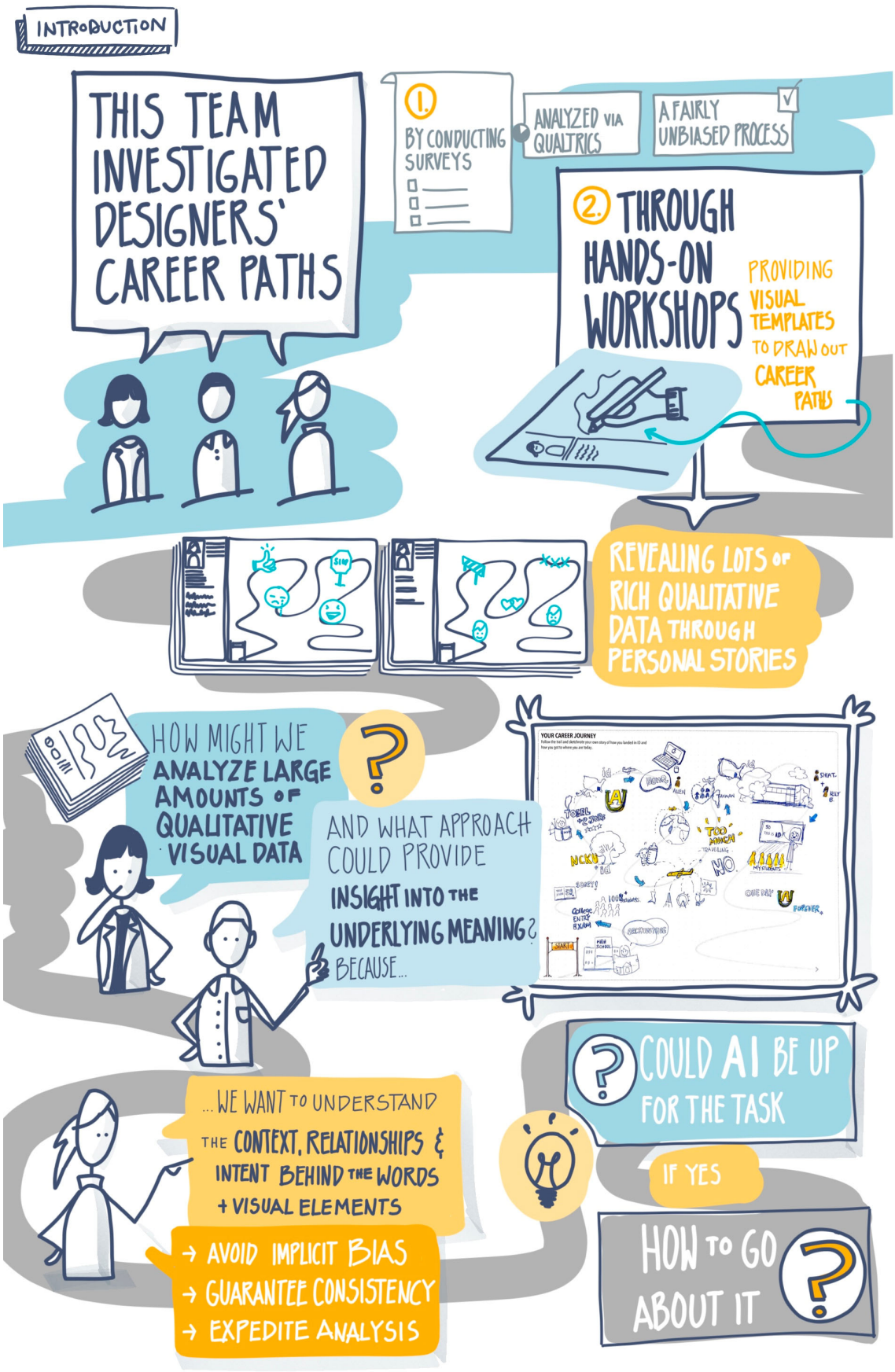
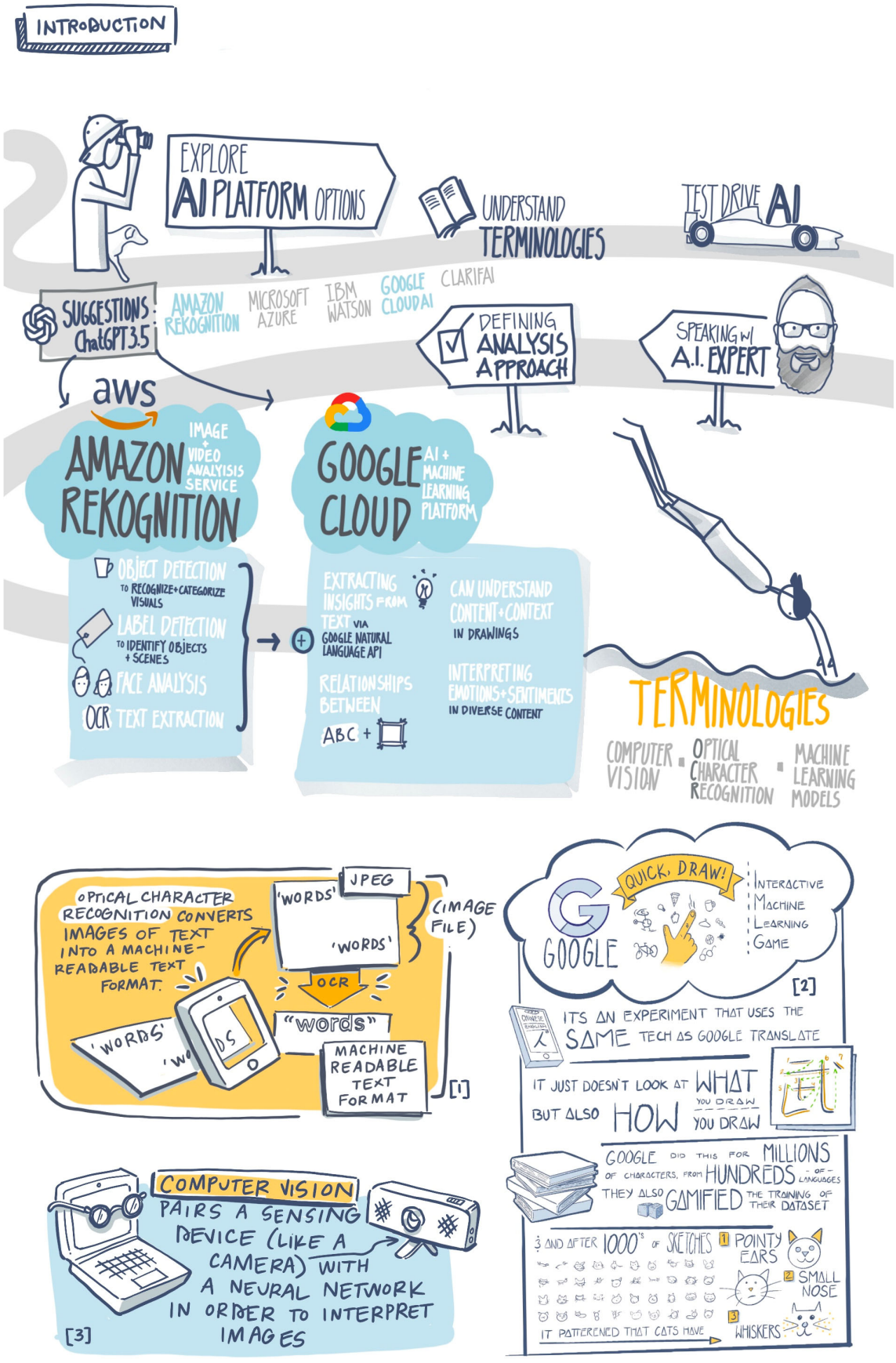
26TH INTERNATIONAL CONFERENCE ON ENGINEERING AND PRODUCT DESIGN EDUCATION
5-6 SEPTEMBER 2024, ASTON UNIVERSITY, BIRMINGHAM, UNITED KINGDOM

