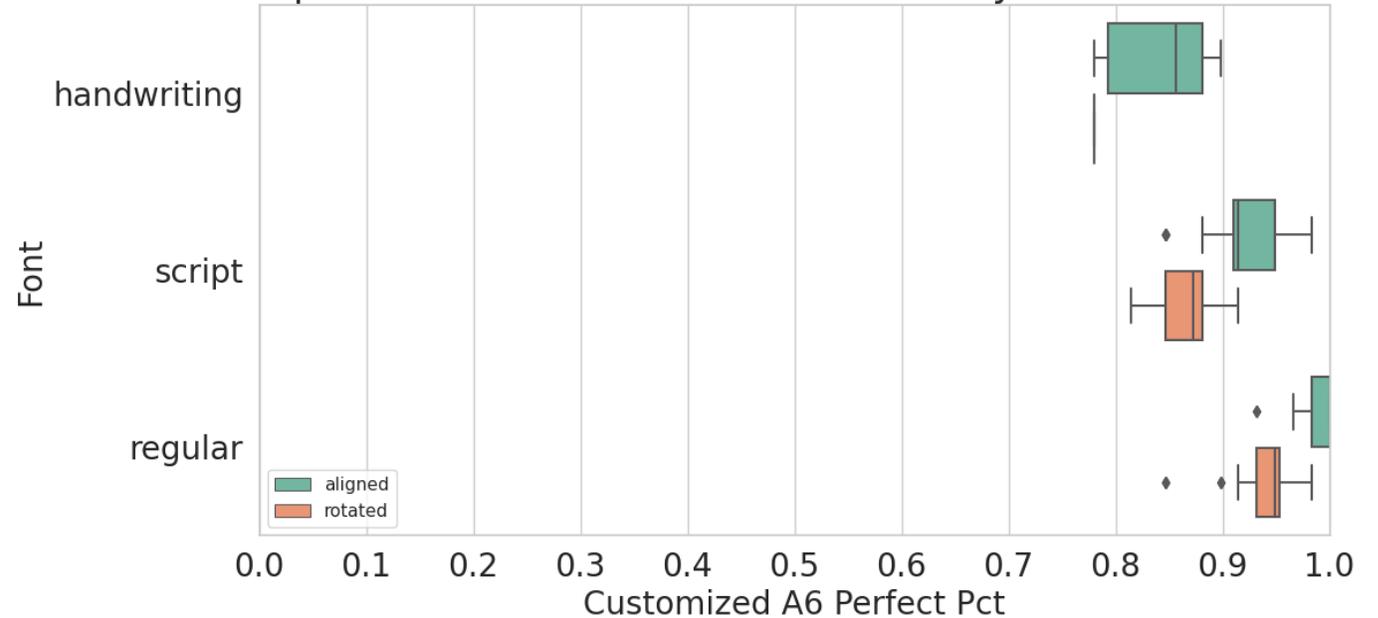
Performance by Perfectly Parsed Questions & Answers for Text Fields Only

In average, (for text fields only)

- 99% of the forms with regular font and aligned layout
- 92% from the forms with script font and aligned layout
- 84% from the forms with handwriting and aligned layout are perfectly parsed.

These averages drop down by 5-6% when the forms are rotated/distorted due to scanning.

Boxplots of Customized A6 Perfect Pct by Font and Rotation



rotation	font	Count	mean	std	min	25%	50%	75%	max
aligned	regular	40	0.989	0.017	0.932	0.983	1	1	1
	script	40	0.921	0.035	0.847	0.911	0.915	0.949	0.983
	handwriting	6	0.842	0.053	0.78	0.793	0.856	0.881	0.898
rotated	regular	40	0.94	0.032	0.847	0.932	0.949	0.953	0.983
	script	40	0.867	0.029	0.814	0.847	0.872	0.881	0.915
	Handwriting	1	0.78		0.78	0.78	0.78	0.78	0.78

Findings



- FDA's MLaaS offering is a strong foundation to be able to read documents and extract data from them.
- For this Form 4056 use case, MLaaS - and staff versed in the offering - we can account for a lot of variability in how the form is submitted and extract its data.

Questions

- We can extract, but where is the data stored for analysis?
- What amount of poorly read data is acceptable?

