“What do I do with these images?”

A practical guide to the classification of images sent by survey participants

PATRICIA A. IGLESIAS | RECSM - UPF

CARLOS OCHOA | RECSM - UPF

MELANIE REVILLA | RECSM - UPF

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No849165), PI: Melanie Revilla.

Introduction

web data opp

- Images are getting more attention between survey methodologists and researchers in general.

- Because they have several potential advantages (Revilla, 2022):

	Expected benefits
Participants	Reduce time/efforts
	More enjoyable
Researchers	Avoid participants not knowing
	Avoid relying on remembering self
	Reduce human error
	Reduce satisficing
	Reduce effects of social desirability
	Data for concepts not measured so far

- Empirical evidence still limited but already some studies on the feasibility of asking respondents to share images in the frame of web surveys (Bosch et al., 2019; Ilić et al., 2020; Jäckle et al., 2019; Revilla et al., 2019; Wenz, 2017).

Introduction

- For the advantages to materialize:

    1) People need to accept to participate and successfully send the images
        - this depends on participants having the skills + availability + willingness to do it (Iglesias & Revilla, 2021)

    2) The information needs to be extracted from the images
        - Depending how this is done, more or less information can be obtained, and more or less accurate one
        - So this directly affects the quality of the results

- Process of extracting information and assigning labels to the items contained in an image = "classification" (Bandyopadhyay, 2022)

- Our goal = practical guide about how to deal with images required in the frame of web surveys to extract the best information possible

What type of items can be extracted from images?

Types of items to be extracted from images

Objects

Types of items to be extracted from images

Objects

Category:
dog, blanket, floor

Types of items to be extracted from images

Objects

Category:
dog, blanket, floor

Attributes:
white, soft, furry

Types of items to be extracted from images

Objects

Scenes

Types of items to be extracted from images

Objects

Scenes

Scene:
Parking lot

Types of items to be extracted from images

Objects

Scenes

Types of items to be extracted from images

Objects

Scenes

Scene:
Family meeting, celebration, dinner.

BACKGROUND

Types of items to be extracted from images

web data opp

Objects

Scenes

Text

	EUR
Banana	1,44 A
1,148 kg x 1,25 EUR/kg	
Freshona/Espinacas	1,15 A
Vemondo/Bebida soja 0%	1,60 B
2 x 0,80	
Vemondo/Tofu ecológico	0,95 B
Mandarina Ebre	2,79 A
Dentalux/Crema dental	0,95 C
Chef Select/Trio de humm	2,19 B
Edulis/Ensalada dúo	1,15 A
Alesto/Mezcla frutos sec	1,89 B
Floralys/Servill 2capas	0,95 C
Favorina/Huevos chocolat	1,49 B
Champiñón	0,65 A
Huevos L suelo	1,79 A
Floralys/Papel higiénico	2,55 C

Total	**21,54**
	============

Types of items to be extracted from images

Objects

Scenes

Text

	EUR	
Banana	1,44	A
1,148 kg x 1,25 EUR/kg		
Freshona/Espinacas	1,15	A
Vemondo/Bebida soja 0%	1,60	B
2 x 0,80		
Vemondo/Tofu ecológico	0,95	B
Mandarina Ebre	2,79	A
Dentalux/Crema dental	0,95	C
Chef Select/Trio de humm	2,19	B
Edulis/Ensalada dúo	1,15	A
Alesto/Mezcla frutos sec	1,89	B
Floralys/Servill 2capas	0,95	C
Favorina/Huevos chocolat	1,49	B
Champiñón	0,65	A
Huevos L suelo	1,79	A
Floralys/Papel higiénico	2,55	C
Total	**21,54**	

