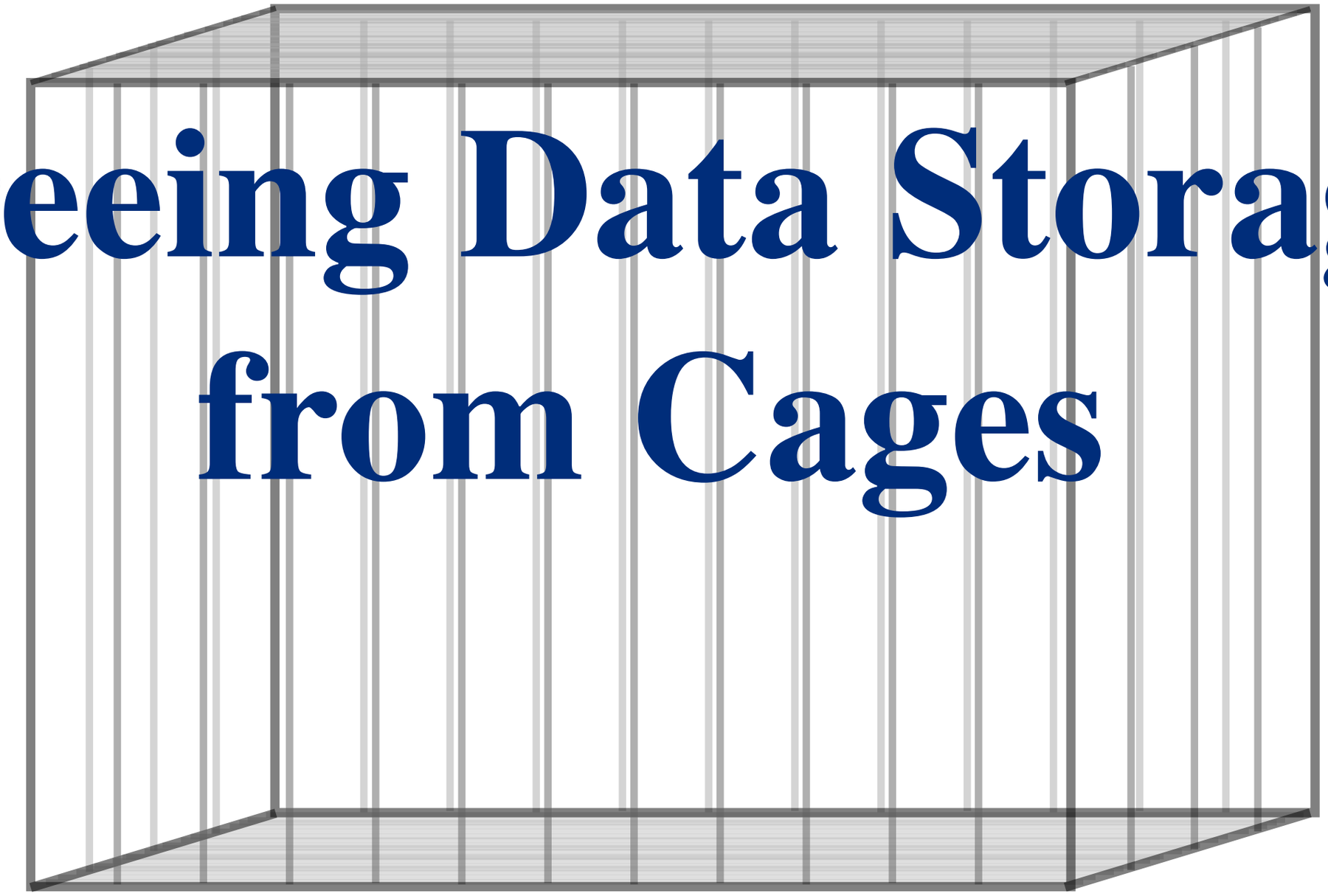


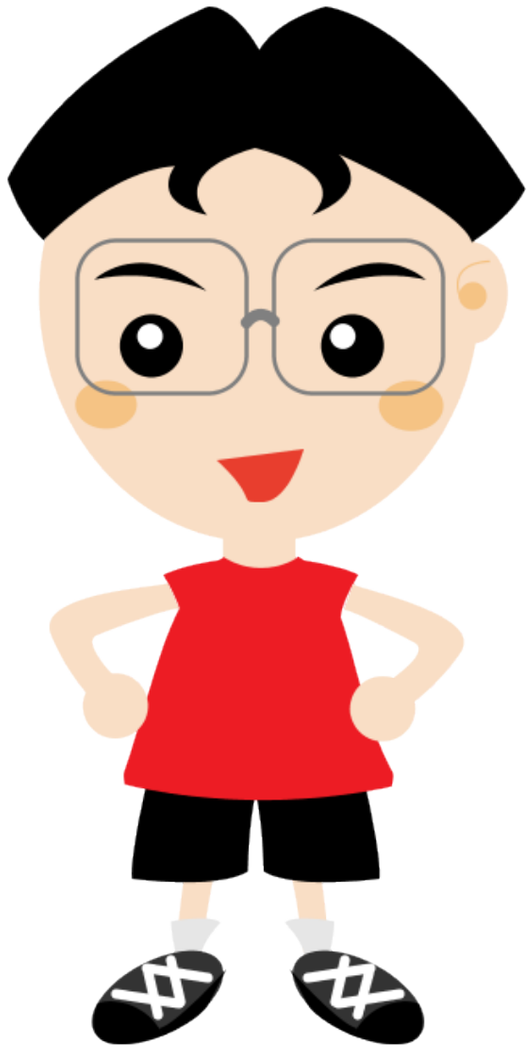
Freeing Data Storage from Cages

Alekh Jindal[★] Jorge Quiané[◆] Jens Dittrich[★]

