Adversarial Learning of “Deepfakes” in Accounting

M. Schreyer[1], T. Sattarov[2], B. Reimer[3], and D. Borth[1]

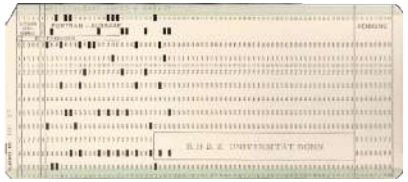
[1]University of St. Gallen, [2]Deutsche Bundesbank, [3]PricewaterhouseCoopers AG

Opinions expressed in this work are solely those of the authors, and do not necessarily reflect the view of Deutsche Bundesbank or PricewaterhouseCoopers (PwC) International Ltd. and its network firms.

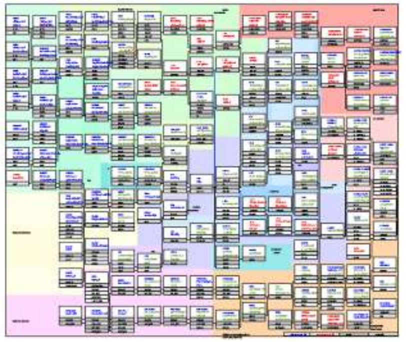
Evolution of Financial Accounting

1920‘s

1960‘s



1990‘s



Data Volume

- Continuous digitization of business activities and processes.
- Accumulation of exhaustive transactional and business process data.
- „Every“ activity within an organization leaves a **digital trace.**

