



The AAAI 2021 KDF Workshop

***“From insight
to impact”*** 

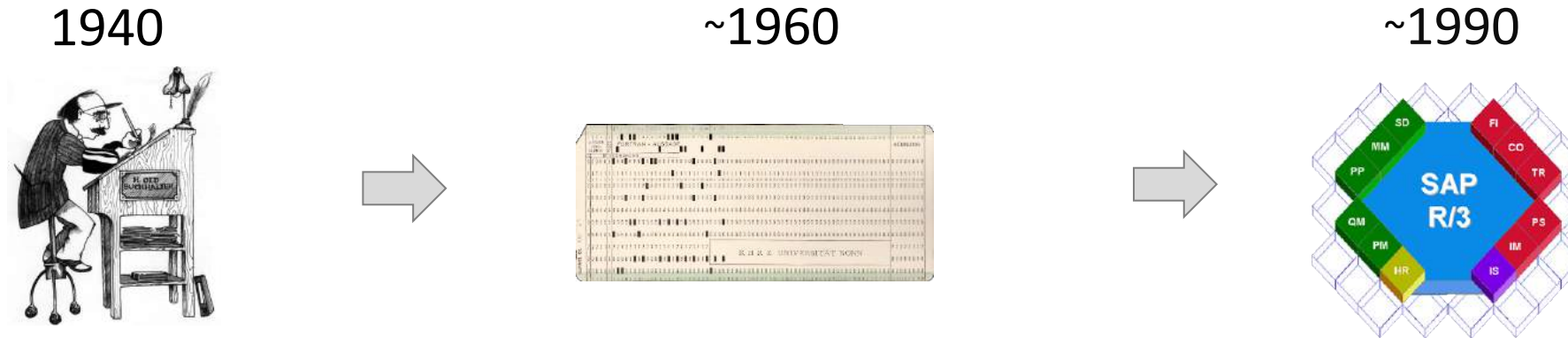
February 9th, 2021

Leaking Sensitive Accounting Data in Plain Sight using Deep Autoencoder Neural Networks

Marco Schreyer¹, Christian Schulze² and Damian Borth¹¹University of St.Gallen (HSG), ²German Research Center for Artificial Intelligence (DFKI)

Introduction & Background

Evolution of Recording and Processing of Enterprise Resource Planning Data



Data Volume

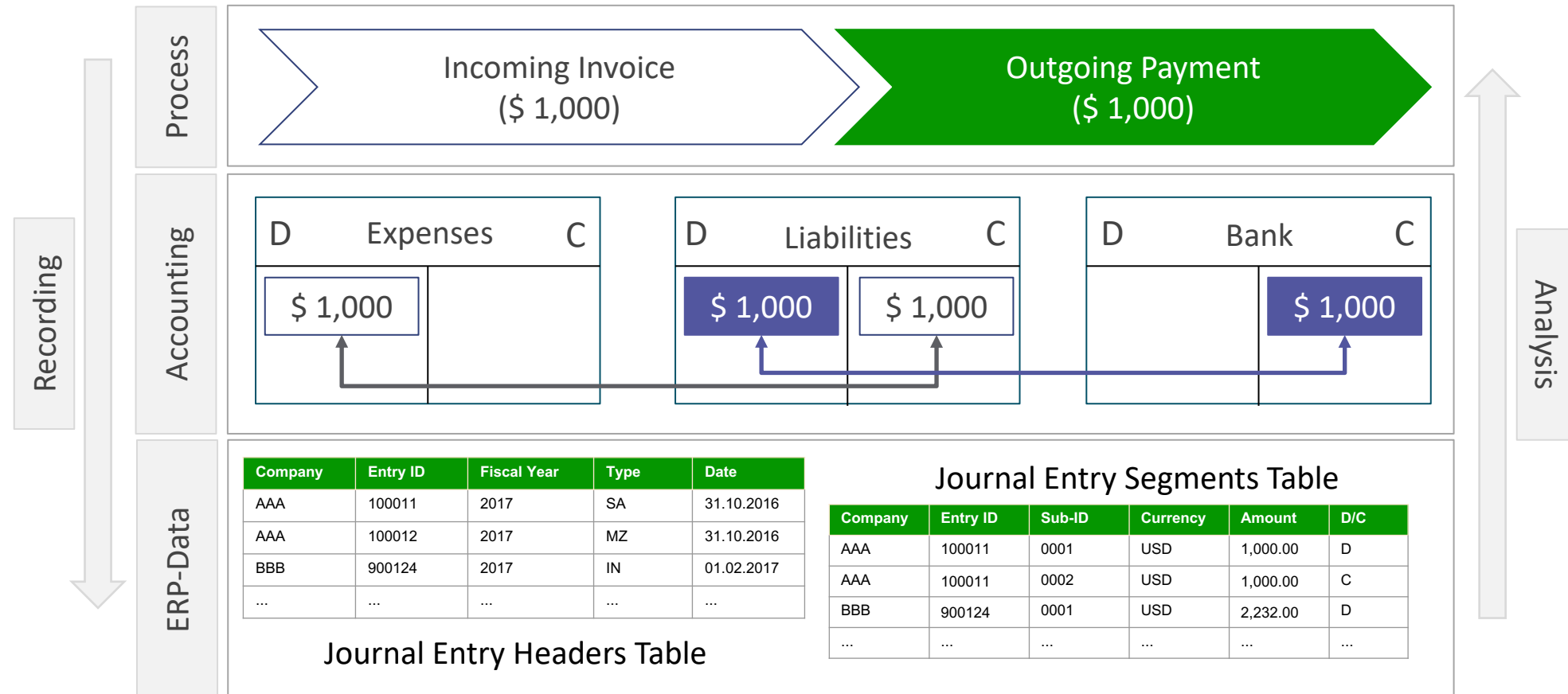
Towards the “Big Data” Driven Economy

- Displacement of the non-digital processing of organizational activities
- Accumulation of exhaustive volumes of transactional and accounting data
- Almost every activity within an organization leaves a **digital trace** ... !

Following: Vasarhelyi, M. A., Kogan, A., Tuttle, B. M., “Big Data in Accounting: An Overview”, Accounting Horizons, Vol. 29, No. 2, 2015
Warren Jr., J. D., Mofitt K. C., Byrnes, P., “How Big Data Will Change Accounting”, Accounting Horizons, Vol. 29, No. 2, 2015
Brown-Liburd, H. and Vasarhelyi, M. A., “Big Data and Audit Evidence”, Journal of Emerging Technologies in Accounting, Vol. 12, 2015

Introduction & Background

Evolution of Recording and Processing of Enterprise Resource Planning Data



Introduction & Background

The Rise of Insider Threats

Deloitte.

“Approx. 59% of employees who leave an organization voluntarily or involuntary say they take sensitive data with them.”

Deloitte Global - “The Rise of Insider Threats Amid COVID-19”
Weekly high-level brief update, Issue 6, May 12, 2020