Focus:
Price per product

Types of items to be extracted from images

Objects

Scenes

Text

	EUR
Banana	1,44 A
1,148 kg x 1,25 EUR/kg	
Freshona/Espinacas	1,15 A
Vemondo/Bebida soja 0%	1,60 B
2 x 0,80	
Vemondo/Tofu ecológico	0,95 B
Mandarina Ebre	2,79 A
Dentalux/Crema dental	0,95 C
Chef Select/Trio de humm	2,19 B
Edulis/Ensalada dúo	1,15 A
Alesto/Mezcla frutos sec	1,89 B
Floralys/Servill 2capas	0,95 C
Favorina/Huevos chocolat	1,49 B
Champiñón	0,65 A
Huevos L suelo	1,79 A
Floralys/Papel higiénico	2,55 C

Total	21,54
	============

Focus:
Products

Types of items to be extracted from images

Objects

Scenes

Text

	EUR
Banana	1,44 A
1,148 kg x 1,25 EUR/kg	
Freshona/Espinacas	1,15 A
Vemondo/Bebida soja 0%	1,60 B
2 x 0,80	
Vemondo/Tofu ecológico	0,95 B
Mandarina Ebre	2,79 A
Dentalux/Crema dental	0,95 C
Chef Select/Trio de humm	2,19 B
Edulis/Ensalada dúo	1,15 A
Alesto/Mezcla frutos sec	1,89 B
Floralys/Servill 2capas	0,95 C
Favorina/Huevos chocolat	1,49 B
Champiñón	0,65 A
Huevos L suelo	1,79 A
Floralys/Papel higiénico	2,55 C

Total	21,54
	============

Focus:
Total expending

BACKGROUND

Types of items to be extracted from images

Objects | Scenes | Text | **More complex analysis**

	EUR
Banana	1,44 A
1,148 kg x 1,25 EUR/kg	
Freshona/Espinacas	1,15 A
Vemondo/Bebida soja 0%	1,60 B
2 x 0,80	
Vemondo/Tofu ecológico	0,95 B
Mandarina Ebre	2,79 A
Dentalux/Crema dental	0,95 C
Chef Select/Trio de humm	2,19 B
Edulis/Ensalada dúo	1,15 A
Alesto/Mezcla frutos sec	1,89 B
Floralys/Servill 2capas	0,95 C
Favorina/Huevos chocolat	1,49 B
Champiñón	0,65 A
Huevos L suelo	1,79 A
Floralys/Papel higiénico	2,55 C

Total	**21,54**
	============

BACKGROUND

Types of items to be extracted from images

Objects

Scenes

Text

More complex analysis

	EUR
Banana	1,44 A
1,148 kg x 1,25 EUR/kg	
Freshona/Espinacas	1,15 A
Vemondo/Bebida soja 0%	1,60 B
2 x 0,80	
Vemondo/Tofu ecológico	0,95 B
Mandarina Ebre	2,79 A
Dentalux/Crema dental	0,95 C
Chef Select/Trio de humm	2,19 B
Edulis/Ensalada dúo	1,15 A
Alesto/Mezcla frutos sec	1,89 B
Floralys/Servill 2capas	0,95 C
Favorina/Huevos chocolat	1,49 B
Champiñón	0,65 A
Huevos L suelo	1,79 A
Floralys/Papel higiénico	2,55 C
Total	21,54

Type of document:
Receipt

How can we classify items?

Options for image classification

MANUAL

Human classification

AUTOMATIC

Already existing algorithms

New algorithm developed externally

New algorithm developed internally

Options for image classification

MANUAL

Human classification

- Performed by 1 or more humans
- Classifier(s) observe images and assign labels
- Similar to coding open-narrative answers in surveys

AUTOMATIC

Already existing algorithms

New algorithm developed externally

New algorithm developed internally

Options for image classification

MANUAL	Human classification	• Performed by 1 or more humans • Classifier(s) observe images and assign labels • Similar to coding open-narrative answers in surveys
AUTOMATIC	Already existing algorithms	• General: for free or against payment, predefined or custom labels • Specific to a topic: eventual collaboration
	New algorithm developed externally	
	New algorithm developed internally	

Options for image classification

M A N U A L

Human classification

- Performed by 1 or more humans
- Classifier(s) observe images and assign labels
- Similar to coding open-narrative answers in surveys