Evolution of Financial Accounting

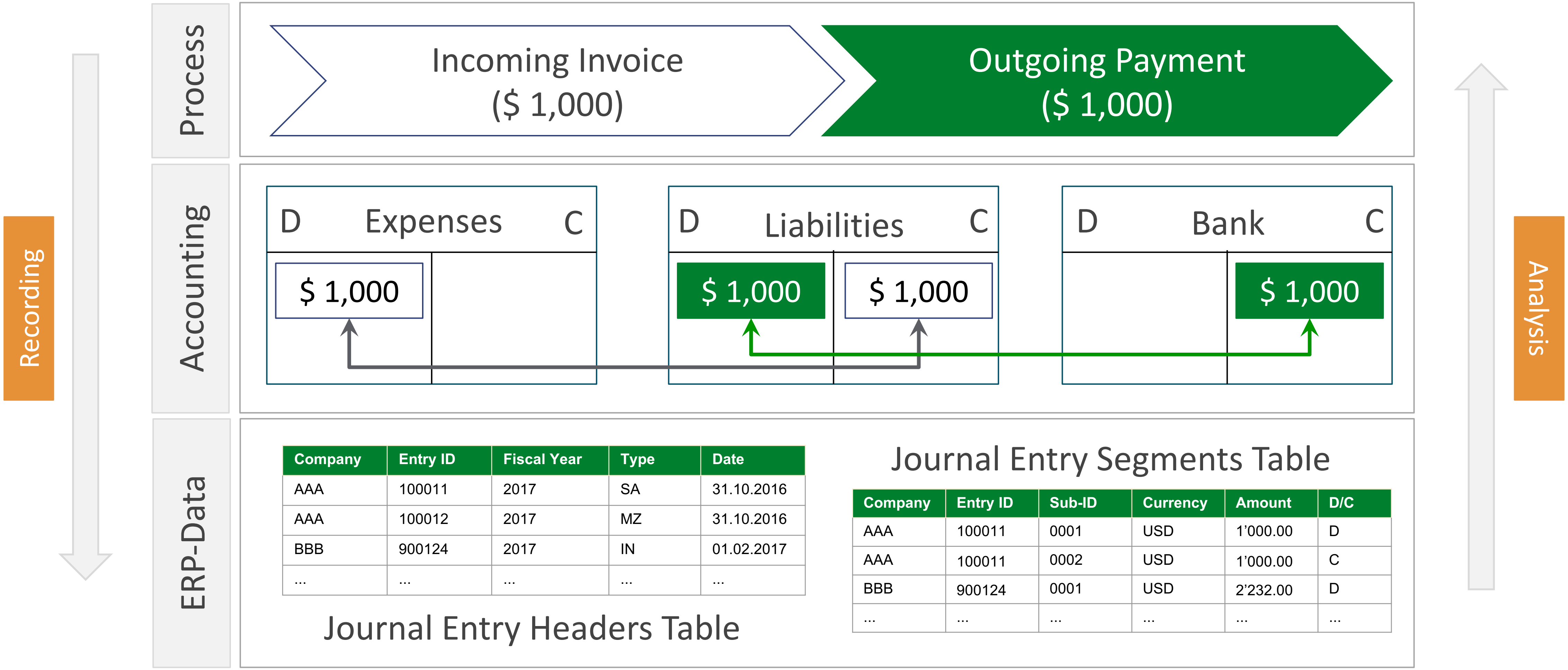
1920‘s **1960‘s** **1990‘s**

> ‘Approx. 77% of the worlds revenue touches one of our ERP systems.’ SAP AG’s Corporate Factsheet 2019

Data Volume

- Continuous digitization of business activities and processes.
- Accumulation of exhaustive transactional and business process data.
- „Every“ activity within an organization leaves a **digital trace.**

Enterprise Ressource Planning (ERP) Systems



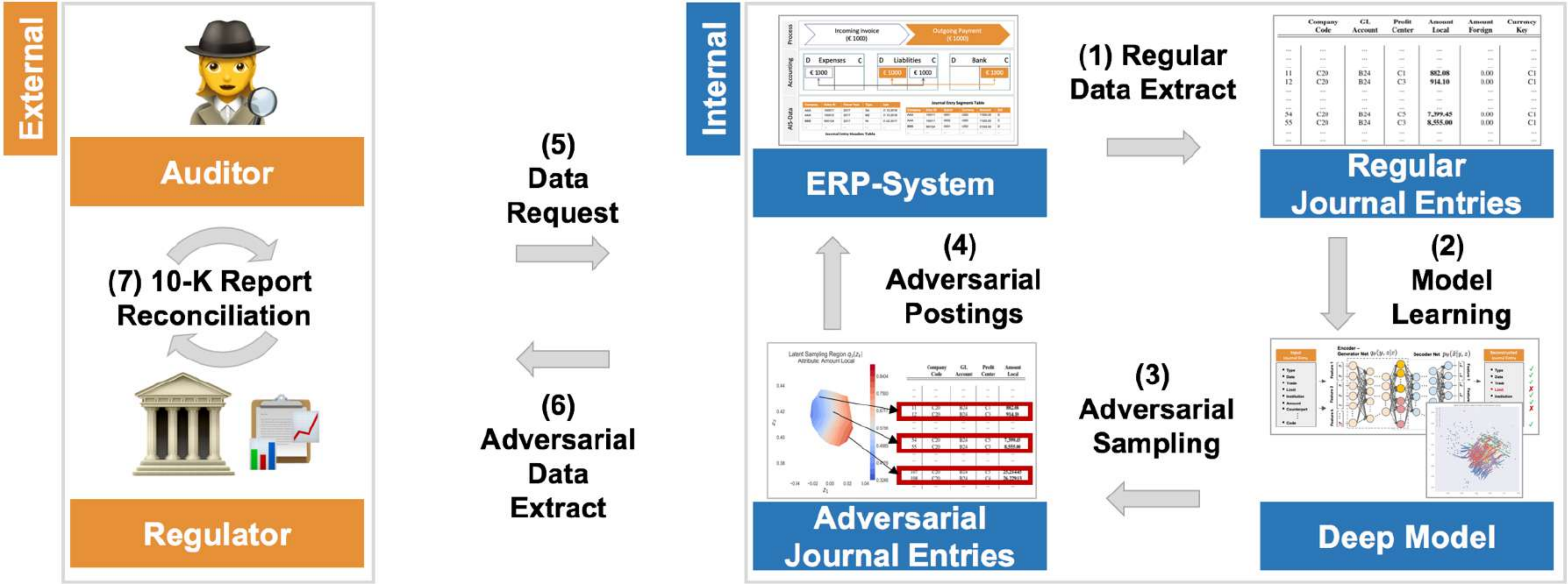
Company	Entry ID	Fiscal Year	Type	Date
AAA	100011	2017	SA	31.10.2016
AAA	100012	2017	MZ	31.10.2016
BBB	900124	2017	IN	01.02.2017
…	…	…	…	…