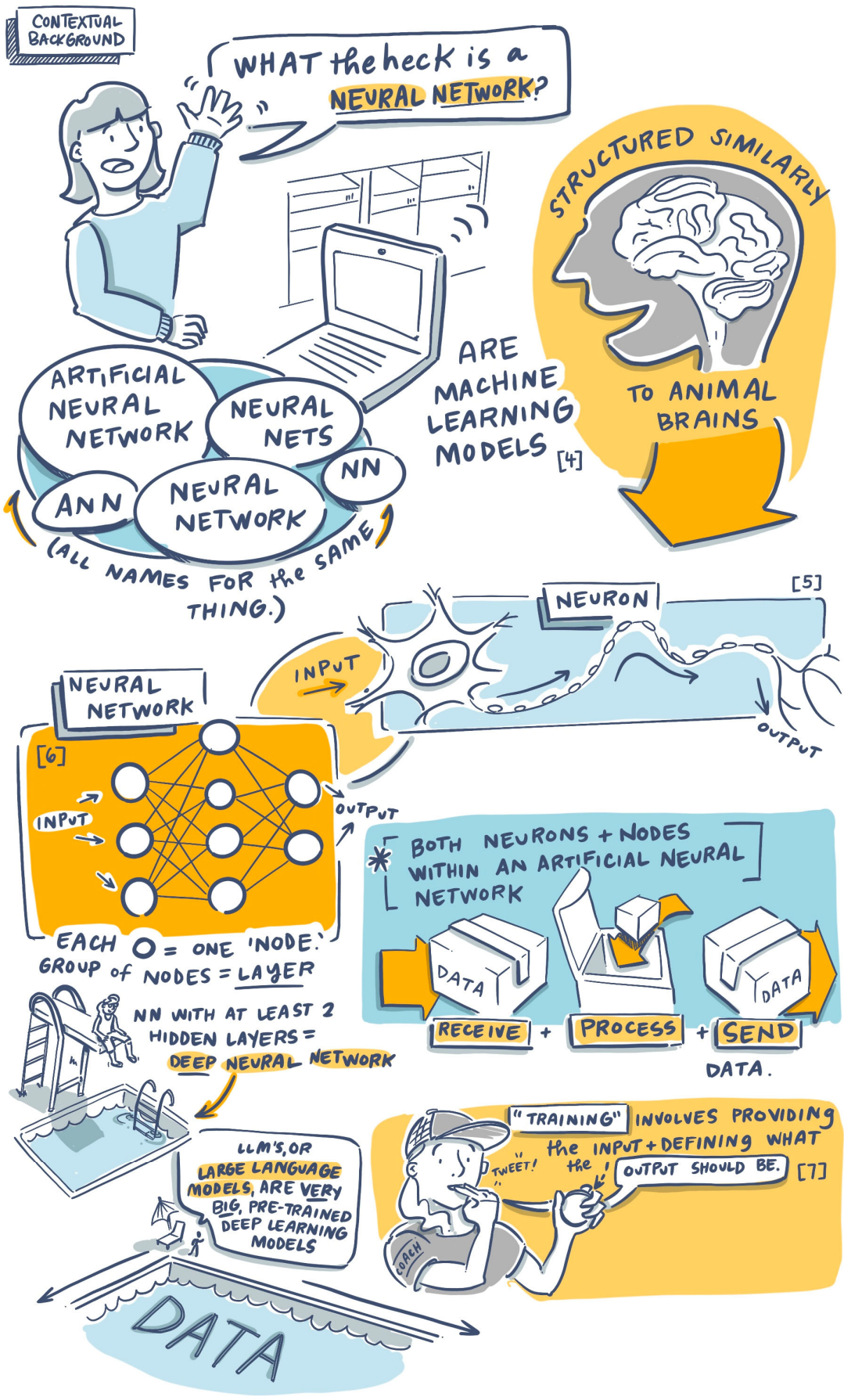
READING VISUALS: EXPLORING THE USE OF AI IN QUALITATIVE RESEARCH ANALYSIS

Carly HAGINS[1], **Verena PAEPCKE-HJELTNESS**[2] **and Benjamin BUSH**[2]
[1]University of Kentucky, United States of America
[2]Auburn University, United States of America

Keywords: Qualitative research, analysis, AI, visual template


INTRODUCTION
THIS TEAM INVESTIGATED DESIGNERS' CAREER PATHS
1. BY CONDUCTING SURVEYS
ANALYZED VIA QUALTRICS
A FAIRLY UNBIASED PROCESS
2. THROUGH HANDS-ON WORKSHOPS
PROVIDING VISUAL TEMPLATES TO DRAW OUT CAREER PATHS
REVEALING LOTS OF RICH QUALITATIVE DATA THROUGH PERSONAL STORIES
HOW MIGHT WE ANALYZE LARGE AMOUNTS OF QUALITATIVE VISUAL DATA
AND WHAT APPROACH COULD PROVIDE INSIGHT INTO THE UNDERLYING MEANING? BECAUSE...
YOUR CAREER JOURNEY
...WE WANT TO UNDERSTAND THE CONTEXT, RELATIONSHIPS & INTENT BEHIND THE WORDS + VISUAL ELEMENTS
→ AVOID IMPLICIT BIAS
→ GUARANTEE CONSISTENCY
→ EXPEDITE ANALYSIS
COULD AI BE UP FOR THE TASK
IF YES
HOW TO GO ABOUT IT


