



Fixing a Schema with Domains and Annotations

Add data intent to describe messy schemas

Alex Nuijten

@alexnuijten

<https://nuijten.blogspot.com/>

Chris Saxon, Oracle Developer Advocate

@ChrisRSaxon.bsky.social & @SQLDaily

<https://blogs.oracle.com/sql>

<https://www.youtube.com/c/TheMagicofSQL>

<https://www.linkedin.com/in/saxonchris/>



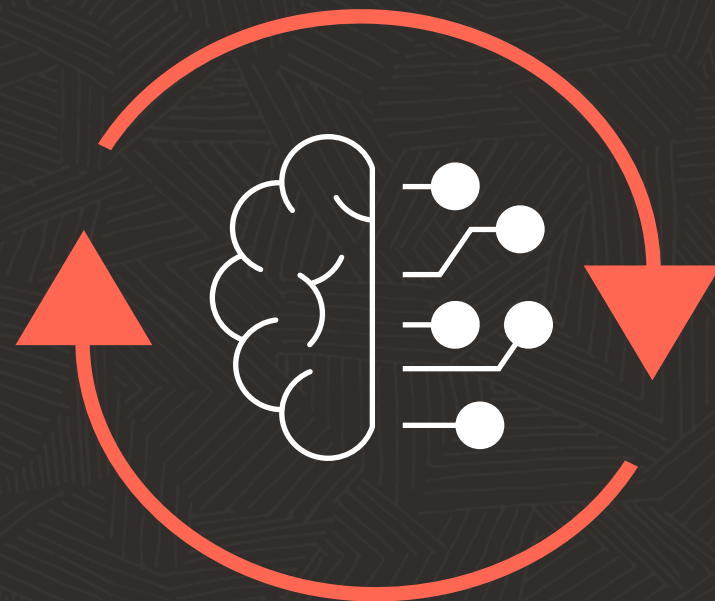
A man in a dark suit, white shirt, and striped tie is speaking in a meeting. He is looking towards the left. Two speech bubbles are overlaid on the image. The background is blurred, showing other people in the meeting.

Mike talked
about GenDev

We need 23ai
to unleash
productivity!



Say what you
want an app to do



AI generates
the app

A man in a dark suit, white shirt, and striped tie is sitting at a table in a meeting. He has a confused or questioning expression on his face. Two speech bubbles are overlaid on the image. The background is blurred, showing other people in the meeting.

We can't
understand
the schema

How will
chatgpt?

A photograph of three people in an office setting. A woman with glasses and a striped shirt stands on the left, looking at a computer monitor. A man with glasses and a maroon shirt stands in the middle, also looking at the monitor. A man with short brown hair and a white shirt sits at the desk on the right, pointing at the screen. A large speech bubble on the left contains the text "I thought we had column comments?". A smaller speech bubble on the right contains the text "Yes, but...".

I thought we
had column
comments?

Yes, but...



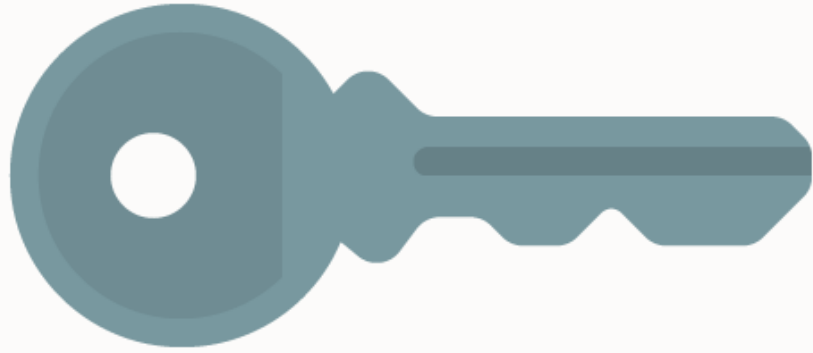
```
comment on column products.id is 'obvious';
comment on column products.name is
'this means it is the name, I guess; allowed operations:
insert; update; delete; read; display;';
comment on column products.i_date is
'{"meaning":"I don't know",
  "dm1":"delete; read; display;"}';
comment on column products.description is
'<meaning>something about products</meaning>
<operations>insert; update; delete; read; display;
</operations>';
comment on column products.sku is '...';
comment on column products.contacteaddr is '~\_(\ツ)\_/~';
```

A photograph of three people in an office setting. A woman with glasses and a striped shirt stands on the left, looking at a computer monitor. A man with glasses and a maroon shirt stands in the middle, also looking at the monitor. A man in a white shirt sits at the desk on the right, pointing at the monitor. The background is a bright, out-of-focus office space.