Journal Entry Headers Table

Journal Entry Segments Table

Company	Entry ID	Sub-ID	Currency	Amount	D/C
AAA	100011	0001	USD	1'000.00	D
AAA	100011	0002	USD	1'000.00	C
BBB	900124	0001	USD	2'232.00	D
…	…	…	…	…	…

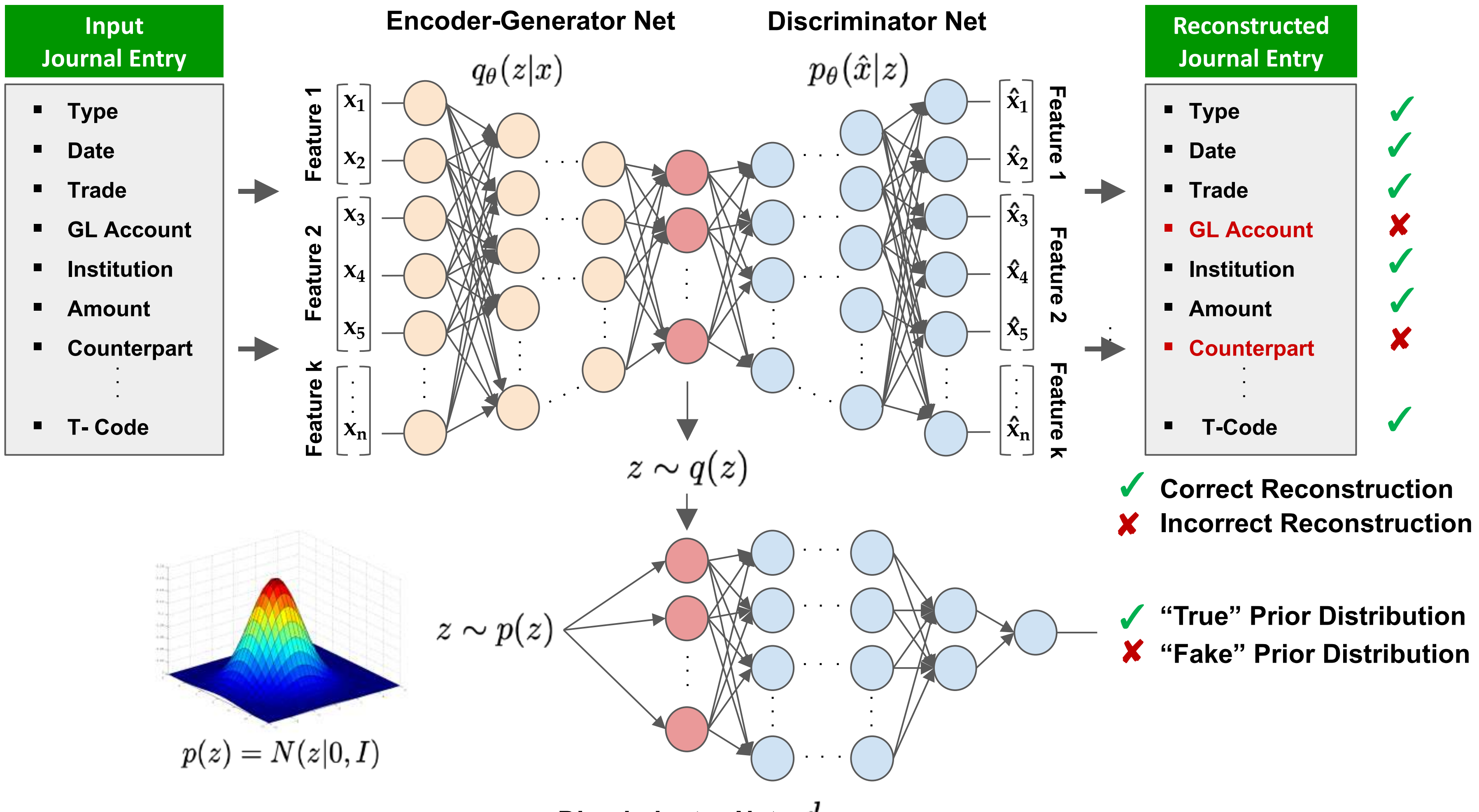
Audit Threat Model



Journal Entry Audit

Adversarial Model Learning & Journal Entry Sampling

Adversarial Autoencoder NNs[1,2]