INTRODUCTION
EXPLORE AI PLATFORM OPTIONS
UNDERSTAND TERMINOLOGIES
TEST DRIVE AI
SUGGESTIONS: ChatGPT 3.5
AMAZON REKOGNITION
MICROSOFT AZURE
IBM WATSON
GOOGLE CLOUD AI
CLARIFAI
DEFINING ANALYSIS APPROACH
SPEAKING W/ A.I. EXPERT
aws
AMAZON REKOGNITION
IMAGE + VIDEO ANALYSIS SERVICE
OBJECT DETECTION
TO RECOGNIZE + CATEGORIZE VISUALS
LABEL DETECTION
TO IDENTIFY OBJECTS + SCENES
FACE ANALYSIS
OCR TEXT EXTRACTION
GOOGLE CLOUD
AI + MACHINE LEARNING PLATFORM
EXTRACTING INSIGHTS FROM TEXT VIA GOOGLE NATURAL LANGUAGE API
CAN UNDERSTAND CONTENT + CONTEXT IN DRAWINGS
RELATIONSHIPS BETWEEN
INTERPRETING EMOTIONS + SENTIMENTS IN DIVERSE CONTENT
TERMINOLOGIES
COMPUTER VISION
OPTICAL CHARACTER RECOGNITION
MACHINE LEARNING MODELS
OPTICAL CHARACTER RECOGNITION CONVERTS IMAGES OF TEXT INTO A MACHINE-READABLE TEXT FORMAT.
'WORDS'
JPEG
(IMAGE FILE)
OCR
"words"
MACHINE READABLE TEXT FORMAT
[1]
GOOGLE
QUICK, DRAW!
INTERACTIVE MACHINE LEARNING GAME
[2]
ITS AN EXPERIMENT THAT USES THE SAME TECH AS GOOGLE TRANSLATE
IT JUST DOESN'T LOOK AT WHAT YOU DRAW BUT ALSO HOW YOU DRAW
GOOGLE DID THIS FOR MILLIONS OF CHARACTERS, FROM HUNDREDS OF LANGUAGES
THEY ALSO GAMIFIED THE TRAINING OF THEIR DATASET
AND AFTER 1000'S OF SKETCHES
1 POINTY EARS
2 SMALL NOSE
3 WHISKERS
IT PATTERNED THAT CATS HAVE
COMPUTER VISION
PAIRS A SENSING DEVICE (LIKE A CAMERA) WITH A NEURAL NETWORK IN ORDER TO INTERPRET IMAGES
[3]


CONTEXTUAL BACKGROUND
"WHAT the heck is a NEURAL NETWORK?"
ARTIFICIAL NEURAL NETWORK
NEURAL NETS
NN
ANN
NEURAL NETWORK
(ALL NAMES FOR the SAME THING.)
ARE MACHINE LEARNING MODELS [4]
STRUCTURED SIMILARLY
TO ANIMAL BRAINS
NEURON
[5]
INPUT
OUTPUT
NEURAL NETWORK
[6]
INPUT
OUTPUT
EACH O = ONE 'NODE'
GROUP of NODES = LAYER
NN WITH AT LEAST 2 HIDDEN LAYERS = DEEP NEURAL NETWORK
* BOTH NEURONS + NODES WITHIN AN ARTIFICIAL NEURAL NETWORK
DATA
DATA
RECEIVE + PROCESS + SEND DATA.
LLM'S, OR LARGE LANGUAGE MODELS, ARE VERY BIG, PRE-TRAINED DEEP LEARNING MODELS
"TRAINING" INVOLVES PROVIDING the INPUT + DEFINING WHAT the OUTPUT SHOULD BE. [7]
TWEET!
COACH
DATA

TEST DRIVE

OCR

OPTICAL CHARACTER RECOGNITION

WITH GOOGLE CLOUD AI

WE SEE THIS DATA AS *beautiful* ✧ NUANCED ✧ INSIGHTFUL

BUT

GOOGLE SEES IT AS "DARK DATA"!