What about
annotations?

Schema Annotations

Metadata for applications



[value]

Annotations are a series of
keys with *optional* values

comment on column products.i_date is
'{"meaning":"I don't know",
 "dml":"delete; read; display;"}';

alter table products modify (i_date
 annotations (
 system_generated, --default add
 add display 'Insert timestamp',
 replace allowed_select '["filter", "sort"]',
 add or replace disallowed_dml '["insert", "update"]'
)
)





Let's
annotate!

DEMO

Data Intention Language (DIL) - Schema Annotations

Metadata

Key-value pairs

Freeform

No DB validation

A photograph of two men in a meeting. The man on the left, wearing glasses and a beard, is pointing at a laptop screen. The man on the right is leaning in, resting his chin on his hand, looking at the screen. A third person's hand is visible on the right side of the frame, also pointing at the screen. Two speech bubbles are overlaid on the image: one on the left containing the text 'Our annotations are different ...' and one on the right containing the text '*sigh*'.

Our
annotations
are different
...

sigh

A woman with dark hair is sitting at a desk in a server room, looking at a computer monitor. Two men are standing behind her, looking at the same screen. The monitor displays a dashboard with charts and data. A speech bubble is overlaid on the image, containing the text "Can we share annotations?".

Can we share
annotations?

Data Use Case Domains

Extend base types with use case information



```
create [ usecase ] domain <domain_name> as <datatype>
    [ strict ]
    [ <default> ]
    [ <collation> ]
    [ not null ]
    [ <check constraints> ]
    [ <display expr> ]
    [ <order expr> ]
    [ <annotations> ]
```

Optional properties



Is this SQL
standard?

...it's
important to
some folks

SQL SYNTAX

Oracle

VS

ANSI





Is this SQL
standard?

...it's
important to
these folks

It extends
the standard



create *domain* *<domain_name>* *as* *<datatype>*

[<default>]

[<collation>]

[not null]

[<check constraints>]

SQL standard

create [*usecase*] *domain* *<domain_name>* *as* *<datatype>*
[*strict*]
[<default>]
[<collation>]
[not null]
[<check constraints>]
[*<display expr>*]
[*<order expr>*]
[*<annotations>*]

SQL standard
Oracle extension



```
create usecase domain insert_timestamp as timestamp
```

```
  annotations (  
    system_generated,  
    display 'Insert timestamp',  
    allowed_reads '["filter", "sort"]' ,  
    disallowed_writes '["insert", "update"]'  
  );
```

A group of people are gathered around a table, looking at a laptop screen. A man with glasses and a beard is pointing at the screen. Another man is leaning in, looking at the screen with his hand on his chin. A woman is also visible, looking at the screen. The background is blurred, showing an office environment.

How is it
system
generated?

...we can
define it



```
create domain insert_timestamp as timestamp
default systimestamp
not null
annotations (
    system_generated,
    display 'Insert timestamp',
    allowed_reads '["filter", "sort"]' ,
    disallowed_writes '["insert", "update"]'
);
```

comment on column products.contacteaddr is '-_(ツ)_/-' ;

Contact email address



A photograph of two women in an office environment. The woman on the left, wearing a yellow cardigan, is looking towards the woman on the right. The woman on the right, wearing a blue patterned sari, is gesturing with her hands while speaking. Three dark grey speech bubbles are overlaid on the image, containing text related to SQL validation and regex.

Can we
validate it?

Can you
write the
regex?

...here
you go



```
REGEXP_LIKE (  
    email,  
    '^([a-zA-Z0-9!#$%&*+=?^_`{|}~ -]+  
    (\.[A-Za-z0-9!#$%&*+=?^_`{|}~ -]+  
    )*)@((a-zA-Z0-9]([a-zA-Z0-9-]*  
    [a-zA-Z0-9])?\.)+[a-zA-Z0-9](  
    [a-zA-Z0-9-]*[a-zA-Z0-9])?)$'  
)
```

A photograph of two women sitting at a table in an office or meeting room. The woman on the left is wearing a yellow cardigan over a white collared shirt and has her hand on her chin, looking thoughtful. The woman on the right is wearing a blue patterned sari and is gesturing with her hands while speaking. Three speech bubbles are overlaid on the image, containing text related to a technical discussion.