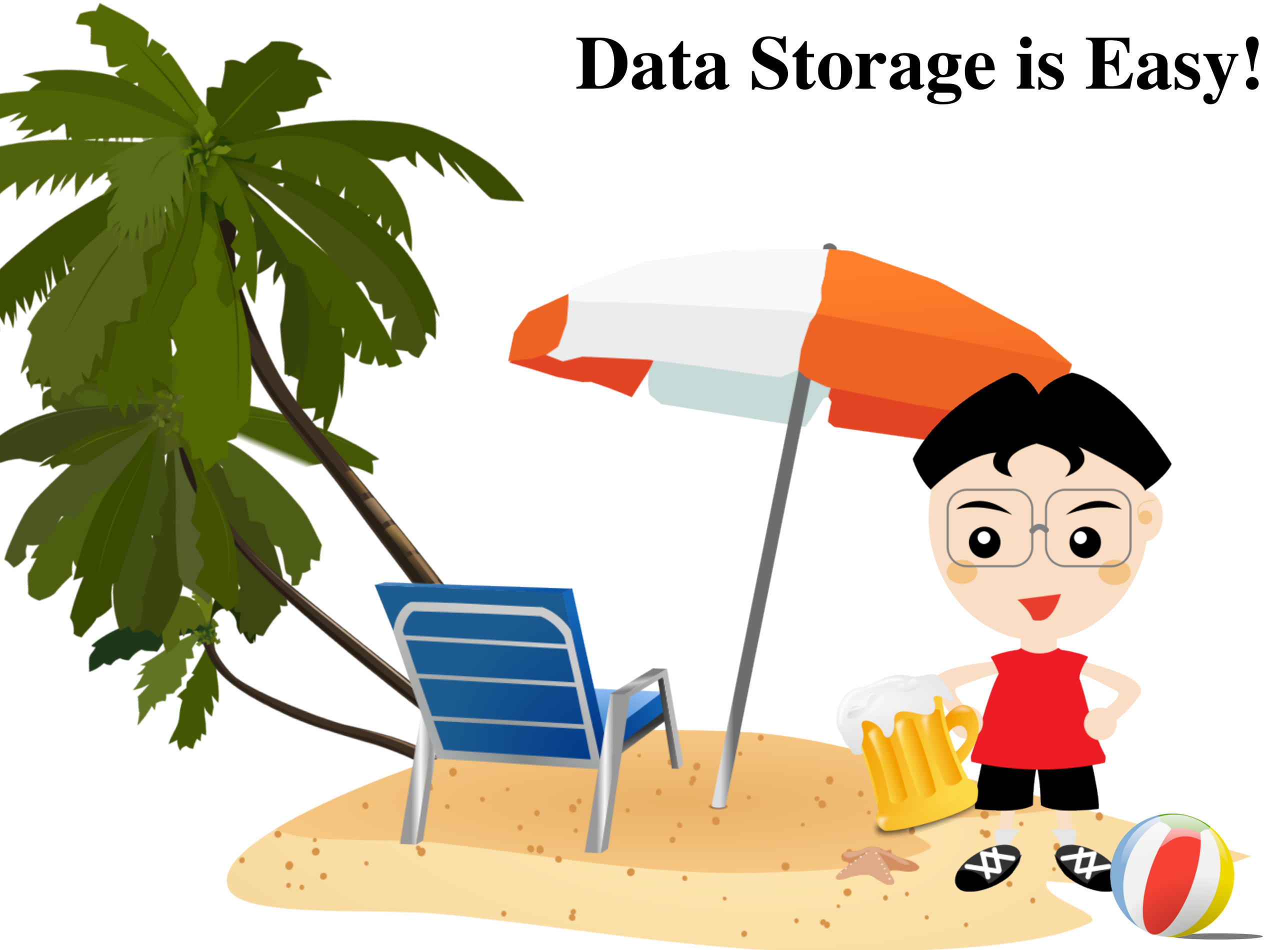
[★]Saarland University, Germany