EUROPEAN UNION AGENCY
FOR CYBERSECURITY

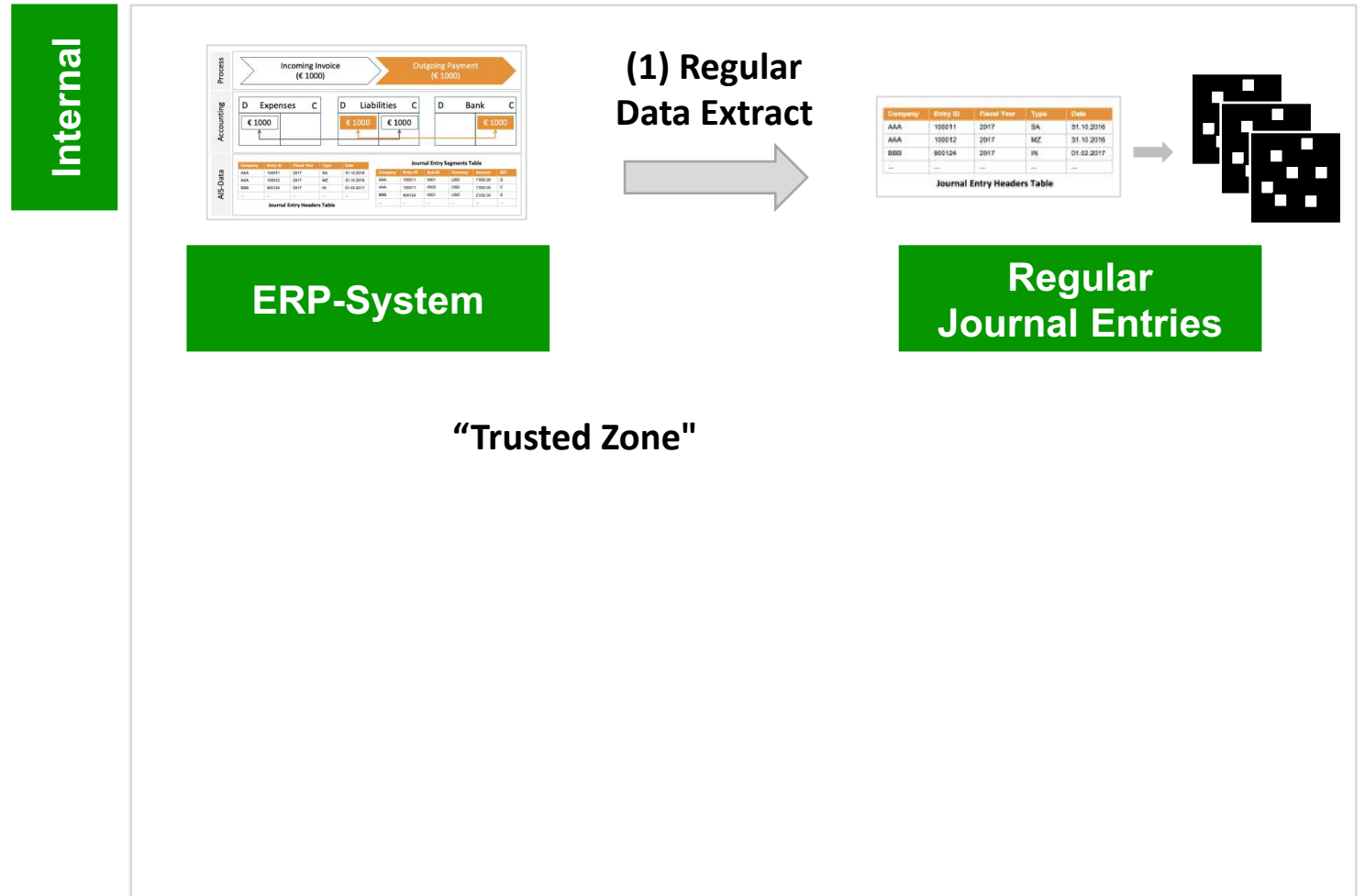


*“The primary attack vector in information leakage is insiders.
This term is used to describe a person with an interest in ‘exfiltrating’
important inside information on behalf of a third party.”*

“Information Leakage - ENISA Threat Landscape 2019 – 2020”
The European Union Agency for Cybersecurity

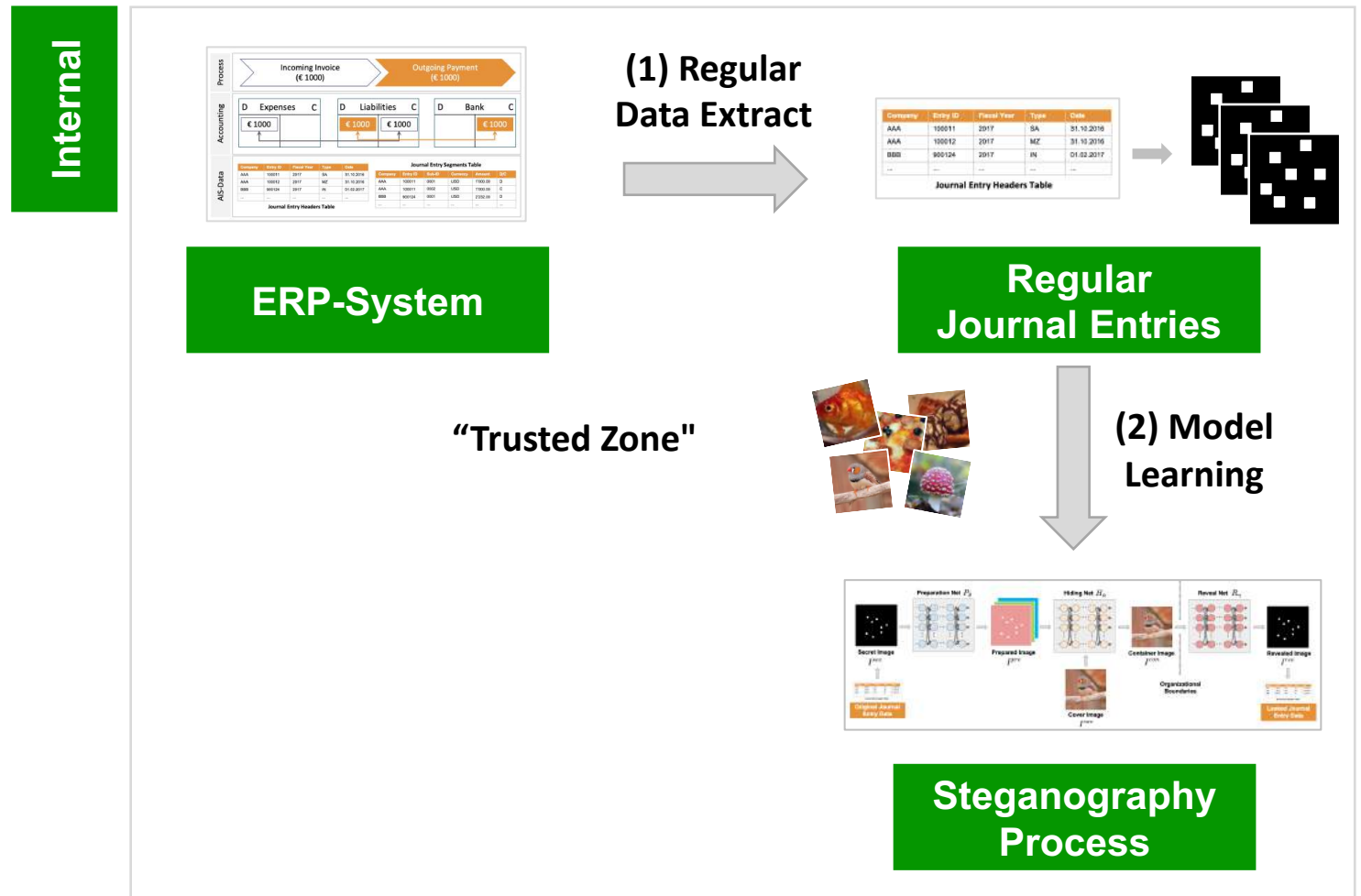
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