I think you missed a [

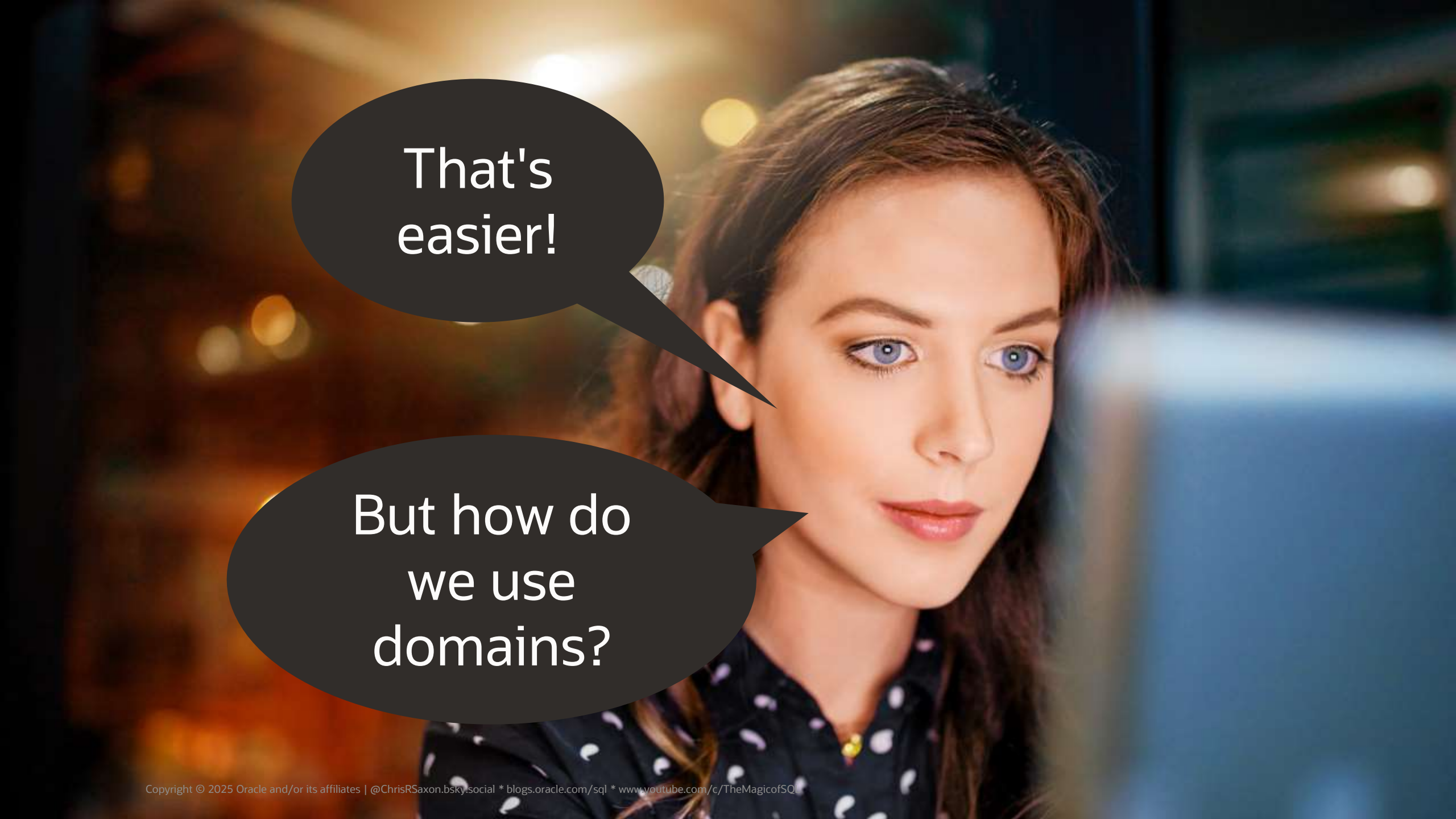
Can you fix it?