A U T O M A T I C

Already existing algorithms

- General: for free or against payment, predefined or custom labels
- Specific to a topic: eventual collaboration

New algorithm developed externally

New algorithm developed internally

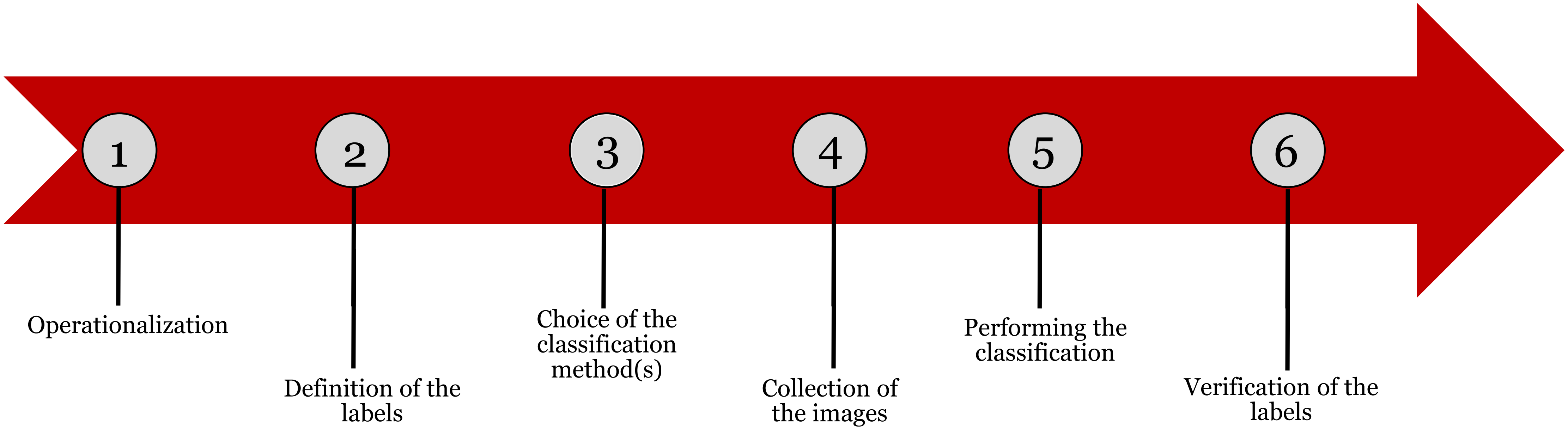
- Specific for the project needs.

Options for image classification

MANUAL	Human classification	• Performed by 1 or more humans • Classifier(s) observe images and assign labels • Similar to coding open-narrative answers in surveys
AUTOMATIC	Already existing algorithms	• General: for free or against payment, predefined or custom labels • Specific to a topic: eventual collaboration
	New algorithm developed externally	• A external provider can be contracted
		• Specific for the project needs.
	New algorithm developed internally	• Open libraries available for R and Python to classify images and read text