UNSTRUCTURED, UNTAPPED OR NOT EASILY SEGMENTED DATA

GOOGLE USES 2 TYPES of OCR

DOCUMENT AI [8]

- OPTIMIZED FOR DOCUMENTS & FORMS

VERY GOOD @ EXTRACTING TEXT FROM GOOGLE OR MICROSOFT DOCS, SCANNED TEXT OR PDFS

CAN ALSO RECOGNIZE

$a^2+b^2=c^2$ MATH FORMULAS

& EXTRACT ☑ CHECK BOXES

GEARED TOWARDS QUALITY

CLOUD VISION [9]

- OPTIMIZED FOR IMAGES, HAND WRITING, AND VIDEO

VERY GOOD @ CLASSIFYING IMAGES

DISCERNING HANDWRITING

PROCESSING REAL TIME VIDEO DATA

CAN ALSO ACCESS MULTIPLE DATA BASES USING R.E.S.T. [REPRESENTATIONAL STATE TRANSFER]

GEARED TOWARDS QUANTITY

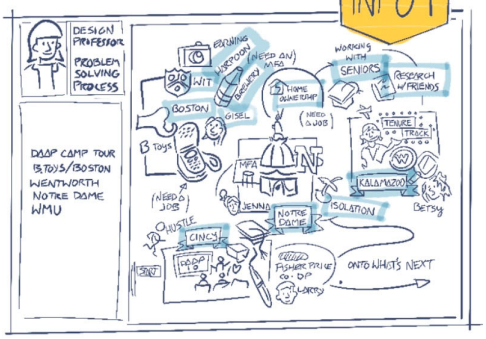
TEST DRIVE

OUTCOMES

FROM OUR ROADMAPS

DOCUMENT Δi

INPUT

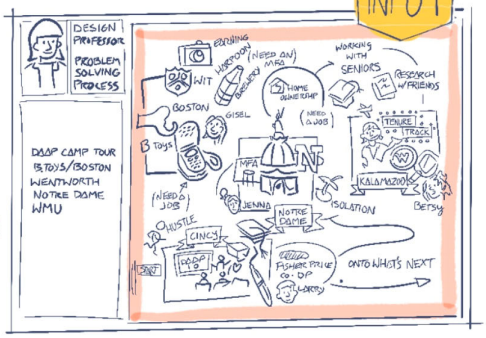


OUTPUT

- SENIORS
- HARPOON
- BREWERY
- BOSTON
- NOTRE DAME
- ISOLATION
- WIT
- RESEARCH W/ FRIENDS
- HOME OWNERSHIP
- GISELA
- CINCY
- KALAMAZOO

CLOUD VISION

INPUT



OUTPUT

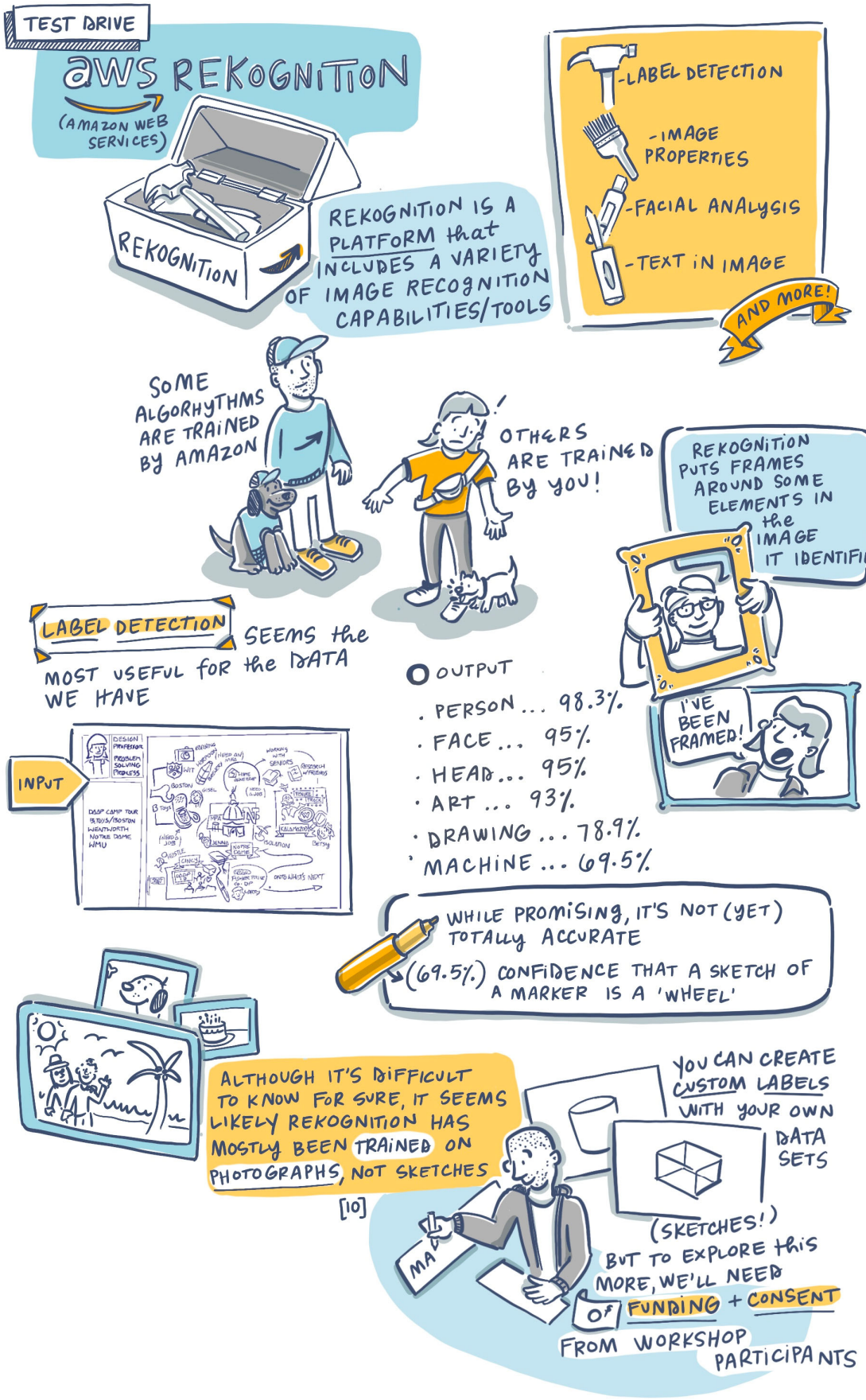
- PRODUCT 91%
- HANDWRITING 89%
- FONT 84%
- PARALLEL 73%
- ILLUSTRATION 67%