I'll use the built-ins



```
select domain_name, search_condition
from    dba_domain_constraints
where   domain_owner = 'SYS';
```

DOMAIN_NAME	SEARCH_CONDITION
-----	-----
PHONE_NUMBER_D	REGEXP_LIKE (phone_number_d, '^...\$')
EMAIL_D	REGEXP_LIKE (email_d, '^...\$')
DAY_SHORT_D	UPPER(day_short_d) IN (...)
DAY_D	UPPER(day_d) IN (...)
...	



That's
easier!

But how do
we use
domains?

```
create table customers ( ...,  
    insert_datetime insert_timestamp  
);
```

SQL standard

```
create table invoices ( ...,  
    insert_datetime domain insert_timestamp  
);
```

domain => Oracle extension

```
create table orders ( ...,  
    insert_datetime timestamp domain insert_timestamp  
);
```

A photograph of three men in an office setting. Two men are in the foreground, looking at a laptop screen. A third man is in the background, looking towards them. Overlaid on the image are four speech bubbles containing text related to database schema management.

We have a
schema!

Shall we
rebuild it?

dbms_redef?

Calm down!

```
alter table products  
  modify ( contacteaddr )  
  add domain sys.email_d;
```

Applies constraint,
default, etc. to the
columns

```
alter table orders  
  modify ( e_addr domain sys.email_d );
```



```
select table_name, search_condition
from   user_constraints
where  domain_name = 'EMAIL_D';
```

TABLE_NAME	SEARCH_CONDITION
-----	-----
ORDERS	REGEXP_LIKE ("E_ADDR", '^...\$')
INVOICES	REGEXP_LIKE ("CUSTOMER_EMAIL", '^...\$')
CUSTOMERS	REGEXP_LIKE ("EMAIL", '^...\$')
PRODUCTS	REGEXP_LIKE ("CONTACTEADDR", '^...\$')

DEMO



Data Intention Language (DIL) - Use Case Domains

Domains

Standardize

Built-ins

Common vals



The sales
report is a
mess...

ORDER_ID	PRODUCT_ID	TOTAL_PAID	CURRENCY_CODE
3	3	8.99	EUR
2	2	8.99	GBP
1	1	8.99	NA
5	5	9.99	USD
4	4	1399.00	JPY
Total		1435.96	

Need standard
currency (EUR)

ORDER_ID	PRODUCT_ID	TOTAL_PAID	CURRENCY_CODE	
3	3	8.99	EUR	2 char currency?!
2	2	8.99	GBP	
1	1	8.99	NA	
5	5	9.99	USD	
4	4	1399.00	JPY	
Total		1435.96		



Fix it!



Mandatory

Amount
Currency
EUR ex rate

Codes

3 uppercase
letters

Normalize

Convert to EUR



So we need
3 domains?

... hold on ...

Multicolumn Domains

Describe composite values



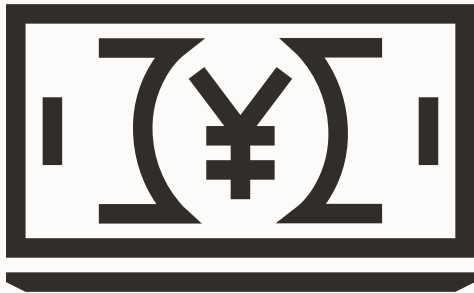
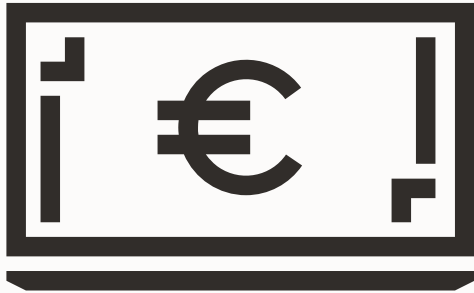
```
create domain <multicolumn_domain_name> as (  
    <column1> as <datatype> [ properties ],  
    <column2> as <datatype> [ properties ],  
    <column3> as <datatype> [ properties ],  
    ...  
)
```



