Strict validation versus accepting anything

Evan Jones





Personalized e-commerce marketing

- ~4 years
- ~140 employees
- ~35 engineers

Hand Picked For You

We thought you might like...

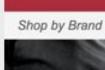
SHOP DRESSES



Amy Knit Dress



Parkdale Combo Dress



Video Gallery

Rebate Center

STILL NEED THIS FOR YOUR 1991 CHEVROLET K2500

Exact Fit



Accel DstCapandRtr



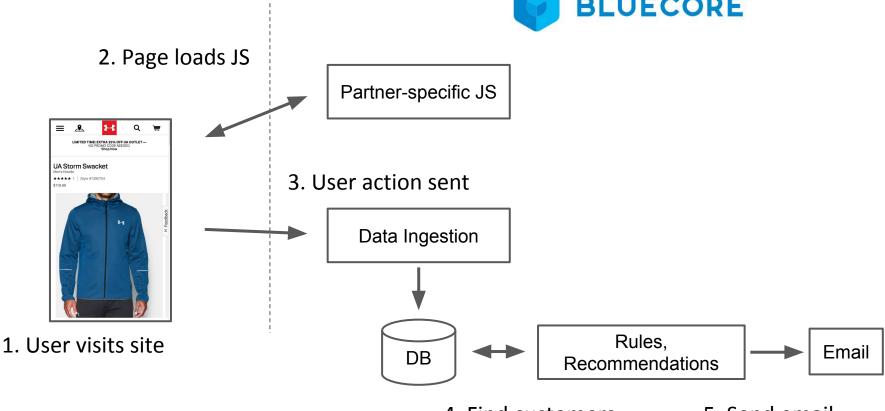


ACDelco Distributor Cap

Exact Fit

Recommendations need product data





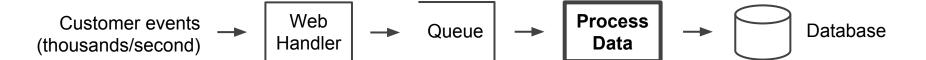
4. Find customers

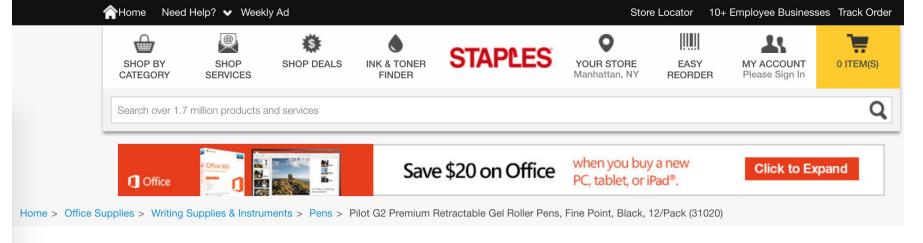
5. Send email

Data ingestion



Data ingestion

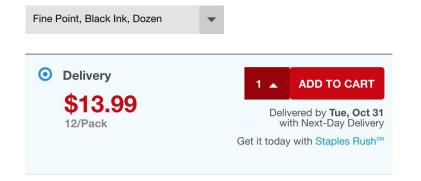




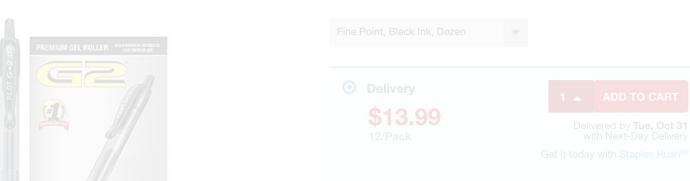
Pilot G2 Premium Retractable Gel Roller Pens, Fine Point, Black, 12/Pack (31020)