Input Journal Entry

- Type
- Date
- Trade
- GL Account
- Institution
- Amount
- Counterpart
- T- Code

Encoder-Generator Net $q_\theta(z|x)$

Discriminator Net $p_\theta(\hat{x}|z)$

$z \sim q(z)$

$z \sim p(z)$

$p(z) = N(z|0, I)$

Discriminator Net $d_{\phi(z)}$

Reconstructed Journal Entry

- Type ✓
- Date ✓
- Trade ✓
- GL Account ✗
- Institution ✓
- Amount ✓
- Counterpart ✗
- T-Code ✓

✓ **Correct Reconstruction**

✗ **Incorrect Reconstruction**

✓ **“True” Prior Distribution**

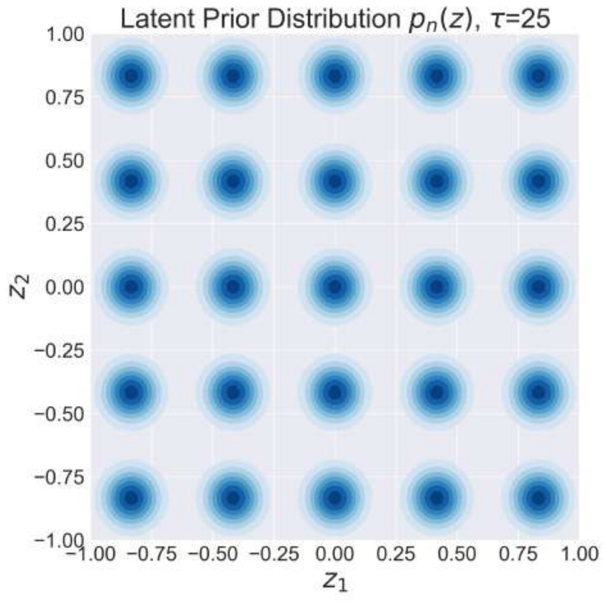
✗ **“Fake” Prior Distribution**

[1] Makhzani, Alireza, et al. "Adversarial autoencoders." *arXiv preprint arXiv:1511.05644,* 2015