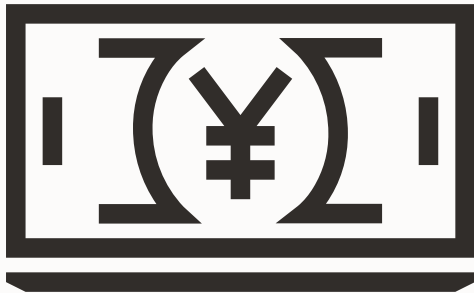
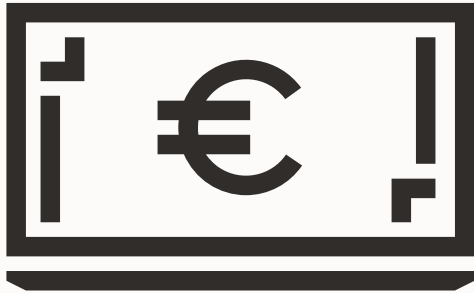
```
create domain currency as (  
    amount                as number(10,2)  
    not null,  
    iso_currency_code as char(3 char) strict  
    constraint curr_code_3_chars_c  
    check ( regexp_like ( iso_currency_code, '[A-Z]{3}' ) )  
    not null,  
    exchange_rate        as number  
    not null  
)
```



What's with
strict?



Amount/price (number(10,2))
ISO currency code (3 chars)
Exchange rate (number)

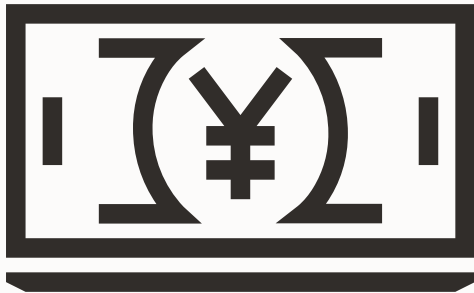
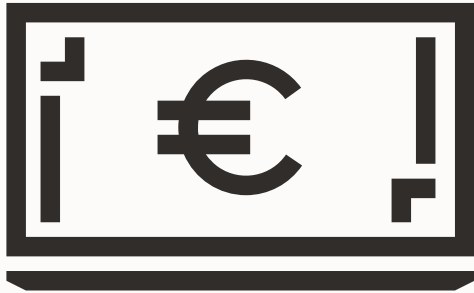


Varying units for
monetary values;
€1.99 or 199¢?

Amount/price (number(10,2))

ISO currency code (3 chars)

Exchange rate (number)



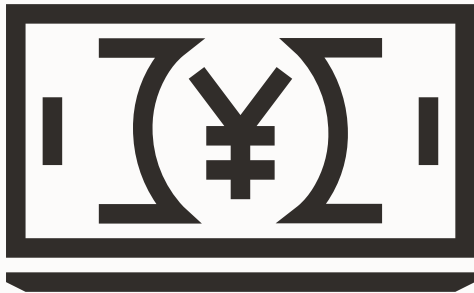
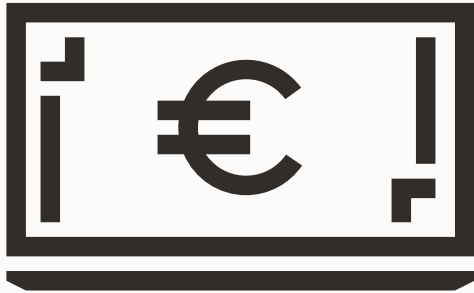
Amount/price (number(10,2))

ISO currency code (3 chars)

Exchange rate (number)

Match source API





Amount/price (number(10,2))

ISO currency code (3 chars)

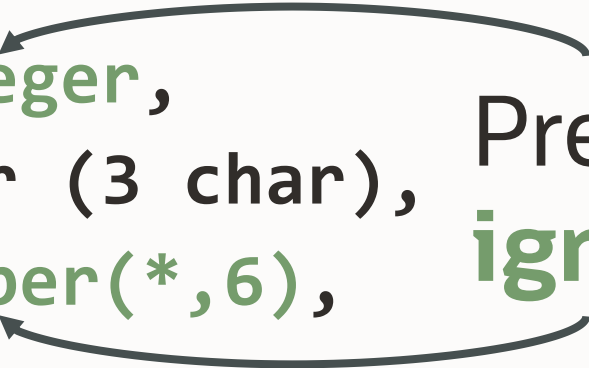
Exchange rate (number)

Must be 3 characters

strict => exact match

```
create table order_items (  
    ...,  
    total_paid          integer,  
    currency_code       char (3 char),  
    eur_exchange_rate   number(*,6),  
    domain currency (  
        total_paid, currency_code, usd_exchange_rate  
    )  
);
```