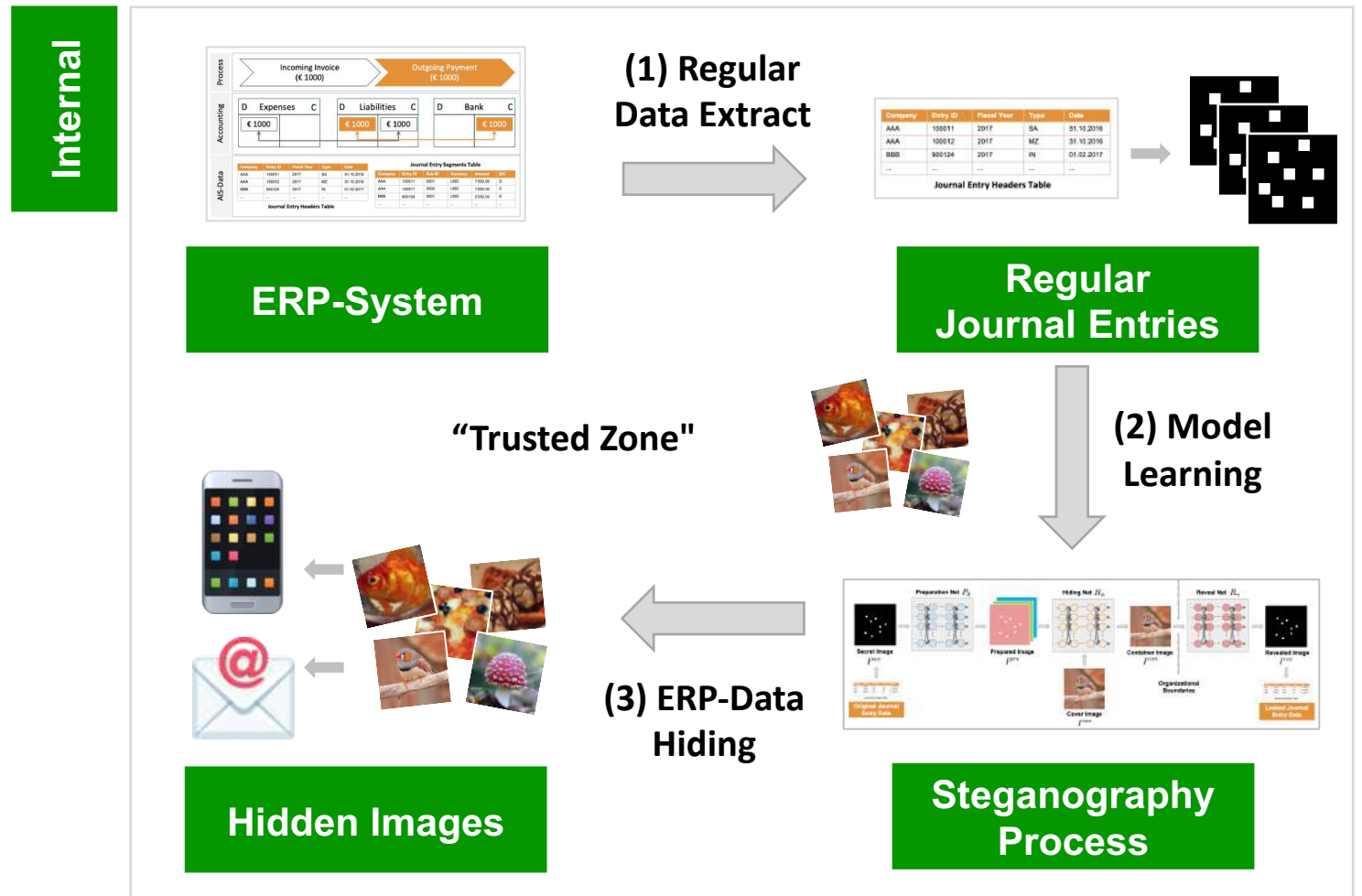
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



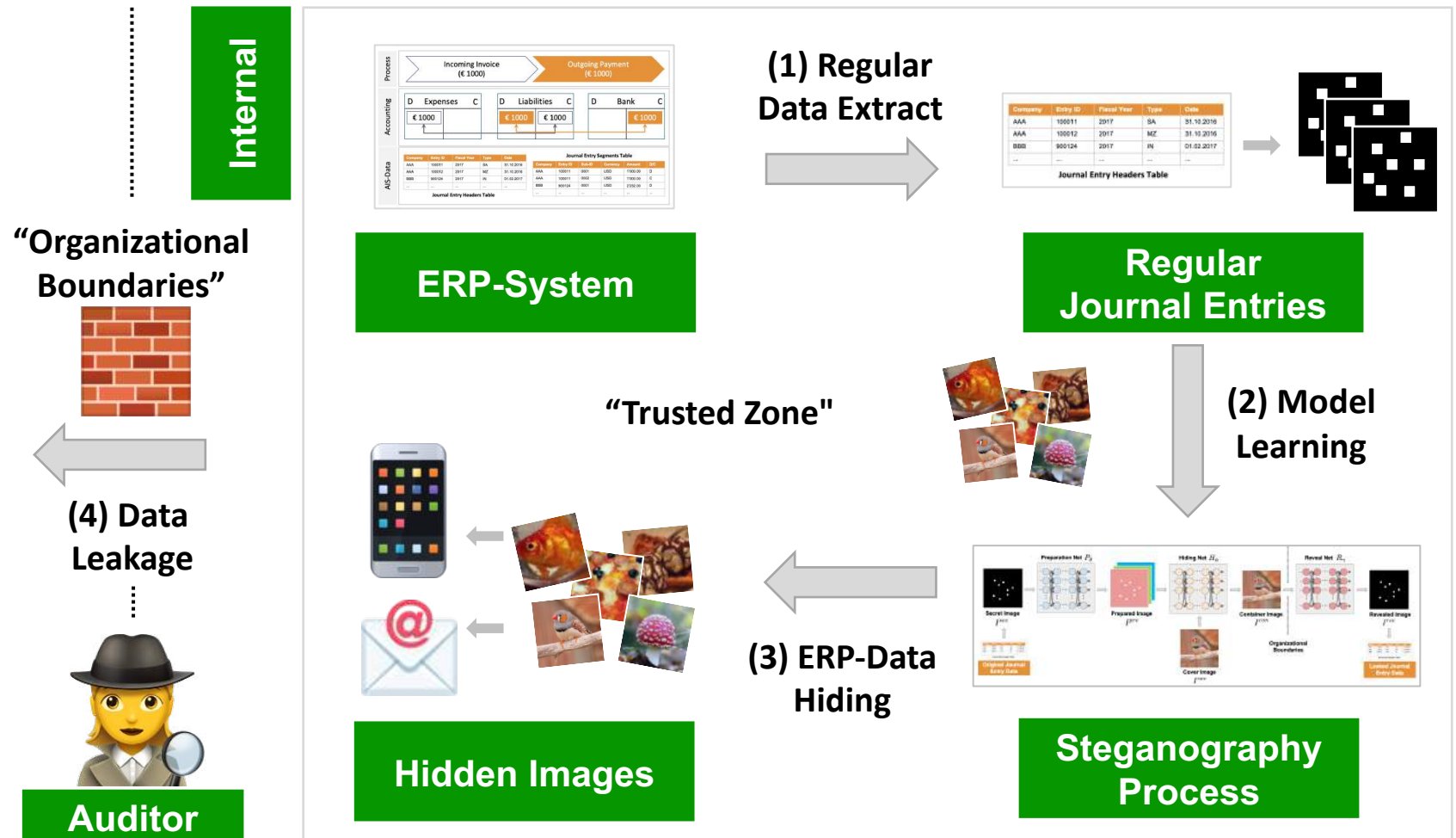
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



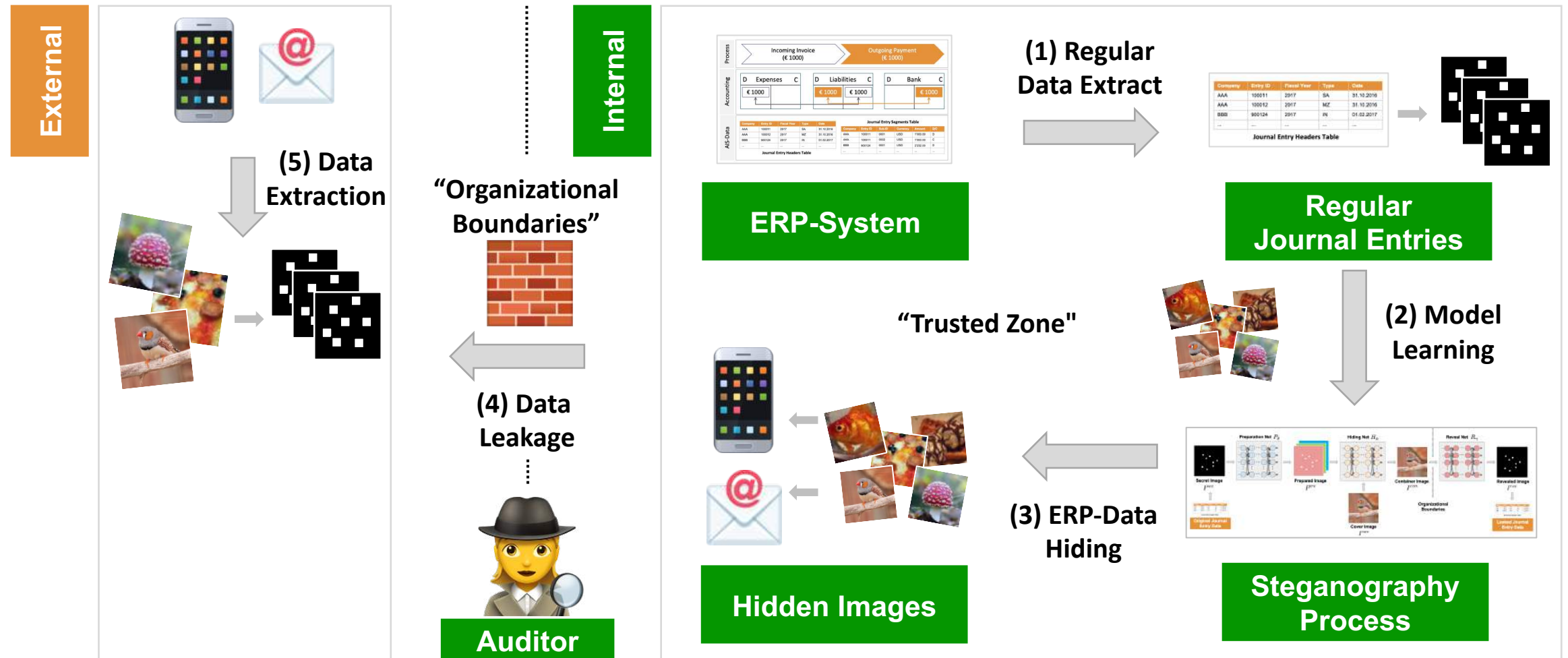
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