CLOSER TO OUR NEEDS
3 READS HANDWRITING WELL

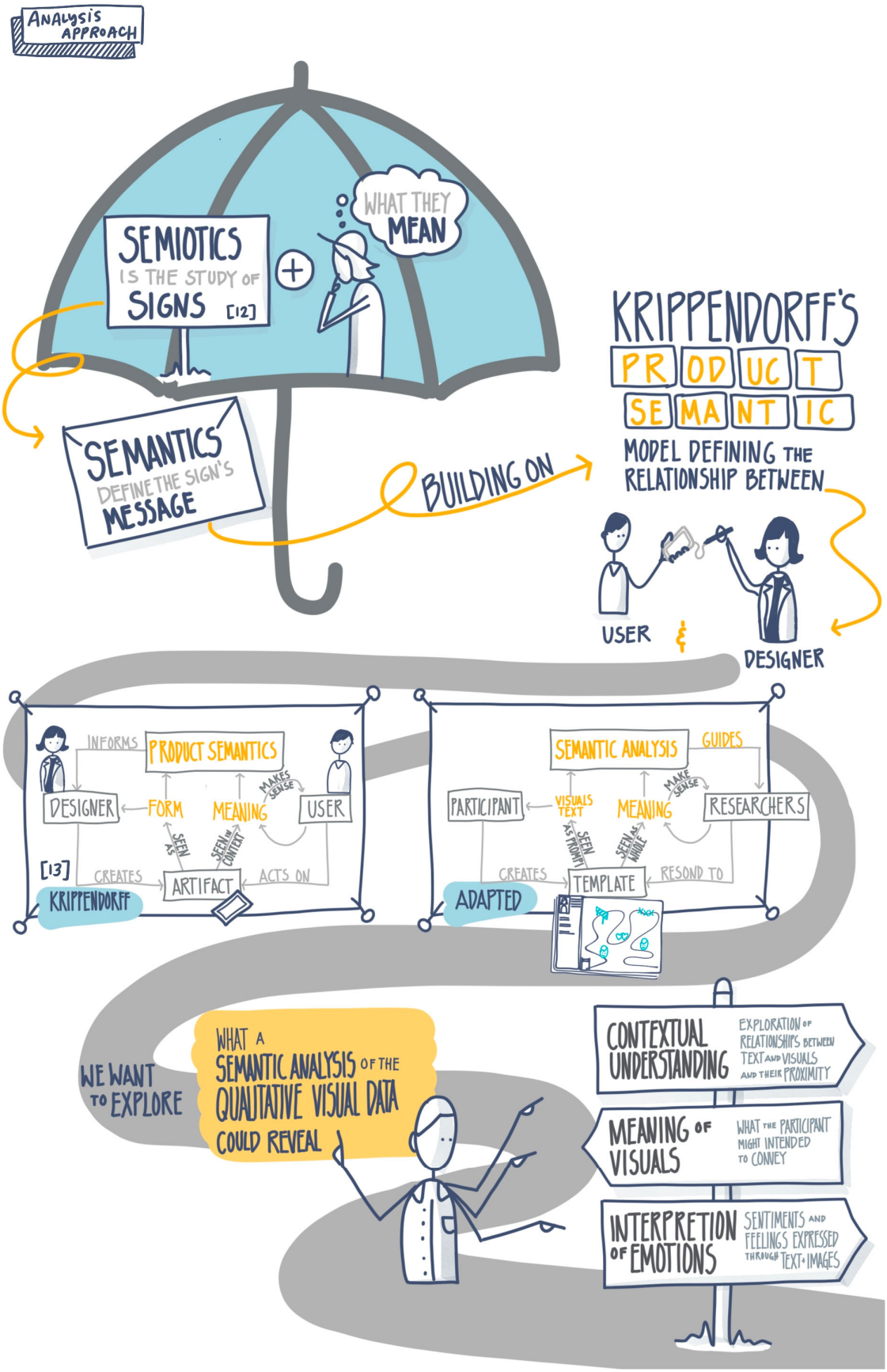
IT IS CORRECT BUT... IT READS THE WHOLE IMAGE RATHER THAN PORTIONS OF THE ROAD MAP