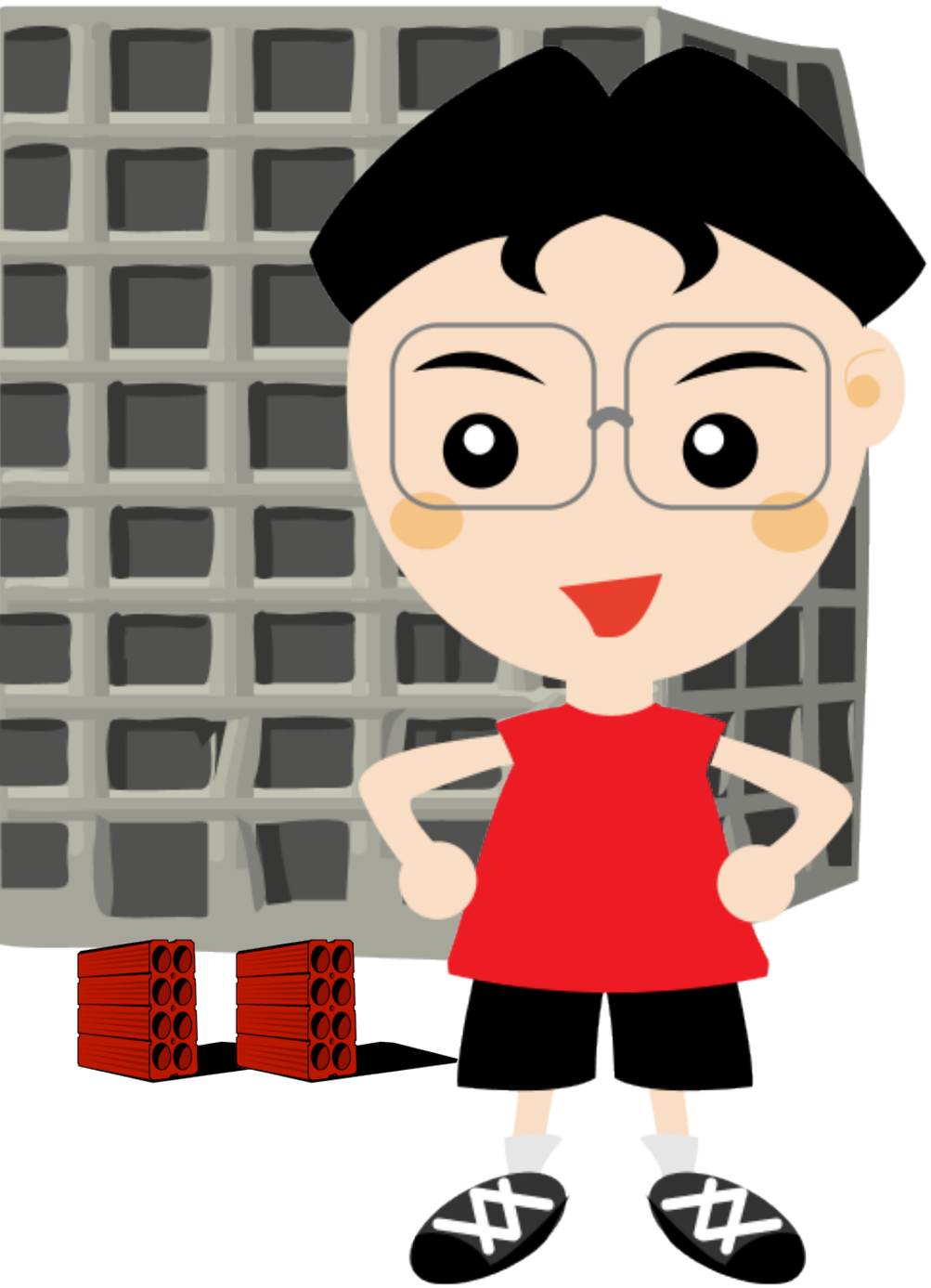
[◆]QCRI, Qatar

Data Storage is Easy!