Image classification in surveys

Our proposal to classify images: a six-step process



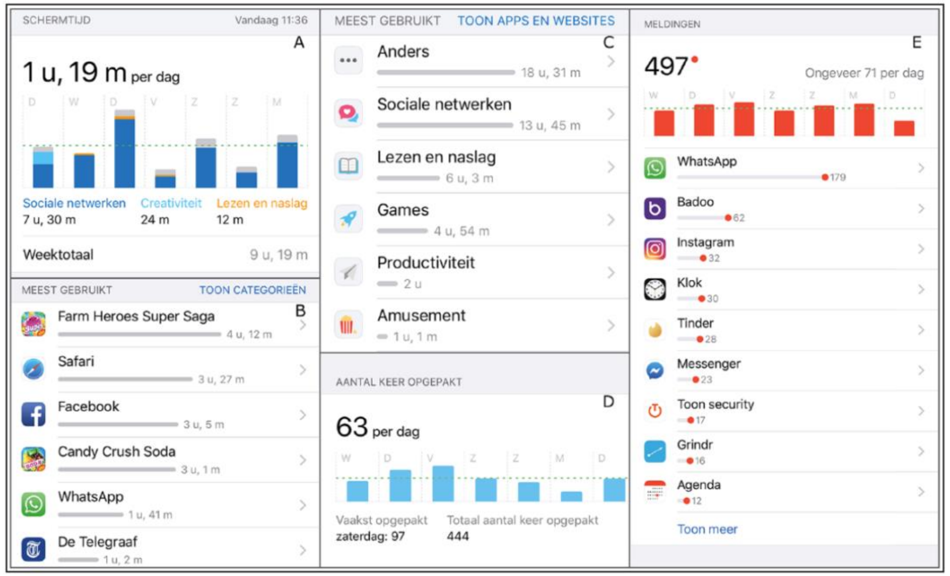
A SIX-STEP PROCESS

Operationalization

Step 1

- Definition of the type(s) of item to be classified.
  - Categories and/or attributes

- Delimitation of the item(s) to be classified.
  - Some or all items in the image.



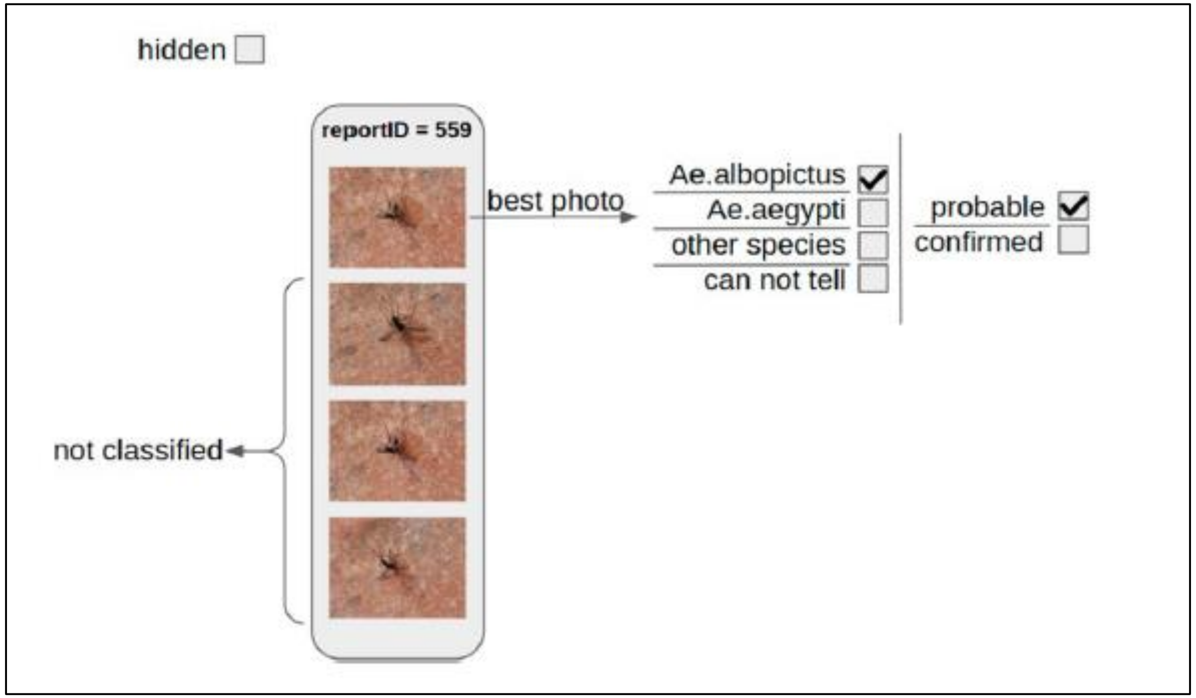
Ohme et al. (2020)

A SIX-STEP PROCESS

Definition of the labels

Step 2

- Think about the exact labels to be used → definition of the response scale.