HOWEVER

...WE KNOW THAT BOTH Δi MODELS COULD BE TRAINED TO READ OUR ROADMAPS IF WE ADDED TIME, $, EFFORT 3 DATA


TEST DRIVE
aws REKOGNITION
(AMAZON WEB SERVICES)
REKOGNITION
REKOGNITION IS A PLATFORM that INCLUDES A VARIETY OF IMAGE RECOGNITION CAPABILITIES/TOOLS
-LABEL DETECTION
-IMAGE PROPERTIES
-FACIAL ANALYSIS
-TEXT IN IMAGE
AND MORE!
SOME ALGORHYTHMS ARE TRAINED BY AMAZON
OTHERS ARE TRAINED BY YOU!
REKOGNITION PUTS FRAMES AROUND SOME ELEMENTS IN the IMAGE IT IDENTIFIES
I'VE BEEN FRAMED!
LABEL DETECTION SEEMS the MOST USEFUL for the DATA WE HAVE
INPUT
OUTPUT
. PERSON ... 98.3%
. FACE ... 95%
. HEAD ... 95%
. ART ... 93%
. DRAWING ... 78.9%
. MACHINE ... 69.5%
WHILE PROMISING, IT'S NOT (YET) TOTALLY ACCURATE
(69.5%) CONFIDENCE THAT A SKETCH OF A MARKER IS A 'WHEEL'
ALTHOUGH IT'S DIFFICULT TO KNOW FOR SURE, IT SEEMS LIKELY REKOGNITION HAS MOSTLY BEEN TRAINED ON PHOTOGRAPHS, NOT SKETCHES [10]
YOU CAN CREATE CUSTOM LABELS WITH YOUR OWN DATA SETS (SKETCHES!)
BUT TO EXPLORE this MORE, WE'LL NEED FUNDING + CONSENT FROM WORKSHOP PARTICIPANTS


UNEXPECTED BARRIERS
DATA
INFO NUGGET
STACK OF QUESTIONS WITH NO ANSWERS
[11]
AS WE DIG FURTHER INTO THESE AI PROCESSING TECHNOLOGIES, WE DEVELOPED A HANDFUL OF QUESTIONS. AND THESE QUESTIONS... GREW INTO 'VERY REAL' CONCERNS
GOOD STUFF
SOLID INSIGHT
WE WERE WORRIED ABOUT...
WHERE IS THE DATA BEING SENT TO?
WHO OWNS THE DATA IF WE TRAIN "OUR OWN" AI MODEL?
DO WE HAVE AUTHORS' PERMISSION TO USE THE DATA THAT WAY
WE FELT A STRONG IMPERATIVE TO PAUSE & CONSIDER
BOUNDARIES
BEFORE WE OVERSTEP ANY ETHICAL • PRIVACY • DATA CONCERNS