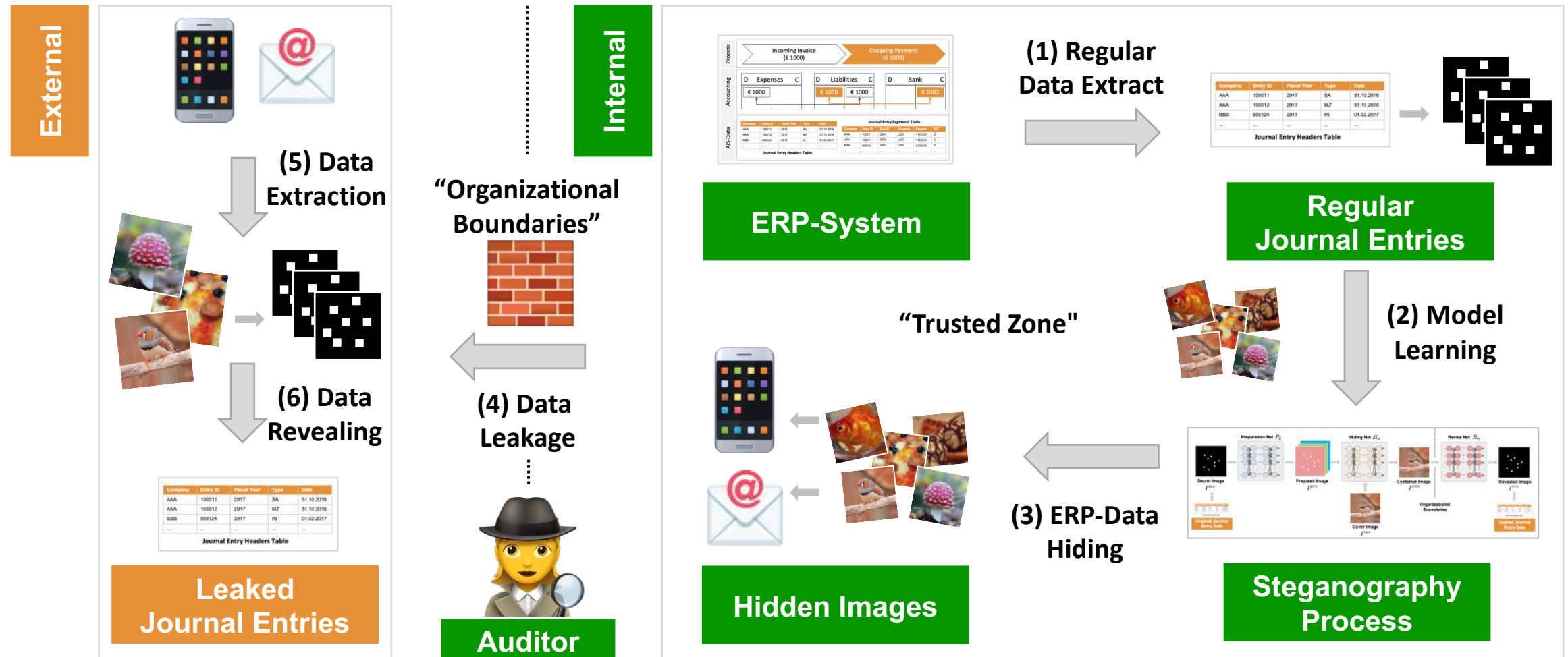
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