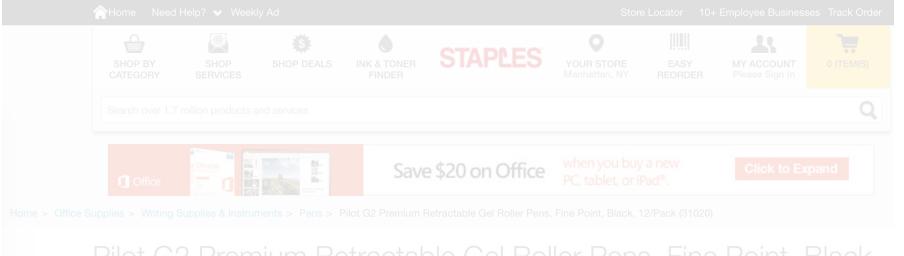
Item: 429174 Model: 31020 ★★★★★ (591) | Write a Review





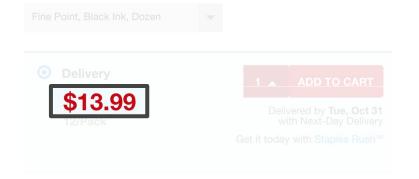


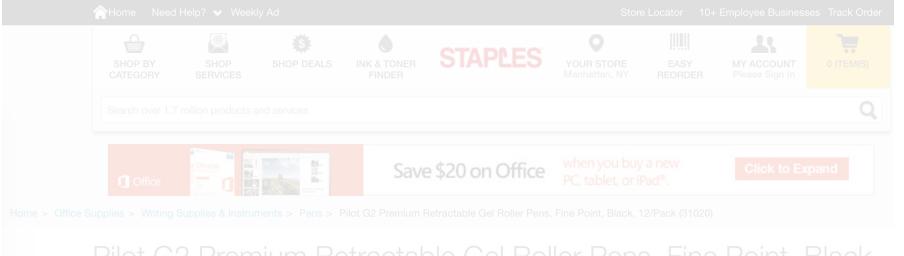




Pilot G2 Premium Retractable Gel Roller Pens, Fine Point, Black, 12/Pack (31020)

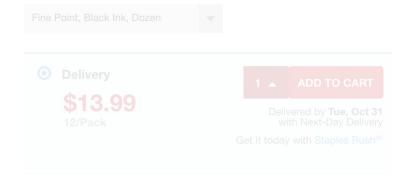






Pilot G2 Premium Retractable Gel Roller Pens, Fine Point, Black, 12/Pack (31020)





Product data

```
"id": 429174,
"name": "Pilot G2 Premium Retractable...",
"price": 13.99,
```

Product data

```
"id": 429174,
"name": "Pilot G2 Premium Retractable...",
"price": 13.99,
```

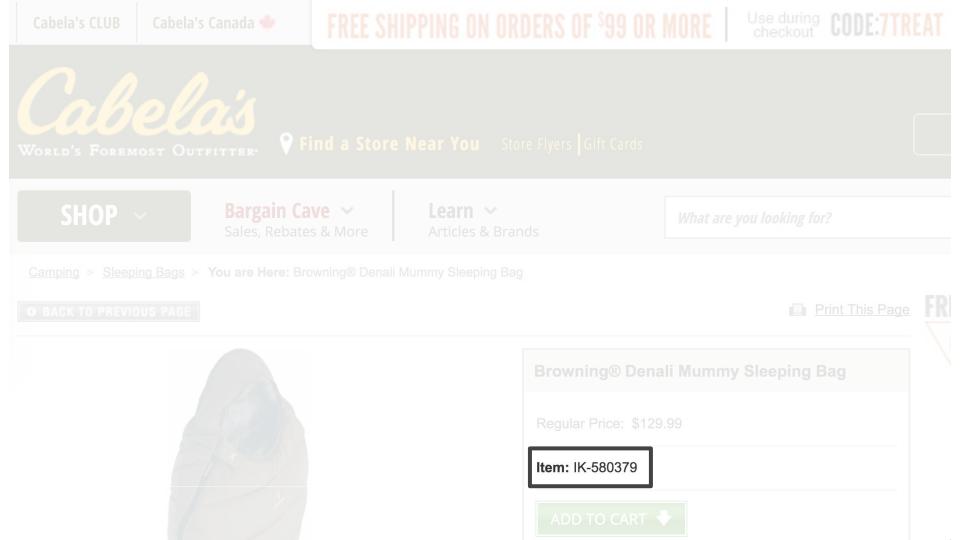




Regular Price: \$129.99

Item: IK-580379

ADD TO CART ▼



Product data

```
"id": "429174", // may contain letters
"name": "Pilot G2 Premium Retractable...",
"price": 13.99,
```