ANALYSIS APPROACH
SEMIOTICS IS THE STUDY OF SIGNS [12]
WHAT THEY MEAN
SEMANTICS DEFINE THE SIGN'S MESSAGE
BUILDING ON
KRIPPENDORFF'S PRODUCT SEMANTIC MODEL DEFINING THE RELATIONSHIP BETWEEN
USER & DESIGNER
INFORMS
PRODUCT SEMANTICS
DESIGNER
FORM
MEANING
USER
MAKES SENSE
SEEN AS
SEEN IN CONTEXT
[13]
CREATES
ARTIFACT
ACTS ON
KRIPPENDORFF
SEMANTIC ANALYSIS
GUIDES
PARTICIPANT
VISUALS TEXT
MEANING
RESEARCHERS
MAKE SENSE
SEEN AS PROMPT
SEEN AS MIRROR
CREATES
TEMPLATE
RESOND TO
ADAPTED
WE WANT TO EXPLORE
WHAT A SEMANTIC ANALYSIS OF THE QUALITATIVE VISUAL DATA COULD REVEAL
CONTEXTUAL UNDERSTANDING
EXPLORATION OF RELATIONSHIPS BETWEEN TEXT AND VISUALS AND THEIR PROXIMITY
MEANING OF VISUALS
WHAT THE PARTICIPANT MIGHT INTENDED TO CONVEY
INTERPRETION OF EMOTIONS
SENTIMENTS AND FEELINGS EXPRESSED THROUGH TEXT+IMAGES


CONCLUSION
OUR HOPE WAS TO OFFLOAD ANALYSIS TO Ai
Ai
BUT THIS INVESTIGATION MADE IT CLEAR the BEST ANALYSIS WILL COME FROM
Ai AND PEOPLE WORKING TOGETHER
Ai TOOLS LIKE COMPUTER VISION CAN QUANTIFY QUALITATIVE DATA
REVEALING PATTERNS that WOULD BE DIFFICULT FOR A HUMAN to IDENTIFY
BUT (for now..) HUMAN LOGIC IS NEEDED to INTERPRET QUALITATIVE INFO TO SUPPORT SEMANTIC ANALYSIS
CONCEPTUAL UNDERSTANDING
MEANING of VISUALS
INTERPRETATION of VISUALS
SO WHAT NEXT?
DATA
LEARN MORE ABOUT WHO OWNS DATA UPLOADED TO SITES LIKE GOOGLE CLOUD Ai OR AWS REKOGNITION
then
BACK to the DRAWING BOARD to REVISE the TEMPLATE
CONSIDERING INTRODUCING STICKERS OR ICONS THAT WOULD FACILITATE Ai INTERPRETATION of COMPLETE TEMPLATES
THEN WE'LL DELIVER MORE WORKSHOPS!

REFERENCES

[1] *What is OCR (Optical Character Recognition)?* Available: https://aws.amazon.com/what-is/ocr/ [Accessed on 2024, 8 January].

[2] Burgess M. (16 November 2016). "You can now play a Pictionary-style game called Quick Draw against Google's AI". *Wired UK*. Wired.co.uk. Available: https://www.wired.co.uk/article/google-draw-ai-artificial-intelligence-experiments [Accessed 2024, 21 January].

[3] *What is Computer Vision?* Available: https://www.ibm.com/topics/computer-vision [Accessed on 2024, 8 January].

[4] *Machine Learning Tutorial*. Available: https://www.geeksforgeeks.org/machine-learning/?ref=gcse [Accessed on 2024, 06 January].

[5] Hardesty L. *Explained: Neural networks*. Available: https://news.mit.edu/2017/explained-neuralnetworks-deep-learning-0414 [Accessed on 2023, 21 December].

[6] *What are Convolutional Neural Networks?* Available: www.ibm.com/topics/convolutional-neural-networks [Accessed on 2024, 15 January].

[7] *Large Language Models Explained*. Available: https://www.nvidia.com/en-us/glossary/large-language-models/ [Accessed on 2023, 10 November].

[8] *What is Document AI?* Available: www.ibm.com/topics/convolutional-neural-networks [Accessed on 2024, 16 January].

[9] *What is Vision AI | Google Cloud?* Available: https://cloud.google.com/vision [Accessed on 2024, 18 January].

[10] *What is Amazon Rekognition?* Available: https://docs.aws.amazon.com/rekognition/latest/dg/what-is.html [Accessed on 2023, 18 December].

[11] Outkast. (2000). Bombs over Baghdad [MP3]. On *Stankonia*. Atlanta, Georgia; Antwan Patton, Andre Benjamin, David Sheats.

[12] Sunde H. Product Semantics. How design can affect user response and behavior, 2018. Available:https://www.ntnu.edu/documents/139799/1279149990/28+Artikkel+Henrik+Sunde.pdf/f9667e2e-7e9d-49e2-b1bd-a962ed35d7a8 [Accessed on 2024, 1 February]

[13] Krippendorff K. On the Essential Contexts of Artifacts or on the Proposition That «Design Is Making Sense (Of Things) ». *Design Issues, 5(2), 9-39,* 1989.