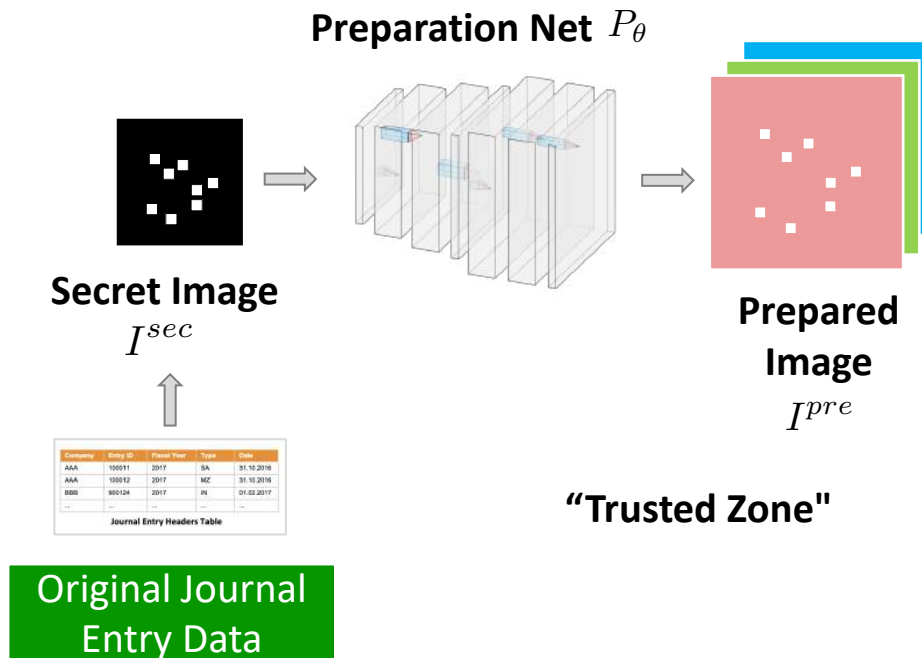
Insider Data Leakage Threat Model

Leaking Sensitive Enterprise Resource Planning Data



Data Leakage Process

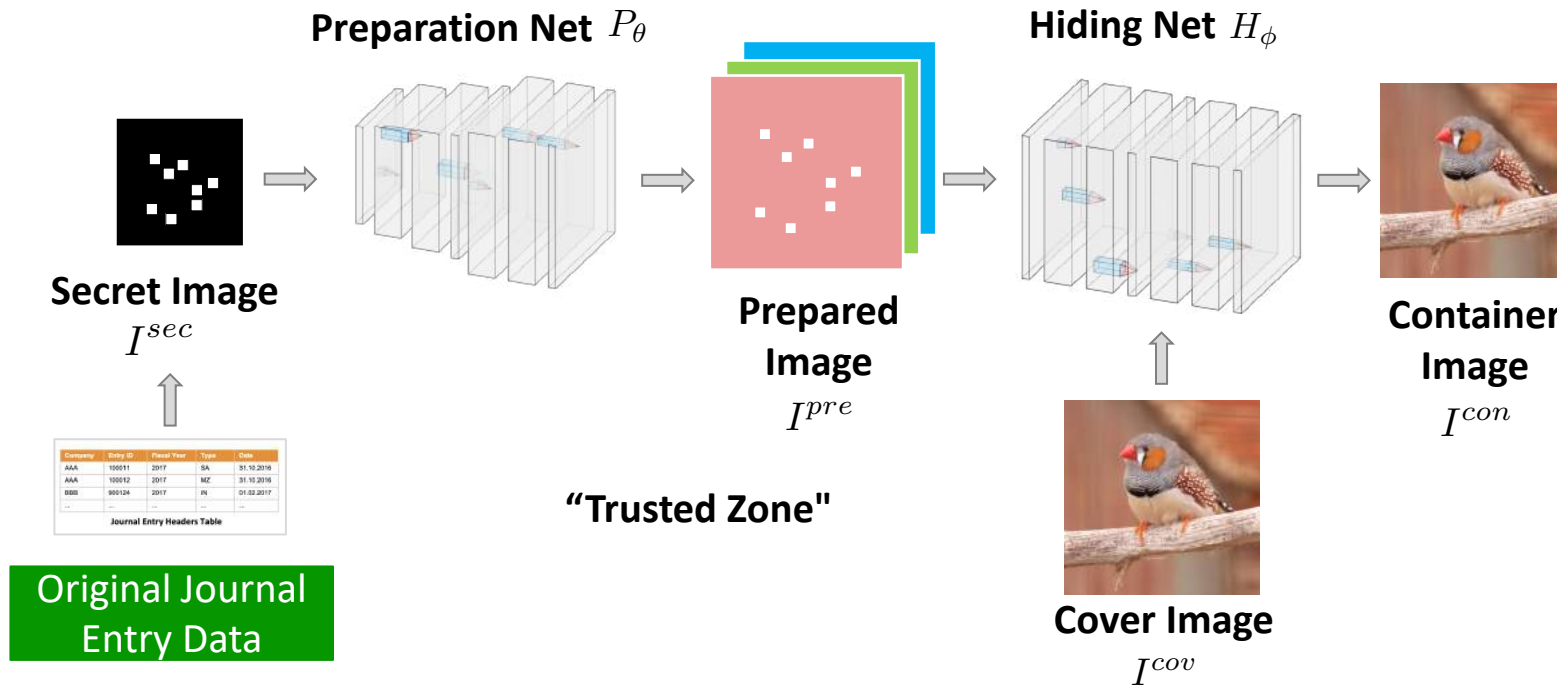
Steganographic Process to Leak Sensitive Enterprise Resource Planning Data



Following: Baluja, S. 2017. Hiding Images in Plain Sight: Deep Steganography. In Advances in Neural Information Processing Systems, 2069–2079.

Data Leakage Process

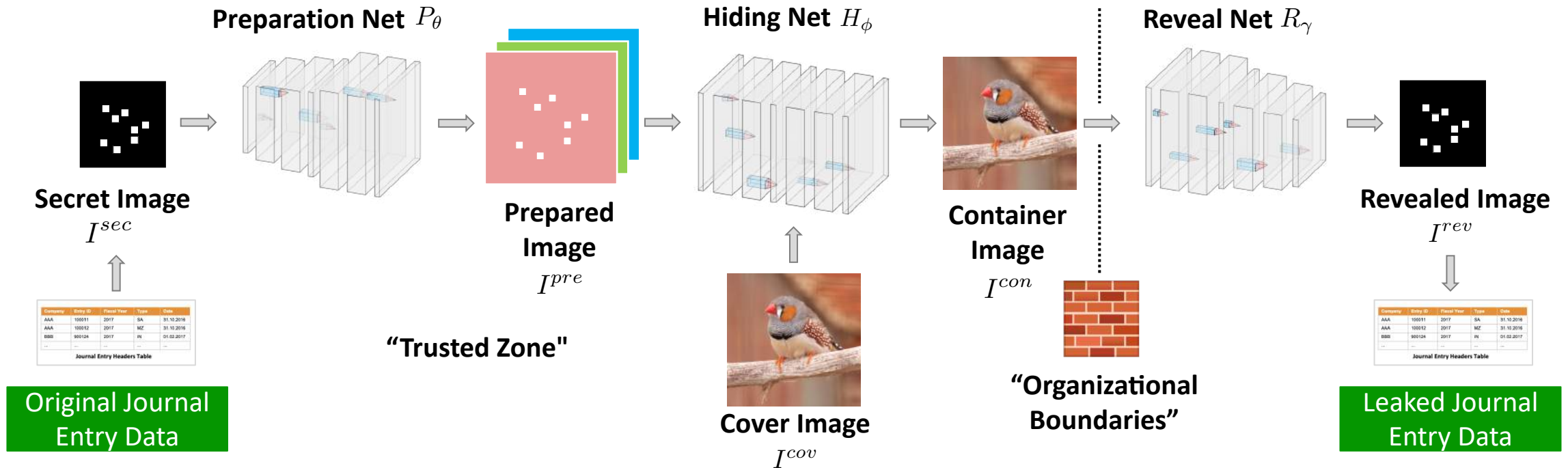
Steganographic Process to Leak Sensitive Enterprise Resource Planning Data



Following: Baluja, S. 2017. Hiding Images in Plain Sight: Deep Steganography. In Advances in Neural Information Processing Systems, 2069–2079.

Data Leakage Process

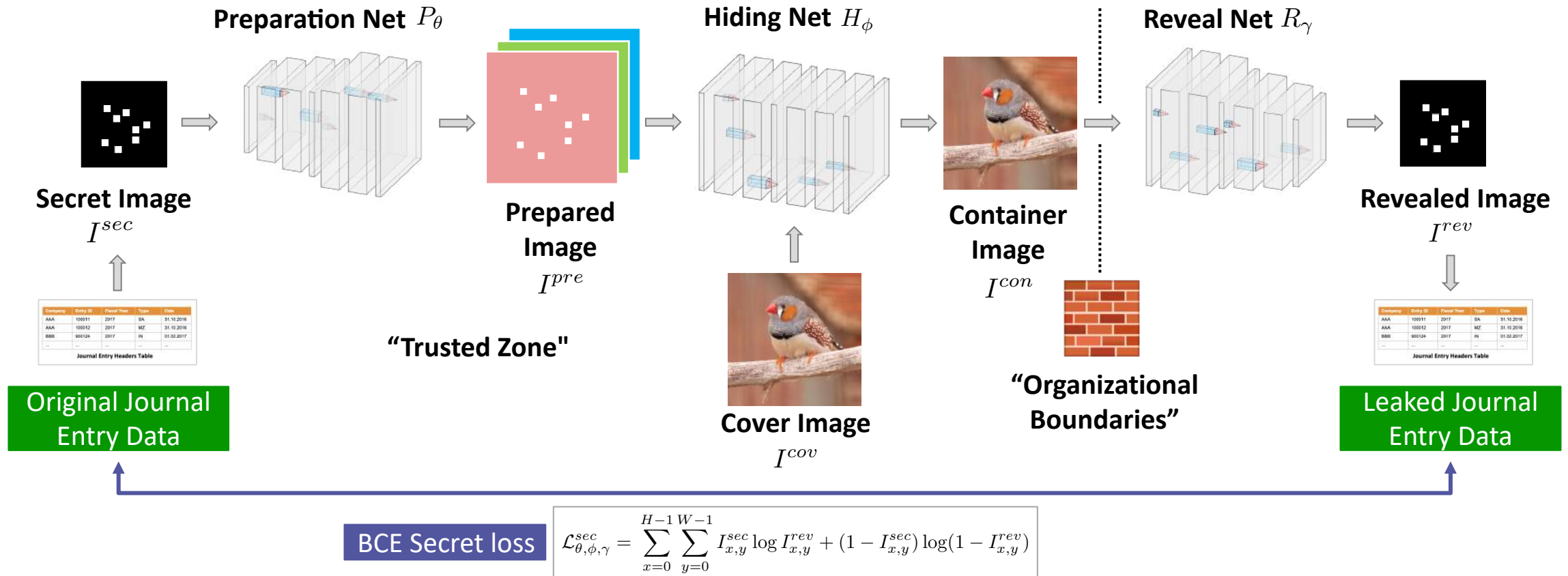
Steganographic Process to Leak Sensitive Enterprise Resource Planning Data



Following: Baluja, S. 2017. Hiding Images in Plain Sight: Deep Steganography. In Advances in Neural Information Processing Systems, 2069–2079.