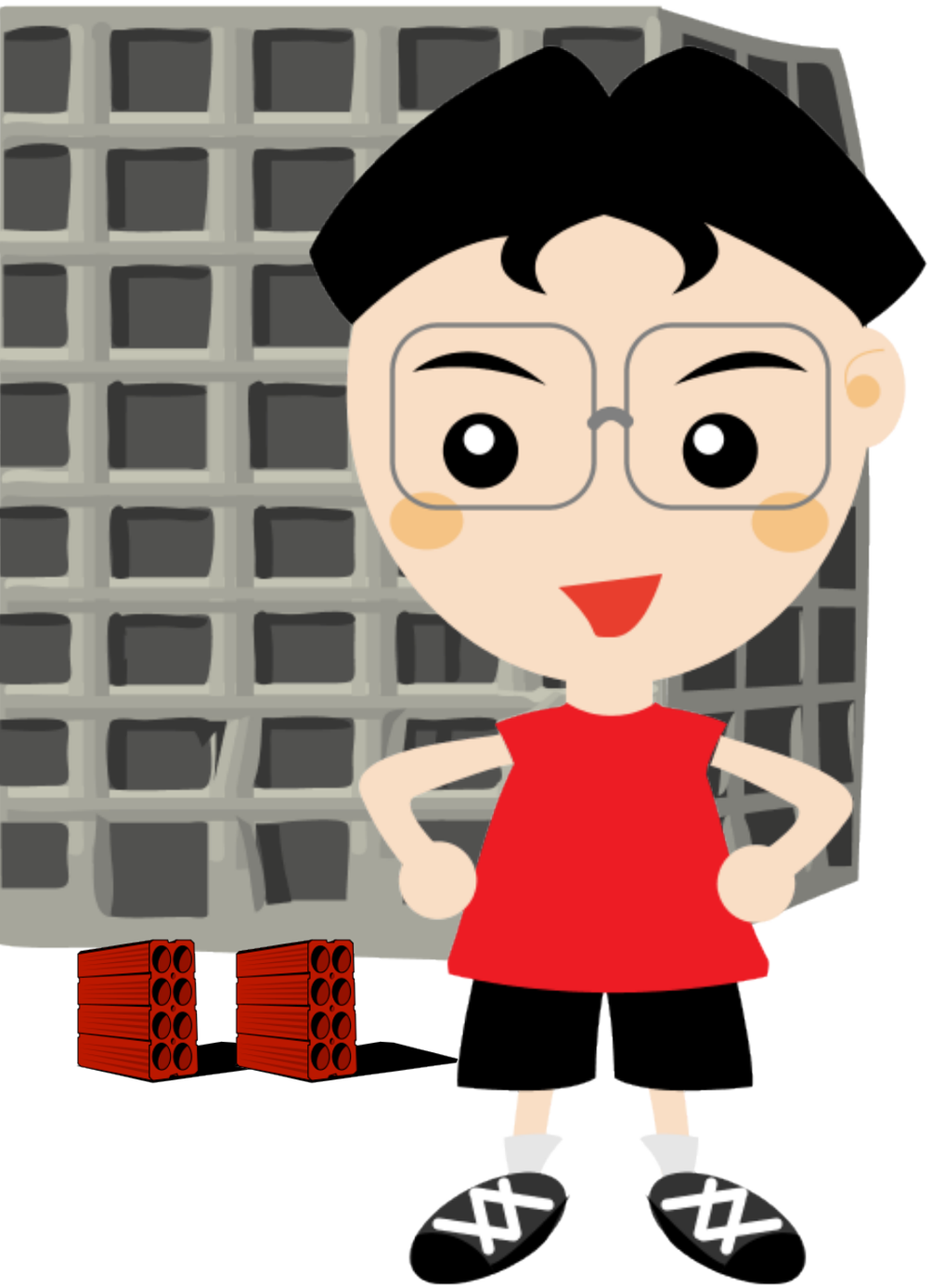


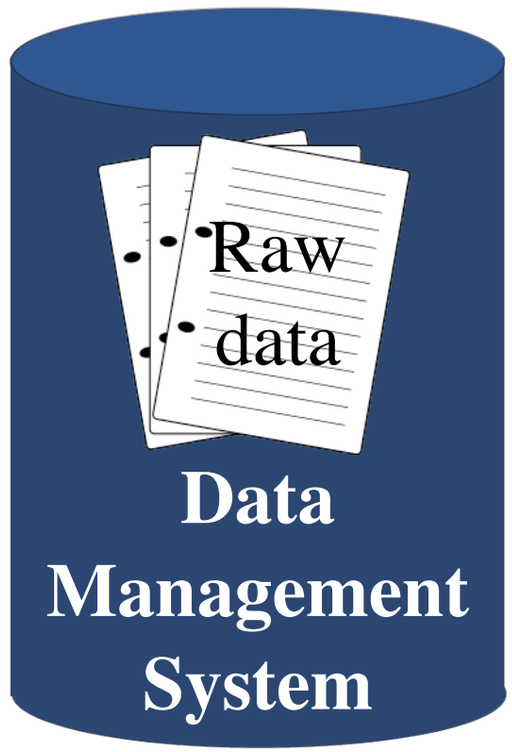
Data Storage is Easy!