Pataki et al. (2021)

Choice of the classification method(s)

Step 3

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

web data opp

Step 3

Factors for choosing the classification model

Classification tasks	Resources	Overall data quality
Total number of images	Human resources	Accuracy
Total number of labels	Infrastructure	Consistency
Kind of labels	Budget	Ethics
	Availability of images	Transparency

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Classification tasks**

- Total number of images
    - **If low (<500)**
    - If more

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Classification tasks**

- Total number of images
  - If low (<500)
  - **If more**

41,900% faster and 21,671% cheaper (Bosch et al., 2019)

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Classification tasks**

- Total number of images
  - If low (<500)
  - If more

- Total number and kind of labels
  - **Large number of labels**
  - Specific labels

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Classification tasks**

  - Total number of images
    - If low (<500)
    - If more
  - Total number and kind of labels
    - Large number of labels
    - **Specific labels**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Resources**

  - Human resources
    - Manual classification: training of classifiers.
    - External algorithm: one person to check labels and one with technical knowledge.
    - Internal algorithm: highly specialized profile.

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Resources**

- Human resources
  - Manual classification: training of classifiers.
  - External algorithm: one person to check labels and one with technical knowledge.
  - Internal algorithm: highly specialized profile.

- Infrastructure
  - **In absence of proper hardware.**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

webdataopp

Step 3

- **Resources**
  - Human resources
    - Manual classification: training of classifiers.
    - External algorithm: one person to check labels and one with technical knowledge.
    - Internal algorithm: highly specialized profile.
  - Infrastructure
    - In absence of proper hardware.
  - Cost
    - **Low budget**
    - Frequency of use will be an important factor

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Resources**

- Human resources
  - Manual classification: training of classifiers.
  - External algorithm: one person to check labels and one with technical knowledge.
  - Internal algorithm: highly specialized profile.
- Infrastructure
  - In absence of proper hardware.
- Cost
  - Low budget
  - Frequency of use will be an important factor
- Availability of images
  - **If images to train a new model are not available**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Overall data quality**

- **Accuracy**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

3

- **Overall data quality**
  - Accuracy
  - **Consistency**
    - **When using only one method**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Overall data quality**
  - Accuracy
  - Consistency
    - When using only one method
  - **Data protection**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Overall data quality**
  - Accuracy
  - Consistency
    - When using only one method
  - Data protection
  - **Transparency**

Human classification

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

AUTOMATIC

Choice of the classification method(s)

Step 3

- **Examples**

Studying incomes and expenditures every 5 years

Classification tasks	Resources	Overall data quality	
High number and complexity of labels. High number of images.	Have images from previous waves. Have hardware and specialized human resources.	Found low accuracy and consistency with human classification. Sensitive information.	**New algorithm developed internally**

Choice of the classification method(s)

Step 3

- **Examples**

Studying recycling, collection of 1,000 images containing 10,000 items

Classification tasks	Resources	Overall data quality	
Specific labels.	Low budget. No specialized human resources. No infrastructure.	Low accuracy when using an already existing algorithm.	**Human classification**

Choice of the classification method(s)

Step 3

- **Examples**

Studying recycling, collection of 1,000 images containing 10,000 items

Classification tasks	Resources	Overall data quality	
Specific labels.	Low budget. No specialized human resources. No infrastructure.	Low accuracy when using an already existing algorithm.	**Human classification**
Specific labels.	Low budget No specialized human resources No infrastructure	**Higher accuracy** based on improving the external algorithm.	**Already existing algorithm (custom labels)**

Choice of the classification method(s)

webdataopp

Step 3

- **Examples**

Studying recycling, collection of 1,000 images containing 10,000 items

Classification tasks	Resources	Overall data quality	
Specific labels.	Low budget. No specialized human resources. No infrastructure.	Low accuracy when using an already existing algorithm.	**Human classification**
Specific labels.	Low budget No specialized human resources No infrastructure	**Higher accuracy** based on improving the external algorithm.	**Already existing algorithm (custom labels)**
Specific labels.	**Specialized human resources. Have the infrastructure. At first no images to train the model.**	High accuracy.	**Manual first, but new algorithm developed internally later**