Data Leakage Process

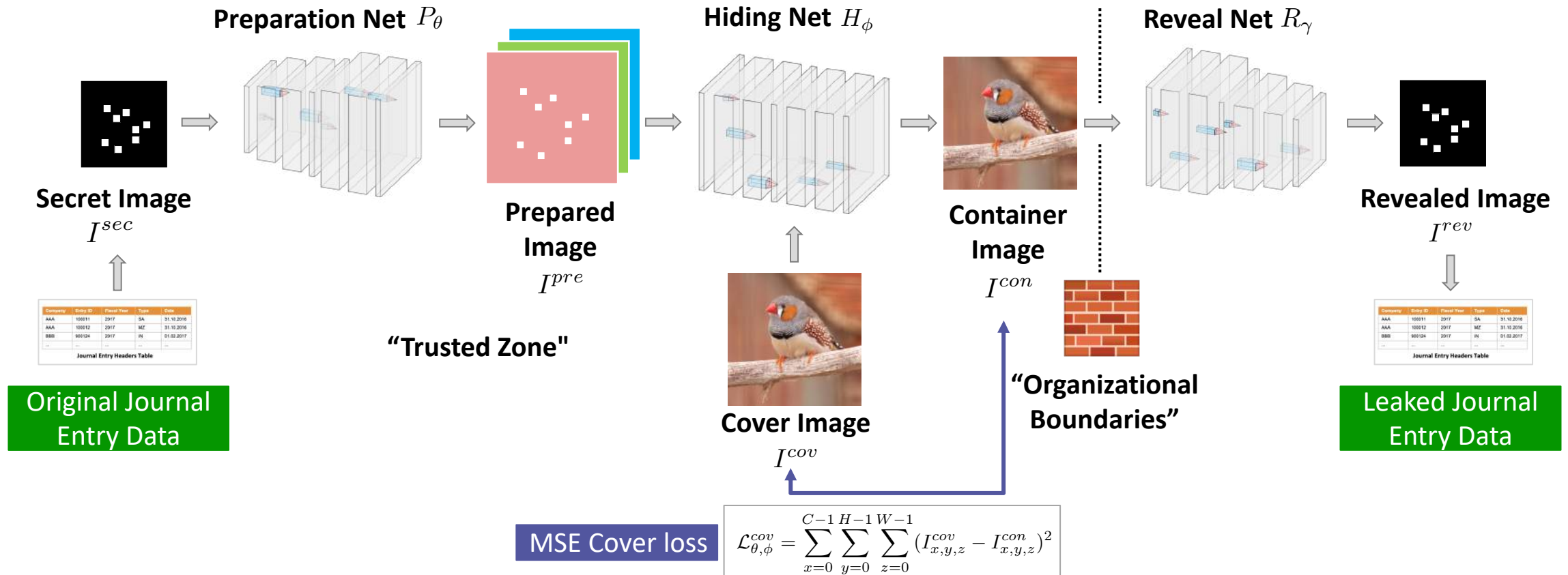
Steganographic Process to Leak Sensitive Enterprise Resource Planning Data



Following: Baluja, S. 2017. Hiding Images in Plain Sight: Deep Steganography. In Advances in Neural Information Processing Systems, 2069–2079.

Data Leakage Process

Steganographic Process to Leak Sensitive Enterprise Resource Planning Data



Following: Baluja, S. 2017. Hiding Images in Plain Sight: Deep Steganography. In Advances in Neural Information Processing Systems, 2069–2079.

Experimental Setup

Utilized “Secret” Datasets of City Payment Data

Dataset: Philadelphia “City Payments”



- \$4.2 billion city payments of the fiscal year 2017
- 60 offices, departments, boards and committees
- N=238,894 payments: 10 categorical, 1 numerical attribute
- 8,565 ‘one-hot’ encoded dimensions: $x^i \in \mathcal{R}^{8,565}$



check_date	document_no	dept	department_title	char	character_title	sub_obj	sub_obj_title
2016-11-25T00:00:00Z	CHEK17083379	1	01 CITY COUNCIL	3	03 MATERIALS AND SUPPLIES	325	PRINTING 0325
2017-01-25T00:00:00Z	CHEK17087297	1	01 CITY COUNCIL	2	02 PURCHASE OF SERVICES	240	ADVERTISING/PROMOTIONAL ACTIVITIES
2016-08-04T00:00:00Z	CHEK17012548	1	01 CITY COUNCIL	2	02 PURCHASE OF SERVICES	240	ADVERTISING/PROMOTIONAL ACTIVITIES
2016-08-26T00:00:00Z	ACHD17028918	1	01 CITY COUNCIL	2	02 PURCHASE OF SERVICES	210	POSTAGE 0210
2017-06-30T00:00:00Z	ACHD17132218	1	01 CITY COUNCIL	3	03 MATERIALS AND SUPPLIES	309	CORDAGE AND FIBERS 0309
2017-06-08T00:00:00Z	CHEK17141080	1	01 CITY COUNCIL	3	03 MATERIALS AND SUPPLIES	325	PRINTING 0325
2017-04-12T00:00:00Z	CHEK17117562	1	01 CITY COUNCIL	2	02 PURCHASE OF SERVICES	255	DUES 0255
2017-04-13T00:00:00Z	ACHD17151640	1	01 CITY COUNCIL	3	03 MATERIALS AND SUPPLIES	304	BOOKS AND OTHER PUBLICATIONS 0304

Source: <https://www.phila.gov/2019-03-29-philadelphias-initial-release-of-city-payments-data/>

Dataset: Chicago “Vendor Payments”



- \$5.3 billion vendor payments of the fiscal years 2009-2019
- 30 departments, 928 contracts
- N=72,814 payments: 7 categorical, 1 numerical attribute
- 2,354 ‘one-hot’ encoded dimensions: $x^i \in \mathcal{R}^{2,354}$



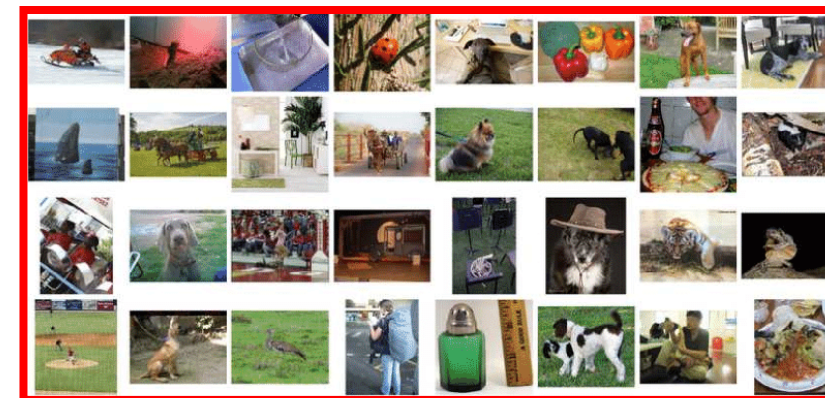
VOUCHER NUMBER	AMOUNT	CHECK DATE	DEPARTMENT NAME	CONTRACT NUMBER	VENDOR NAME	CASHED
CV5419500322	94.80	07/01/2019	DEPT OF COMMUNITY DEV...	50236	18TH STREET, DEVELOPME...	Yes
CVPT174102790	6.56	04/30/2019	DEPARTMENT OF HEALTH	64612	RESPIRATORY HEALTH ASS...	No
CVPT173004687	1,380.00	01/02/2019	DEPT OF FAMILY AND SUPP...	31322	FEATHERIST	Yes
CVPT182500015	2,686.10	01/08/2019		82037	PHALANX FAMILY SERVICES	Yes
CVPT182500016	12,831.19	01/09/2019		82037	PHALANX FAMILY SERVICES	Yes
CVPT182500017	48,588.00	01/25/2019		87409	CHICAGO CITYWIDE LITERA...	Yes
CVPT182500021	115.47	04/29/2019		89573	CATHOLIC CHARITIES OF I...	Yes
CVPT182500022	963.03	04/29/2019		89573	CATHOLIC CHARITIES OF I...	Yes

Source: <https://data.cityofchicago.org/Administration-Finance/Payments/s4vu-giwb/>

Dataset: “Tiny ImageNet” Images



- Contains 100,000 images utilized as "cover" images
- Each image shows an individual objects at low resolution
- Each image is of size 3 x 224 x 244 (C x H x W) pixels
- Grayscale images are reduced to 1 x 224 x 244 pixels



Network Architectural Details

- Each network applies three distinct series of 2D convolutions of different spatial filter sizes.
- Each convolutional layer is followed by a non-linear Rectified Linear Unit (ReLU) activation function.