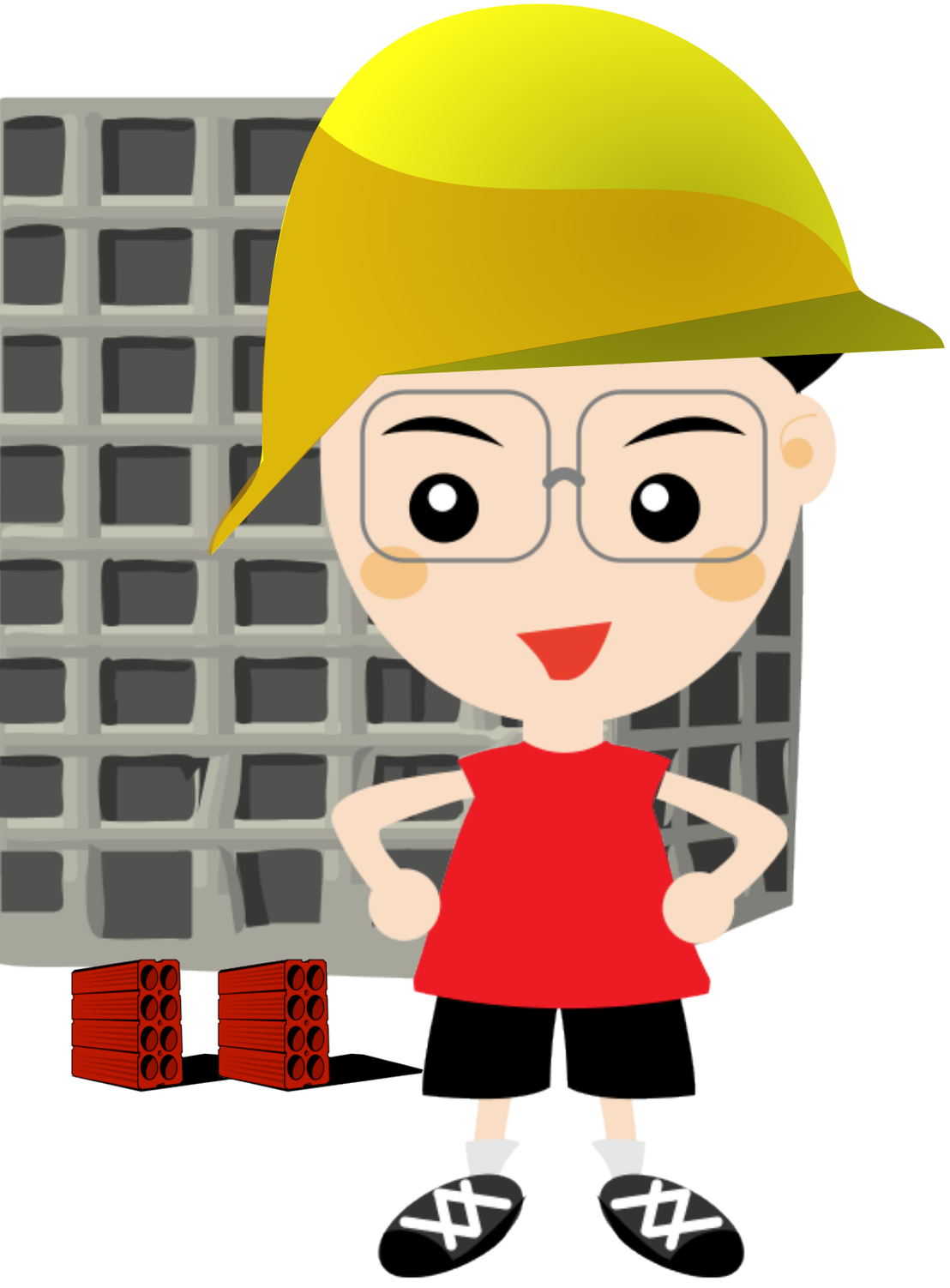




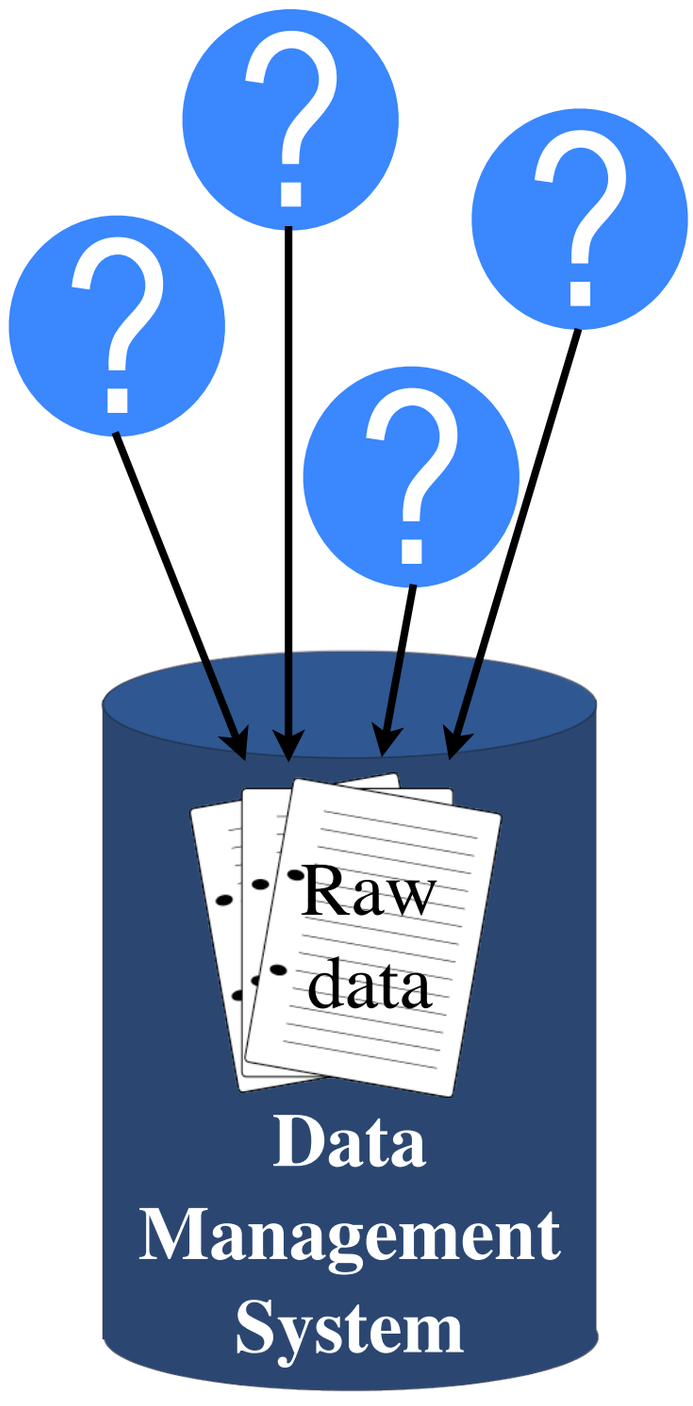

Raw
data