Collection of the images

Step 4

- **Main considerations**

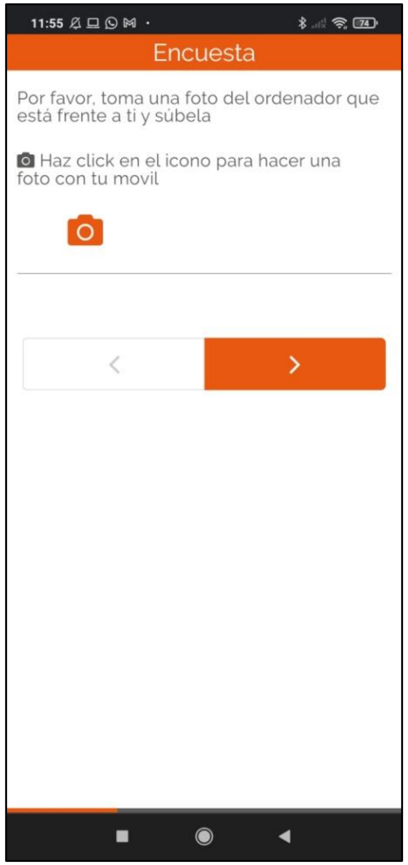
  - Storage of images
    - Will images be deleted?
    - Is storage safe?
  - ID to relate to each participant
  - Size of the files
    - Important for storage and uploading
  - Processing sensitive data and metadata
  - Informed consent (intellectual property, what not to upload)
  - Tools to collect the images