Product data

```
"id": "429174", // may contain letters
"name": "Pilot G2 Premium Retractable...",
"price": 13.99,
```

Eng V Stores

Q My Account



BCBGMAXAZRIA

Spring '18 Preview What's New Dresses Clothing Shoes Accessories Sale Factory

Home > Clothing > Sweaters And Cardigans > Masha Oversized Cropped Pullover





Masha Oversized Cropped Pullover

Style #IQS1Z433-0D2

£151.00



Color: LIGHT DOVE

UK Size:

GET 15% OFF ▶

BCBGMAXAZRIA

Spring '18 Preview What's New Dresses Clothing Shoes Accessories Sale Factory

Home > Clothing > Sweaters And Cardigans > Masha Oversized Cropped Pullover

Next >



Masha Oversized Cropped Pullover

Style #IQS1Z433-0D2

£151.00



Write a Review

Color:



LIGHT DOVE

UK Size:

GET 15% OFF

Product data

```
"id": "429174", // may contain letters
"name": "Pilot G2 Premium Retractable...",
"price": "£13.99, // may contain currency
```

.

What is **valid** product data?

Design A: Accept anything!

Design B: Strict validation!

Design choice: Validation versus flexibility

Programming languages: static versus dynamic typing

Databases: Strict versus flexible schemas (SQL vs NoSQL)

Robustness Principle (Postel's Law)

"Be liberal in what you accept, and conservative in what you send"

Advantage: implementations can interoperate (e.g. TCP)

Disadvantage: bugs can become "standard" (e.g HTML)

Original policy: Accept anything

Rationale: One chance to store the data; fix it later

Implementation: Store any key/value pairs

Fun ensues ...

price: 13 (integer), 13.99 (float), "13.99", "£13.99"

products without ids

products with both "title" and "name"

Evaluation

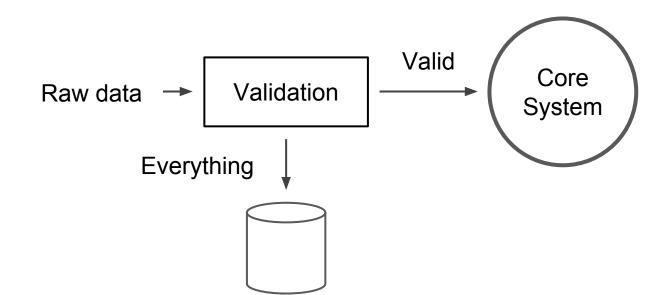
- + Store any e-commerce data
- + Fix any data bugs

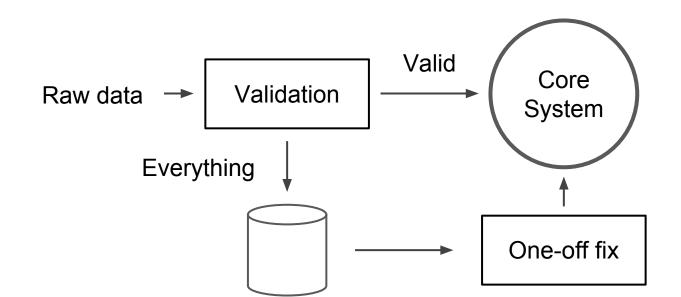
Evaluation

- Store any e-commerce data
- + Fix any data bugs

- Processing is much harder
- Harder to test if we are sending the right data







Conclusion: Err on the side of validation

Find errors sooner

Simplifies the overall system

Easier to relax restrictions than to add them

Want to fix errors later? Record everything

Thanks!

Evan Jones

http://www.evanjones.ca/

Bluecore http://www.bluecore.com/



Store raw but require "core" schema

Store the raw data we receive

Validate "core" fields: return helpful error messages

e.g. must have id, price is a string, use "name" not "title"

Found *many* data bugs