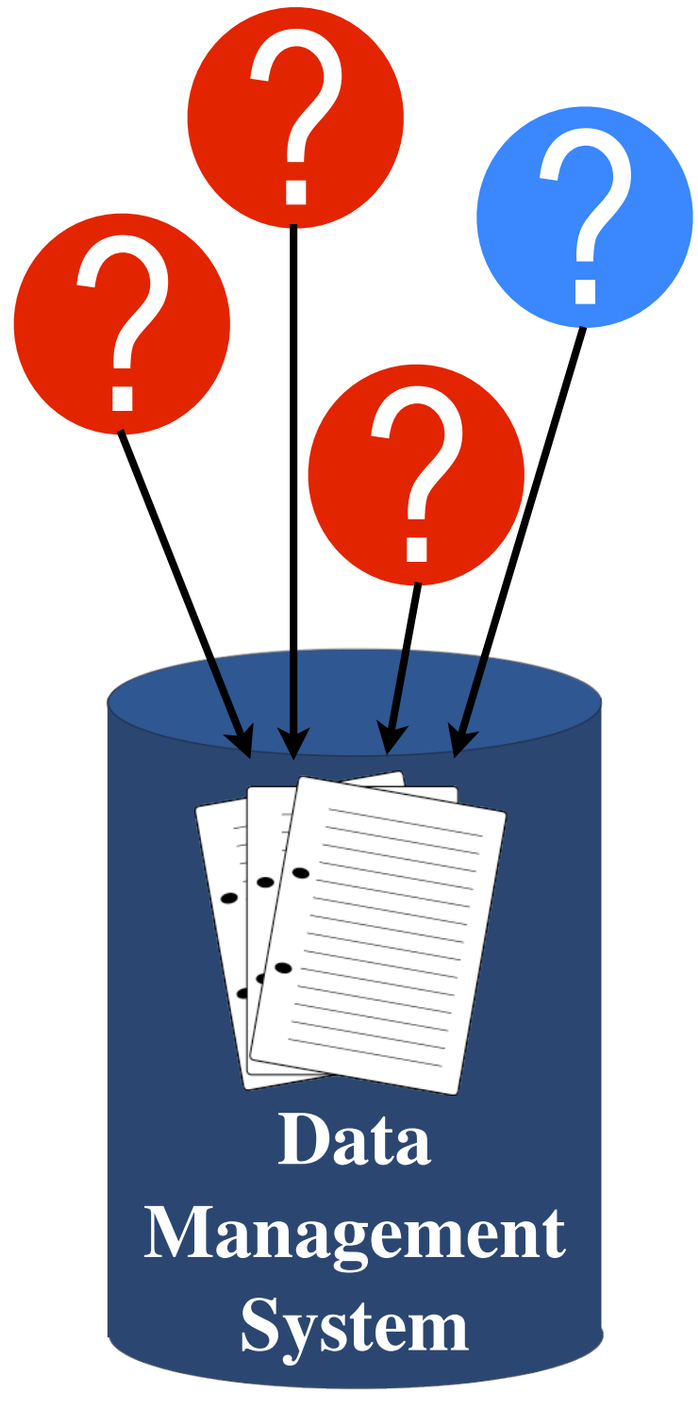

Raw
data
Data
Management
System



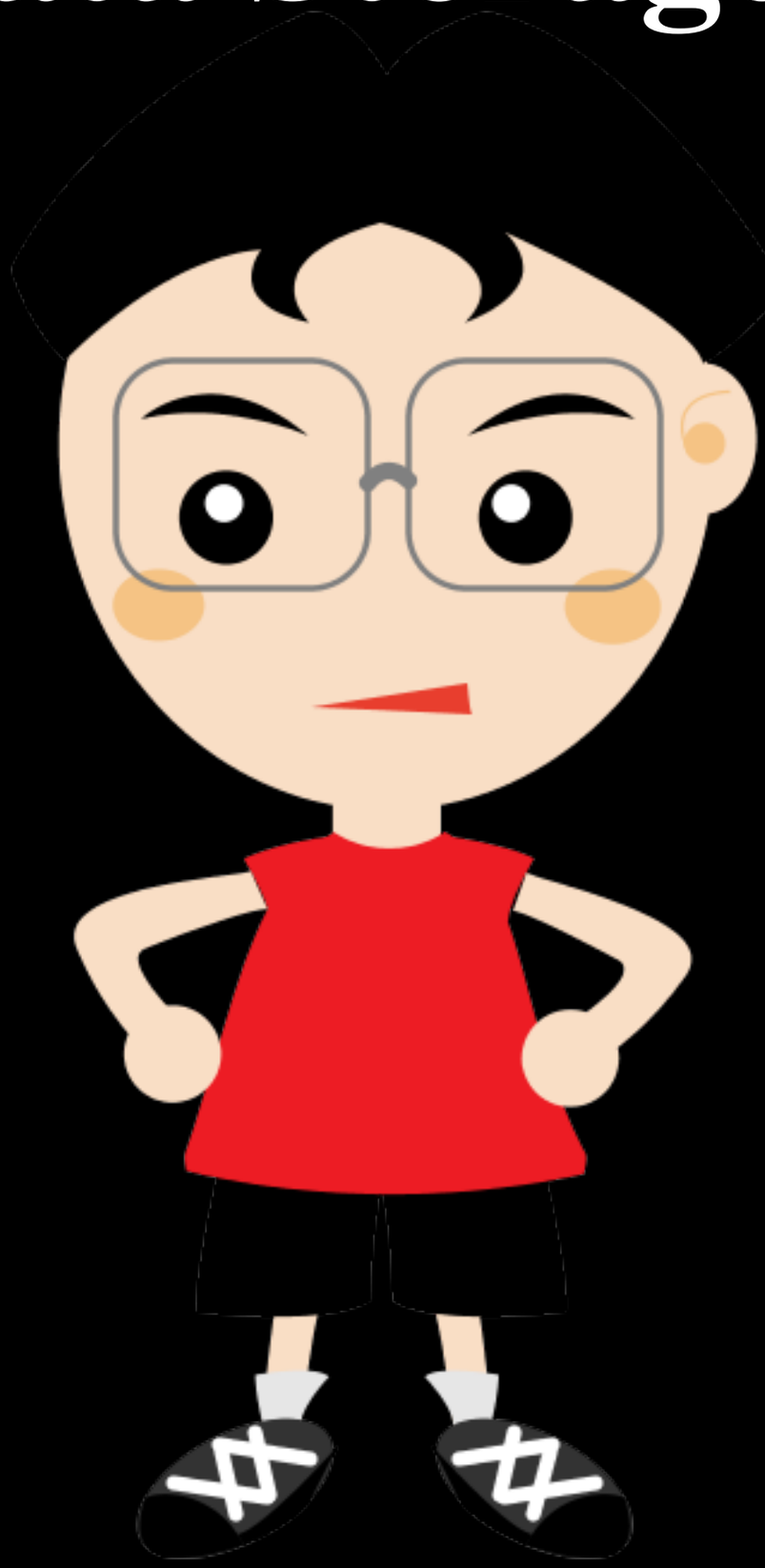