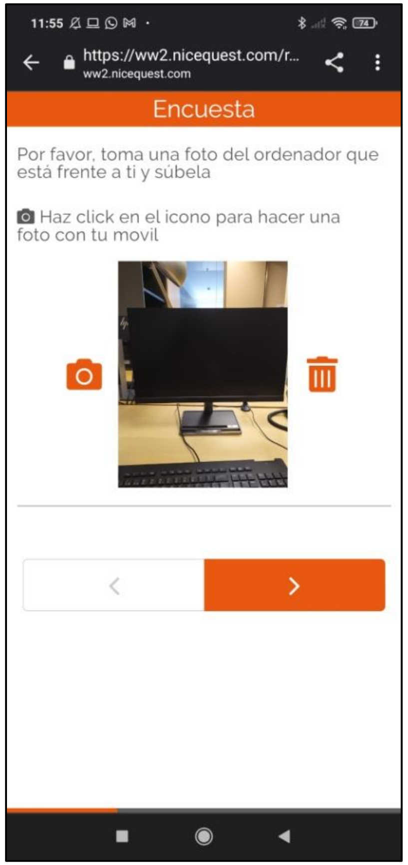
Collection of the images

webdata opp

Step 4

- **Already available tools: webdataVisual**



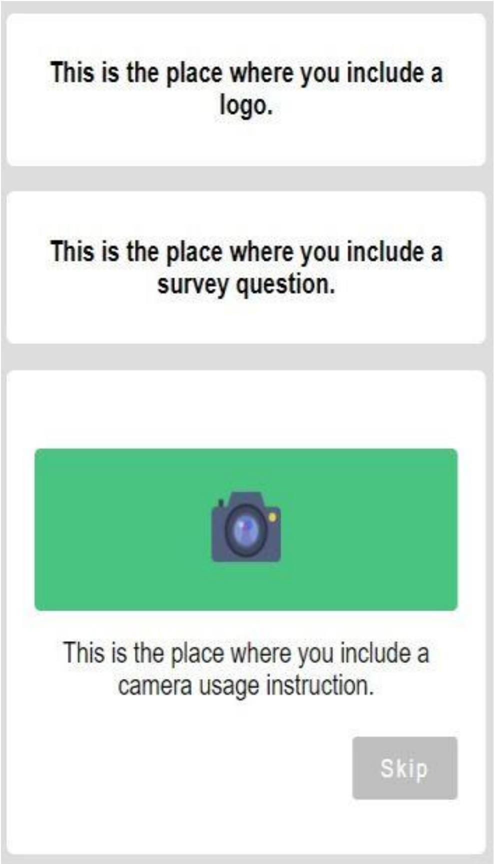


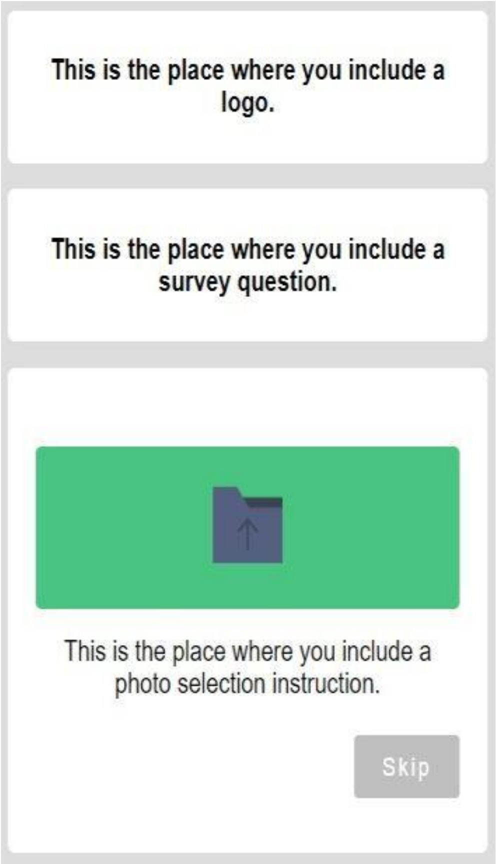
More information available at https://www.upf.edu/web/webdataopp

Collection of the images

Step 4

- **Already available tools: SurveyImage**





Höhne, Qureshi & Gavras (2020).

Performing the classification

Step 5

- **Preliminary checkups**

- Look for inappropriate and conflictive images (e.g., privacy issues).
- Assess if there are other potential labels to be used.
- Image enhancement.

Performing the classification

Step 5

- **Preliminary checkups**

- Look for inappropriate and conflictive images (e.g., privacy issues).
- Assess if there are other potential labels to be used.
- Image enhancement.

Human classification

- Classification guidelines
- Training of classifiers
- Availability of the necessary infrastructure

Already existing algorithm

New algorithm developed externally

New algorithm developed internally

Performing the classification

Step 5

- **Preliminary checkups**

- Look for inappropriate and conflictive images (e.g., privacy issues).
- Assess if there are other potential labels to be used.
- Image enhancement.

Human classification

- Classification guidelines
- Training of classifiers
- Availability of the necessary infrastructure

Already existing algorithm

- Choose the more appropriate one, or can use more than one.
- Check the labels provided
- If available, upload images to the API

New algorithm developed externally

New algorithm developed internally

Performing the classification

Step 5

- **Preliminary checkups**

- Look for inappropriate and conflictive images (e.g., privacy issues).
- Assess if there are other potential labels to be used.
- Image enhancement.

Human classification
- Classification guidelines
- Training of classifiers
- Availability of the necessary infrastructure

Already existing algorithm
- Choose the more appropriate one, or can use more than one.
- Check the labels provided
- If available, upload images to the API

New algorithm developed externally
- Decide who is going to develop the algorithm → guidelines.
- Make sure the algorithm works properly (it is well trained) and provides results as expected.

New algorithm developed internally

Performing the classification

webdataopp

Step 5

- **Preliminary checkups**
  - Look for inappropriate and conflictive images (e.g., privacy issues).
  - Assess if there are other potential labels to be used.
  - Image enhancement.

Human classification
- Classification guidelines
- Training of classifiers
- Availability of the necessary infrastructure

Already existing algorithm
- Choose the more appropriate one, or can use more than one.
- Check the labels provided
- If available, upload images to the API

New algorithm developed externally
- Decide who is going to develop the algorithm → guidelines.
- Make sure the algorithm works properly (it is well trained) and provides results as expected.

New algorithm developed internally
- Development of the algorithm
- Training of the model (availability of images)

Verification of the labels

Step 6

- **Within each method**

- Human classification: swapping images between classifiers.
- Automatic models: manual verification.