AAAI'21 – Workshop on Knowledge Discovery from Unstructured Data in Financial Services

Experimental Setup

Experimental Quality Measures

Peak Signal-to-Noise Ratio (PSNR)

- Defines the ratio between the power of image fidelity and the power of corrupting noise.

➡ The “technical” container image quality.

$$\text{PSNR} = 10 \cdot \log_{10} \frac{\max |I^{cov}|}{\mathcal{L}_{\theta, \phi}^{cov}(I^{cov}, I^{con})}$$

Structural Similarity Index Measure (SSIM)

- Defines a perception-based model considering image degradation as perceived change in structural information.

➡ The “perceived” container image quality.

$$\text{SSIM} = \frac{(2\mu_{cov}\mu_{con} + c_1)(2\sigma_{cov, con} + c_2)}{(\mu_{cov}^2 + \mu_{con}^2 + c_1)(\sigma_{cov}^2 + \sigma_{con}^2 + c_2)}$$

Bit Accuracy (BACC)

- Defines the fraction of identical active bits between the secret image and the revealed image.

➡ The revealed data accuracy.

$$\text{BACC} = 1 - \frac{\sum_{x=0}^{H-1} \sum_{y=0}^{W-1} \mathbb{1}[\|I_{x,y}^{sec} - M \circ I_{x,y}^{rev}\| \leq \delta]}{\sum_{x=0}^{H-1} \sum_{y=0}^{W-1} \mathbb{1}[I_{x,y}^{sec} > 0]}$$

Experimental Results

Quantitative Experimental Results

RGB Cover								Cover Quality		Accuracy
	Dataset	BPP	α	β	$\mathcal{L}_{\theta,\phi,\gamma}^{ALL}$	$\mathcal{L}_{\theta,\phi}^{cov}$	$\mathcal{L}_{\theta,\phi,\gamma}^{sec}$	PSNR	SSIM	BACC
			0.2	1.0						
Philadelphia:	A	0.0436	0.5	1.0	0.74 ± 0.03	0.40 ± 0.04	0.54 ± 0.04	44.23 ± 0.50	0.997 ± 0.001	0.998 ± 0.003
			0.8	1.0	0.81 ± 0.03	0.34 ± 0.03	0.55 ± 0.01	44.97 ± 0.43	0.998 ± 0.001	0.999 ± 0.001
			1.0	1.0	0.92 ± 0.06	0.38 ± 0.05	0.54 ± 0.02	44.42 ± 0.65	0.997 ± 0.001	0.998 ± 0.002
			0.2	1.0	0.63 ± 0.01	0.43 ± 0.06	0.54 ± 0.01	43.91 ± 0.64	0.997 ± 0.001	0.999 ± 0.001
Chicago:	B	0.0120	0.5	1.0	0.77 ± 0.01	0.52 ± 0.02	0.51 ± 0.01	42.94 ± 0.23	0.996 ± 0.001	0.998 ± 0.001
			0.8	1.0	0.87 ± 0.05	0.46 ± 0.07	0.51 ± 0.00	43.51 ± 0.65	0.995 ± 0.003	0.998 ± 0.002
			1.0	1.0	0.99 ± 0.04	0.47 ± 0.04	0.52 ± 0.01	43.35 ± 0.37	0.997 ± 0.001	0.997 ± 0.003
			0.2	1.0	0.67 ± 0.07	0.79 ± 0.03	0.51 ± 0.00	41.38 ± 0.32	0.998 ± 0.002	0.999 ± 0.001

Variances originate from parameter initialization using four distinct random seeds.

Grayscale Cover	Dataset	BPP	α	β	$\mathcal{L}_{\theta,\phi,\gamma}^{ALL}$	$\mathcal{L}_{\theta,\phi}^{cov}$	$\mathcal{L}_{\theta,\phi,\gamma}^{sec}$	PSNR	SSIM	BACC
			0.2	1.0						
			0.5	1.0						
Philadelphia:	A	0.1307	0.8	1.0	0.44 ± 0.00	0.01 ± 0.00	0.54 ± 0.00	60.98 ± 0.13	0.999 ± 0.001	0.999 ± 0.006
			1.0	1.0	0.55 ± 0.01	0.01 ± 0.00	0.54 ± 0.01	60.15 ± 1.79	0.999 ± 0.001	0.999 ± 0.004
			0.2	1.0	0.11 ± 0.01	0.01 ± 0.00	0.54 ± 0.01	64.84 ± 0.49	0.999 ± 0.001	0.997 ± 0.003
			0.5	1.0	0.29 ± 0.00	0.01 ± 0.00	0.54 ± 0.01	60.93 ± 1.58	0.999 ± 0.001	0.997 ± 0.002
Chicago:	B	0.0359	0.8	1.0	0.98 ± 0.05	0.52 ± 0.06	0.58 ± 0.01	43.19 ± 0.53	0.990 ± 0.001	0.967 ± 0.006
			1.0	1.0	0.99 ± 0.13	0.41 ± 0.14	0.58 ± 0.01	44.28 ± 1.34	0.991 ± 0.001	0.966 ± 0.005
			0.2	1.0	0.75 ± 0.02	0.82 ± 0.08	0.53 ± 0.01	39.98 ± 0.33	0.985 ± 0.001	0.983 ± 0.003
			0.5	1.0	0.87 ± 0.04	0.64 ± 0.07	0.55 ± 0.02	42.27 ± 0.49	0.989 ± 0.001	0.977 ± 0.002

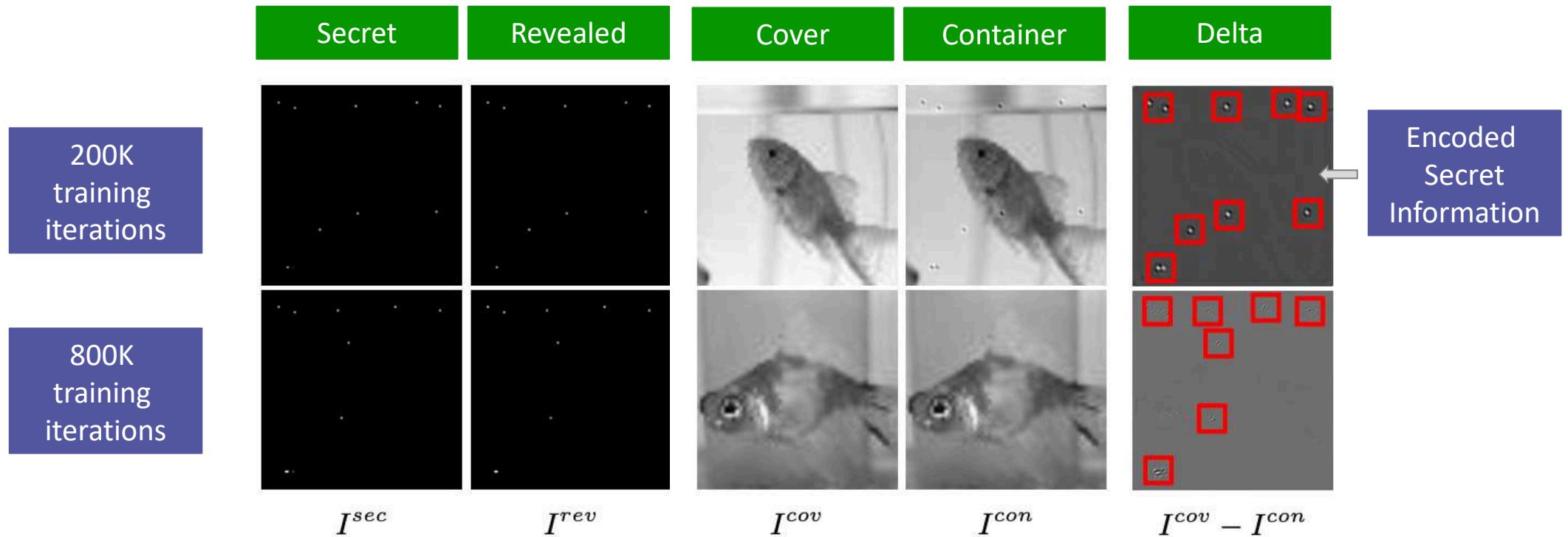
Variances originate from parameter initialization using four distinct random seeds.