Queries

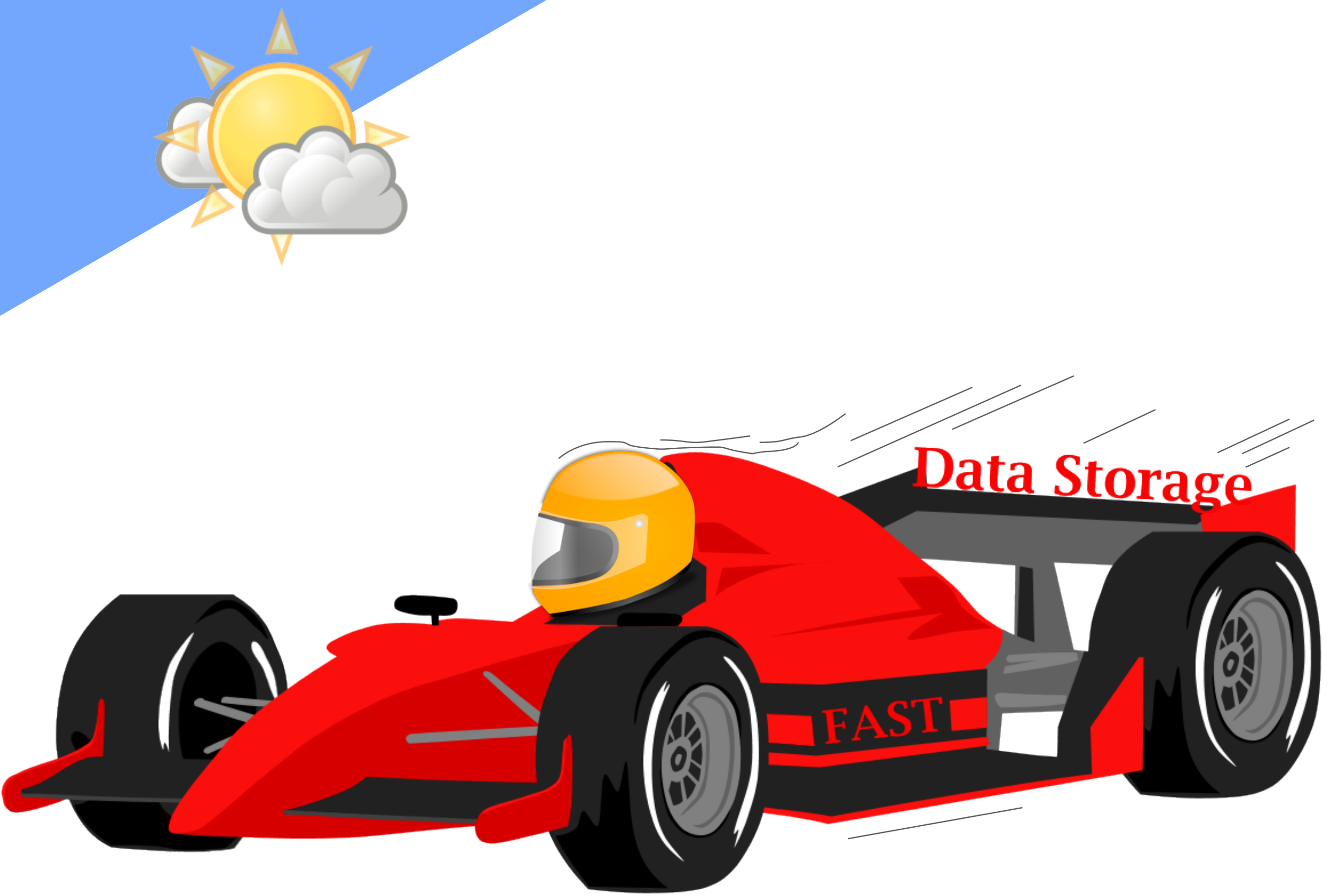


Is performance as expected?