[2] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems,* 2014

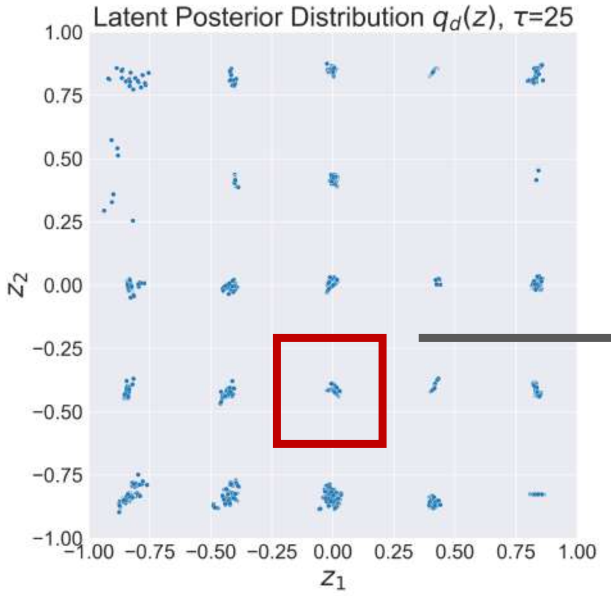
Learning Disentangled Representations

Imposed Latent Prior Distribution



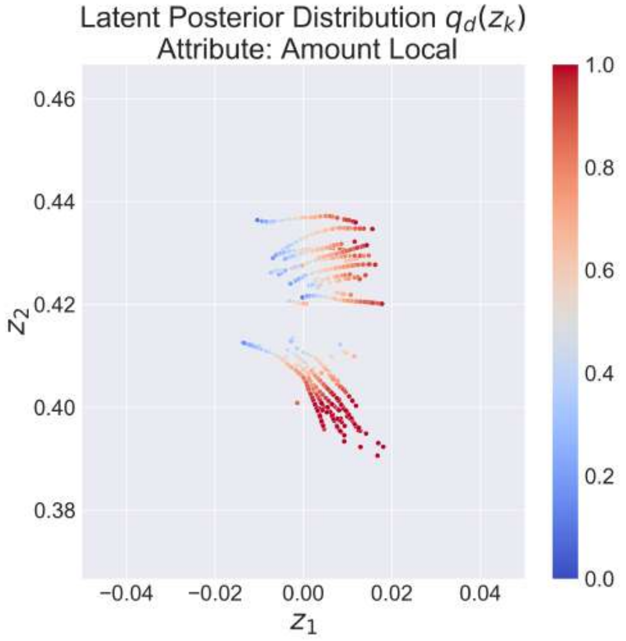
Target Distribution

High-Order Generative Factor Disentanglement



Disentanglement of Accounting Processes

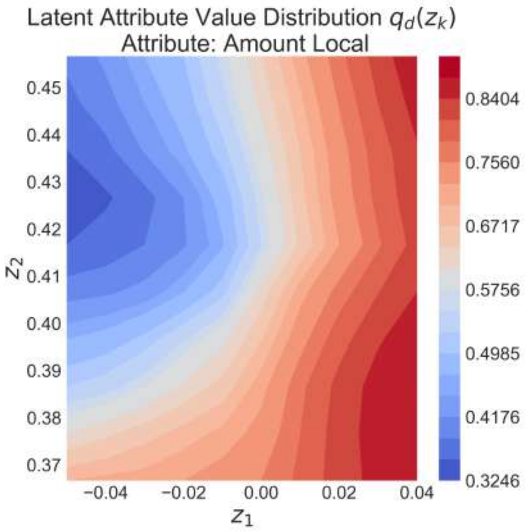
Low-Order Generative Factor Disentanglement