- Suggestions: swap at least 30 images to check accuracy (Zhang et al., 2022)

- **If using more than one method**

- Contrast results between them.

Verification of the labels

Step 6

- **Checking with participants**

- Participants can access the labels and see if they are correctly classified.
- Example: “cleansing gel” in a receipt.

Verification of the labels

Step 6

- **Checking with participants**

- Participants can access the labels and see if they are correctly classified.
- Example: “cleansing gel” in a receipt.

Only for research where participants know the content and are expected to see the produced labels.

Summary

web data opp

1. Even when working with images, the **operationalization** process must be preserved as when posing any type of question (definition of the concept of interest, response scale, codes)

2. Defining the **labels** (similarly to codes in open-ended questions) allow knowing what will be measured on each image, and determine the third step (choice of the classification method)

3. The factors to **choose a classification method** favor one over the other. They should be weighted according to the project's characteristics and needs.

4. When **collecting** images, researchers must me able to link the images with the corresponding participant while protecting their privacy.

5. Preliminary checks must be made before **classifying**, and part of them will depend on the classification method chosen.

6. Labels, whether produced by human or automatic classification, should be **verified**.

Conclusions

Conclusions

- Proper assessment of whether images fit the topic/question.

- It is not easy to work with images.
  → There are many decisions to make during the process and factors to consider.
  → Things can go wrong.

- A precise definition of the information to be collected is crucial…

  … and a tool able to extract such information must be available.

- Verifications are always necessary.

- Still, images have the potential to provide new and/or better insights and to improve the overall respondents' experience

References

- Bandyopadhyay, H. (2022, January 10). *Image Classification Explained: An Introduction*. V7. https://www.v7labs.com/blog/image-classification-guide
- Bosch, O., Revilla, M., & Paura, E. (2019). Answering mobile surveys with images: An exploration using a computer vision API. *Social Science Computer Review*, *37*(5), 669–683. https://doi.org/10.1177/0894439318791515
- Höhne, J.K., Qureshi, D.D., & Gavras, K., (2020). *SurveyImage (SImage): A comprehensive guide for collecting images in surveys. Zenodo*. 10.5281/zenodo.4280543
- Iglesias, P. A., & Revilla, M. (2021, July 2). *When does it make sense to ask respondents for visual data? Insights for (mobile) web surveys* [Presentation]. ESRA Conference. https://www.europeansurveyresearch.org/conf2021/uploads/177/3602/50/210702_2_ESRA_presentation.pdf
- Ilić, G., Struminskaya, B., & Lugtig, P. (2020, November 13). *Giving respondents a choice: Does it increase sharing of sensor data?* BigSurv20 virtual conference. https://www.bigsurv20.org/conf20/uploads/162/217/BigSurv2020_BStruminskaya.pdf
- Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases. *Survey Research Methods*, *13*(1), 23–44. https://doi.org/10.18148/srm/2019.v1i1.7297
- Revilla, M. (2022). Four "new" opportunities to enhance or extend (mobile) web survey data. *Survey Research Methods*.
- Revilla, M., Couper, M. P., & Ochoa, C. (2019). Willingness of online panelists to perform additional tasks. *Methods, Data, Analyses*, *13*(2), 29. https://doi.org/10.12758/mda.2018.01
- Wenz, A. (2017). Completing web surveys on mobile devices: Does screen size affect data quality? *ISER Working Paper Series. Institute for Social and Economic Research*, *2017–05*, 1–23.
- Zhang, R., Gong, J., Ma, S., Firdaus, A., & Xu, J. (2022). Automatic Coding Mechanisms for Open-Ended Questions in Journalism Surveys: An Application Guide. *Digital Journalism*, *0*(0), 1–22. https://doi.org/10.1080/21670811.2022.2037006

Thanks!

Questions?

PATRICIA A. IGLESIAS| RECSM – UPF

patricia.iglesias@upf.edu

https://www.upf.edu/web/webdataopp