The quantitative results demonstrate the steganographic capabilities of the trained models.

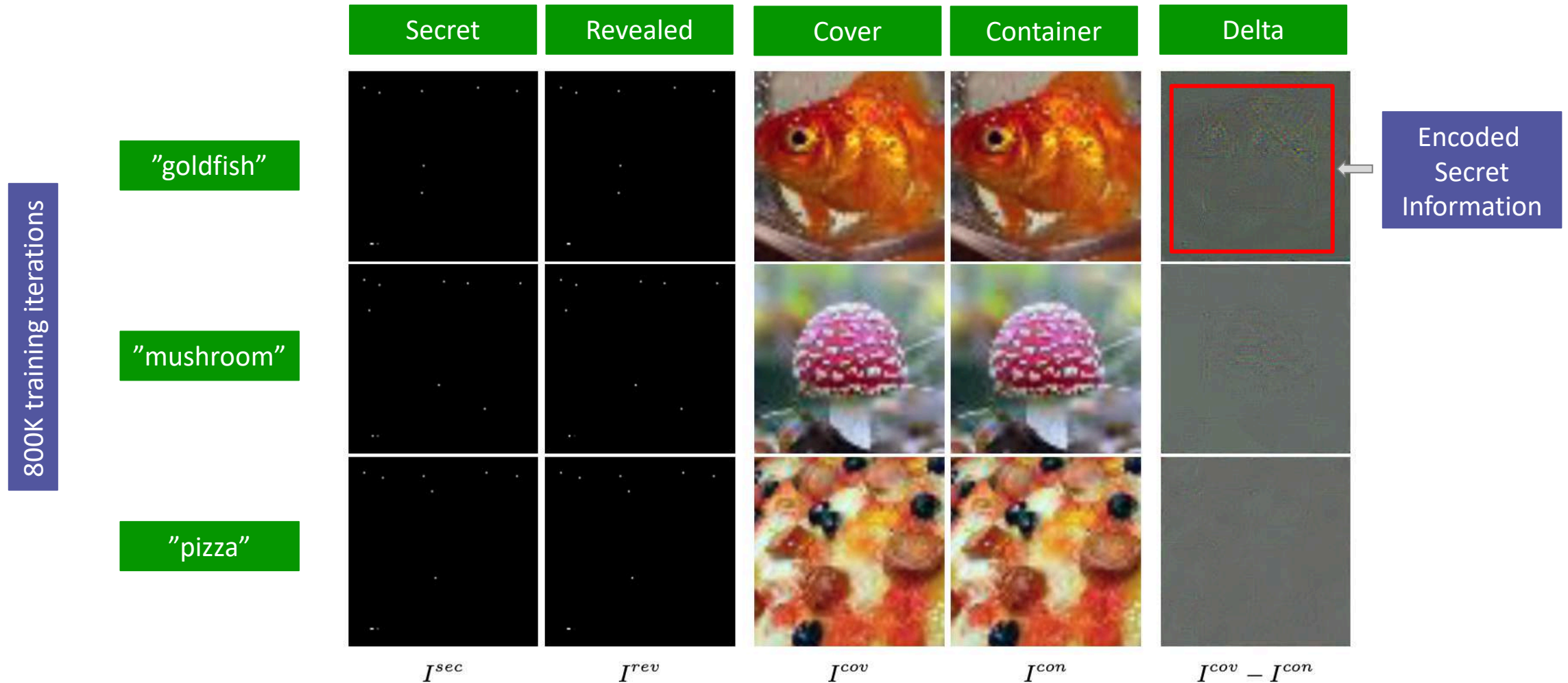
Experimental Results

Qualitative Experimental Results (single Journal Entry, single Grayscale Cover Images)



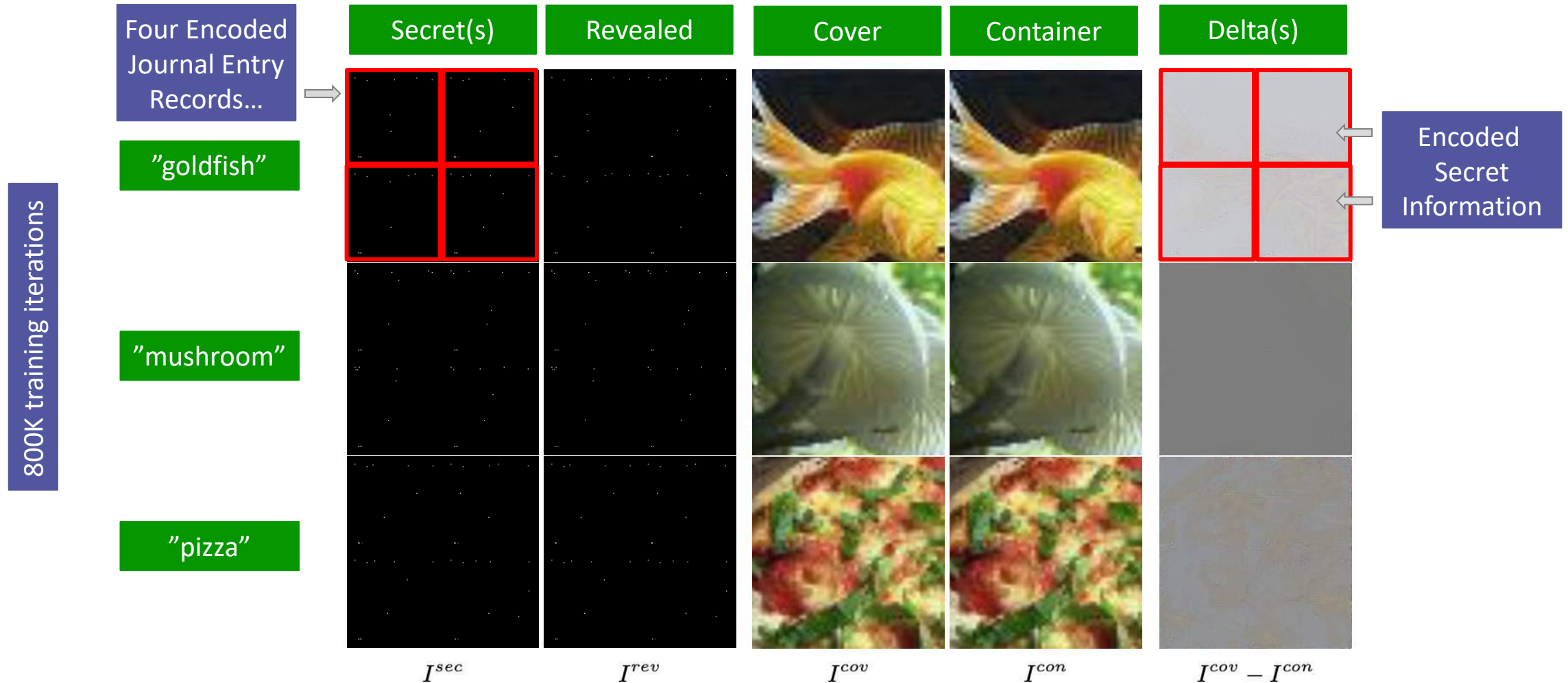
Experimental Results

Qualitative Experimental Results (single Journal Entry, single RGB Cover Images)



Experimental Outlook

Qualitative Experimental Results (four Journal Entries, single RGB Cover Images)



Thank you 🧐



Contact: **Marco Schreyer**
Artificial Intelligence & Machine Learning
University of St.Gallen (HSG)
marco.schreyer@unisg.ch

Paper:



Paper available via:
<https://aaai-kdf.github.io/kdf2021/>