Disentanglement of Journal Entries

Learning Disentangled Representations

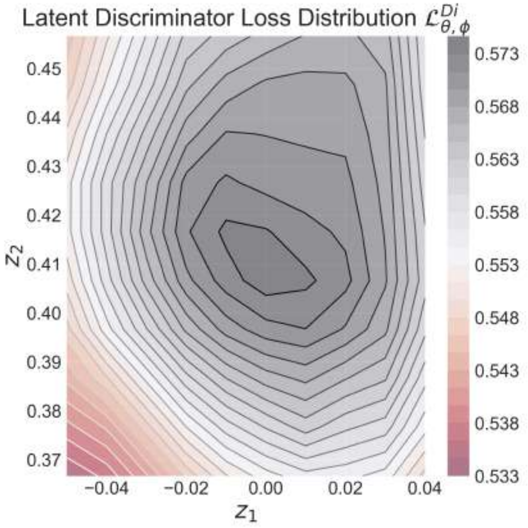
Journal Entry Amount Value Distribution



$$\max p_\theta(\hat{x}_j|z) \neq \max p_\theta(\hat{x}_j|z+\delta)$$

-

Discriminator Loss Distribution

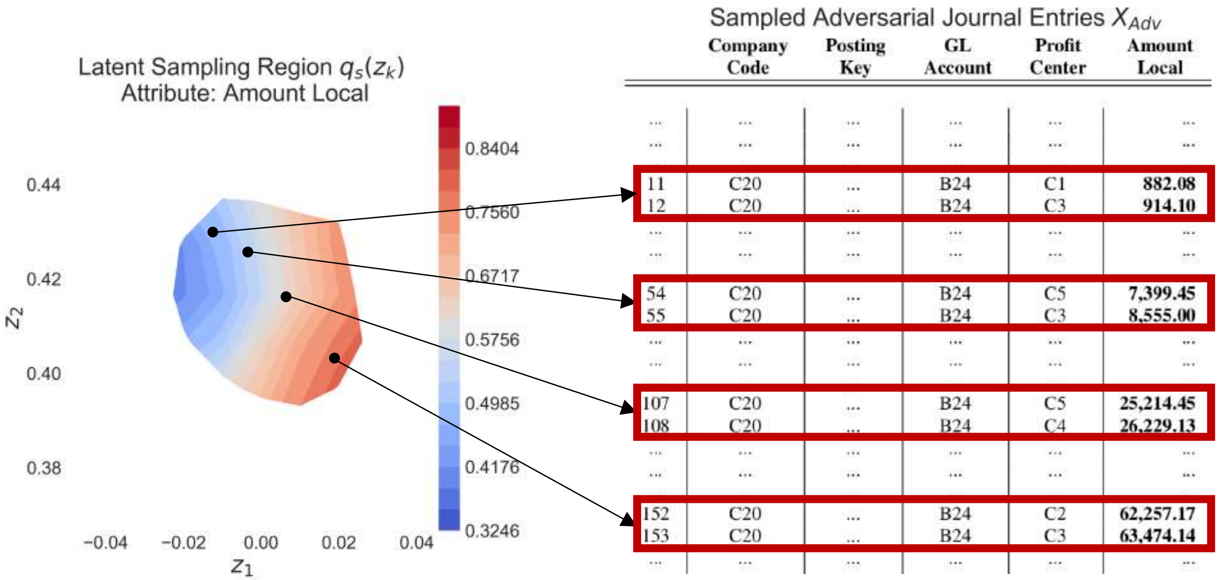


$$\|d_\phi(z+\delta)\| \geq d_\phi(z) + \rho$$

=

"Adversarial Sampling Map"

Adversarial Journal Entries



Sampled Adversarial Journal Entries X_{Adv}

	Company Code	Posting Key	GL Account	Profit Center	Amount Local
...	...	...	...	...	...
...	...	...	...	...	...
11	C20	...	B24	C1	**882.08**
12	C20	...	B24	C3	**914.10**
...	...	...	...	...	...
...	...	...	...	...	...
54	C20	...	B24	C5	**7,399.45**
55	C20	...	B24	C3	**8,555.00**
...	...	...	...	...	...
...	...	...	...	...	...
107	C20	...	B24	C5	**25,214.45**
108	C20	...	B24	C4	**26,229.13**
...	...	...	...	...	...
...	...	...	...	...	...
152	C20	...	B24	C2	**62,257.17**
153	C20	...	B24	C3	**63,474.14**
...	...	...	...	...	...

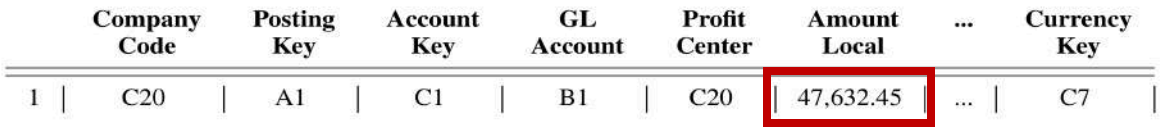
$$z_{adv} \sim q_s(z)$$

$$\hat{x}_{adv}$$

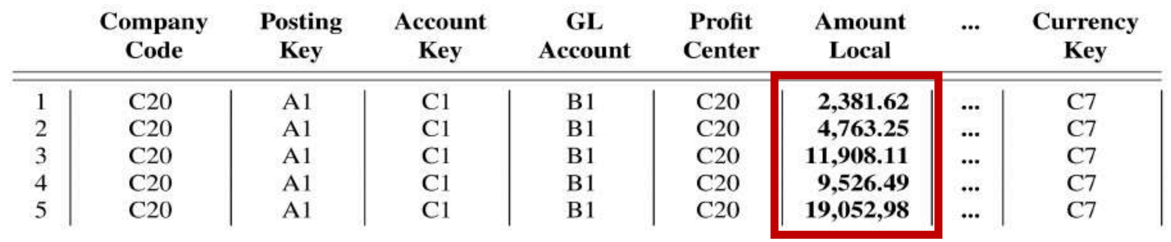
Sampled “Deepfake“ Accounting Records

“Anomaly Replacement” attack scenario:

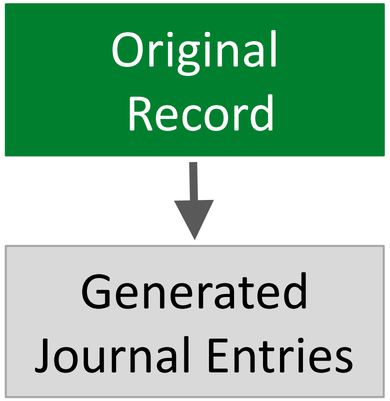
“camouflage the circumvention of an invoice approval limit”



	Company Code	Posting Key	Account Key	GL Account	Profit Center	Amount Local	...	Currency Key
1	C20	A1	C1	B1	C20	47,632.45	...	C7



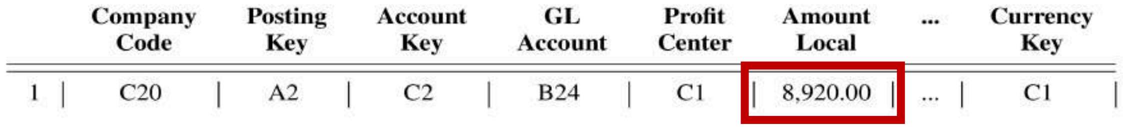
	Company Code	Posting Key	Account Key	GL Account	Profit Center	Amount Local	...	Currency Key
1	C20	A1	C1	B1	C20	**2,381.62**	...	C7
2	C20	A1	C1	B1	C20	**4,763.25**	...	C7
3	C20	A1	C1	B1	C20	**11,908.11**	...	C7
4	C20	A1	C1	B1	C20	**9,526.49**	...	C7
5	C20	A1	C1	B1	C20	**19,052,98**	...	C7


Original Record

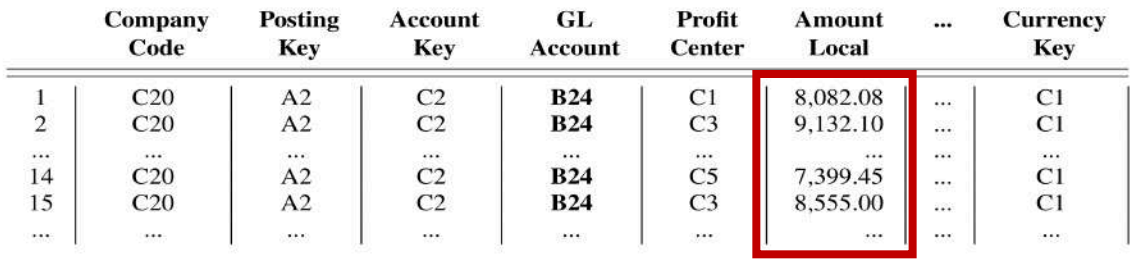
Generated Journal Entries

“Anomaly Augmentation” attack scenario:

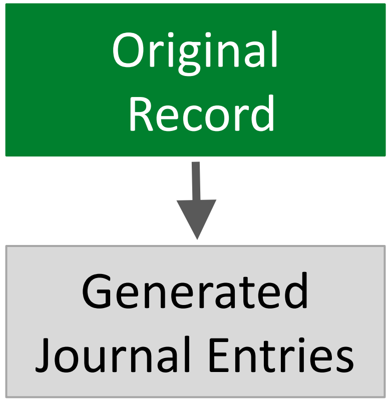
“camouflage the usage of seldom used general ledger accounts”



	Company Code	Posting Key	Account Key	GL Account	Profit Center	Amount Local	...	Currency Key
1	C20	A2	C2	B24	C1	8,920.00	...	C1



	Company Code	Posting Key	Account Key	GL Account	Profit Center	Amount Local	...	Currency Key
1	C20	A2	C2	**B24**	C1	8,082.08	...	C1
2	C20	A2	C2	**B24**	C3	9,132.10	...	C1
...	...	...	...	...	...	...	...	...
14	C20	A2	C2	**B24**	C5	7,399.45	...	C1
15	C20	A2	C2	**B24**	C3	8,555.00	...	C1
...	...	...	...	...	...	...	...	...


Original Record

Generated Journal Entries

Thank you

Marco Schreyer marco.schreyer@unisg.ch
Timur Sattarov timur.sattarov@bundesbank.de

Publication available on arXiv:
https://arxiv.org/abs/1910.03810

SCAN ME