Precision & scale
ignored for non-strict





```
create table invoice_items (  
    ...,  
    total_paid          number(10, 2),  
    currency_code       char (99 char),  
    eur_exchange_rate   number,  
    domain currency (  
        total_paid, currency_code, usd_exchange_rate  
    )  
);
```

Length mismatch
for strict => **error**



```
create table invoice_items (  
    ...,  
    total_paid          number(10, 2),  
    currency_code       char (3 char),  
    eur_exchange_rate  number,  
    domain currency (  
        total_paid, currency_code, usd_exchange_rate  
    )  
);
```



```
alter table order_items
  modify (
    unit_price, currency_code, eur_ex_rate
  )
  add domain currency;
```

ORA-02293: cannot validate (%) - check constraint violated



```
create domain currency as (  
    amount                as number(10,2)  
    not null,  
    iso_currency_code as char(3 char) strict  
    constraint curr_code_3_chars_c  
    check ( regexp_like ( iso_currency_code, '[A-Z]{3}' ) )  
    novalidate  
    not null,  
    exchange_rate        as number  
    not null  
)
```

A photograph of two men in a meeting. The man on the left, wearing glasses and a beard, is pointing at a laptop screen. The man on the right is leaning in, resting his chin on his hand, and looking at the screen. A third person's head is partially visible on the right. Two speech bubbles are overlaid on the image.

Amount in
EUR virtual
columns?

Sounds
good!

Ex rate formulas

Chris

```
alter table order_items
add ( amount_in_eur as (
    unit_price * ex_rate
) );
```

Alex

```
alter table invoice_items
add ( amount_in_eur as (
    amount / ex_rate
) );
```

A group of people are gathered around a table in a meeting. A man with glasses and a beard is looking at a document. Another man is leaning over him, looking at the same document. A woman is partially visible on the right. There are three speech bubbles overlaid on the image.

we're different
again...

can domains
help us?

Where's my
bat...

Domains Functions

Rules for displaying & ordering values



```
alter domain currency  
  add order amount * exchange_rate;
```

```
alter domain currency  
  add display '(' || case iso_currency_code  
    when 'NA' then 'XXX'  
    else iso_currency_code  
  end || ')' || round ( amount * exchange_rate, 2 );
```



```
select order_id, product_id, total_paid,  
       domain_display (  
         total_paid, currency_code, usd_exchange_rate  
       ) eur_amount  
from   order_items  
order  by domain_order (  
         total_paid, currency_code, usd_exchange_rate  
       );
```

Call domain functions to
use **sort & display** rules



ORDER_ID	PRODUCT_ID	TOTAL_PAID	EUR_AMOUNT
----------	------------	------------	------------

4	4	1399 (JPY)	8.46
2	2	8.99 (GBP)	8.54
3	3	8.99 (EUR)	8.99
5	5	8.99 (XXX)	8.99
1	1	9.99 (USD)	11.49

46.47

A photograph of two men in a meeting. The man on the left, with glasses and a beard, is pointing at a laptop screen. The man on the right is leaning in, looking at the screen with his hand on his chin. A third person's head is partially visible on the right. Two speech bubbles are overlaid on the image.

Nice!
But I liked
virtual cols...

good point!



```
alter table order_items add (  
    amount_in_eur_display as ( domain_display (  
        total_paid, currency_code, eur_exchange_rate  
    ) ),  
    amount_in_eur_order as ( domain_order (  
        total_paid, currency_code, eur_exchange_rate  
    ) )  
);
```

DEMO



Data Intention Language (DIL) - Use Case Domains

Types

Single
Multicolumn
Enum
Flexible

Properties

Annotations
Constraints
Functions

Built-ins

Common values



How did this
help again?



Say what you
want an app to do

Add schema
intent

AI generates
code

Data Intention Language (DIL)

Annotations

Key-value pairs

Application metadata



Use Case Domains

Extend base types

Describe data use case



Data Intention Language (DIL)

Consistent

Define rules
once

Discovery

Info in data
dictionary

Prompt

Augment LLM
requests



So will chatgpt
take our jobs?!



Fixing a Schema with Domains and Annotations-Chris Saxon & Alex Nuijten

Please fill in your
evaluations