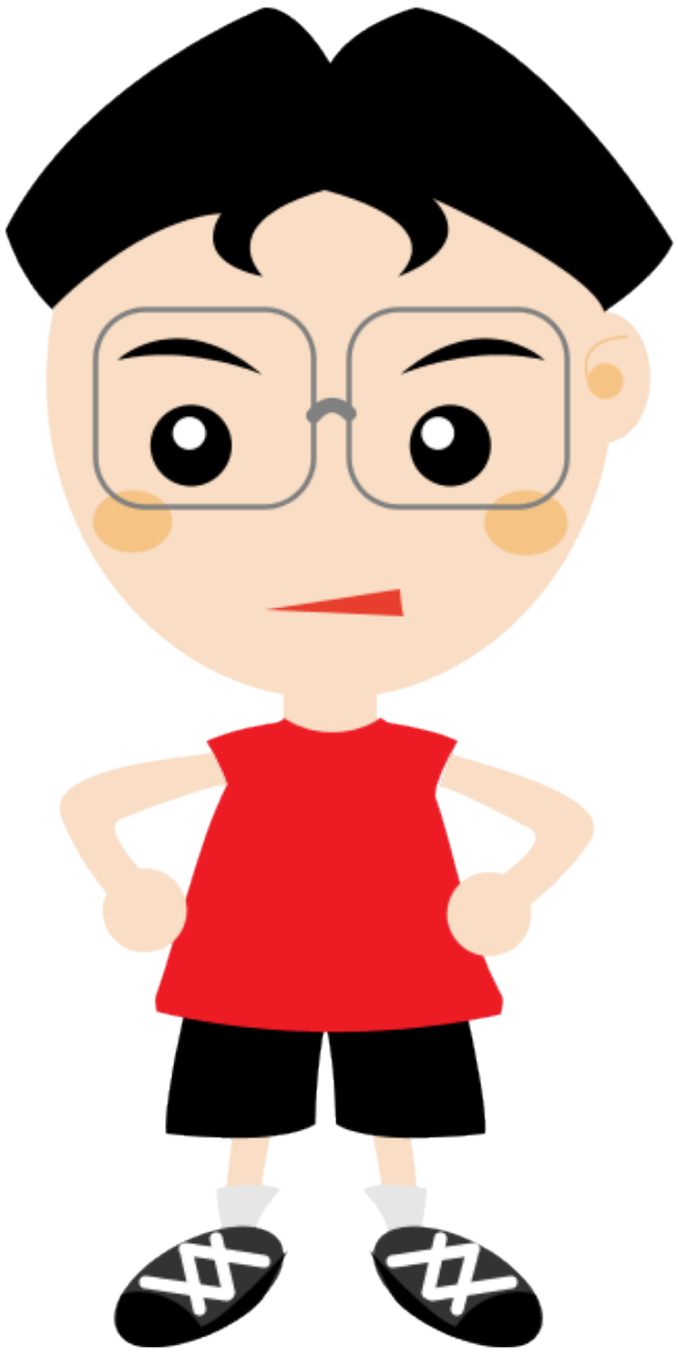


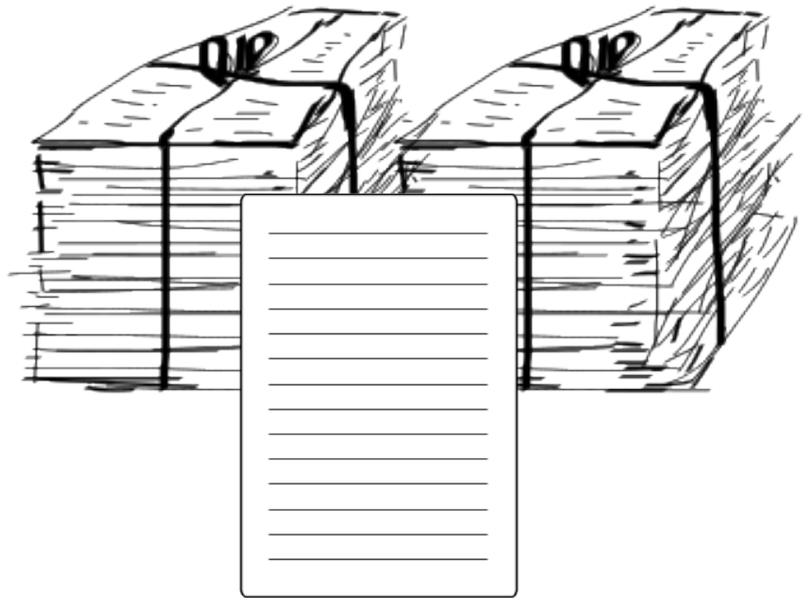
Is your Data Storage Efficient?



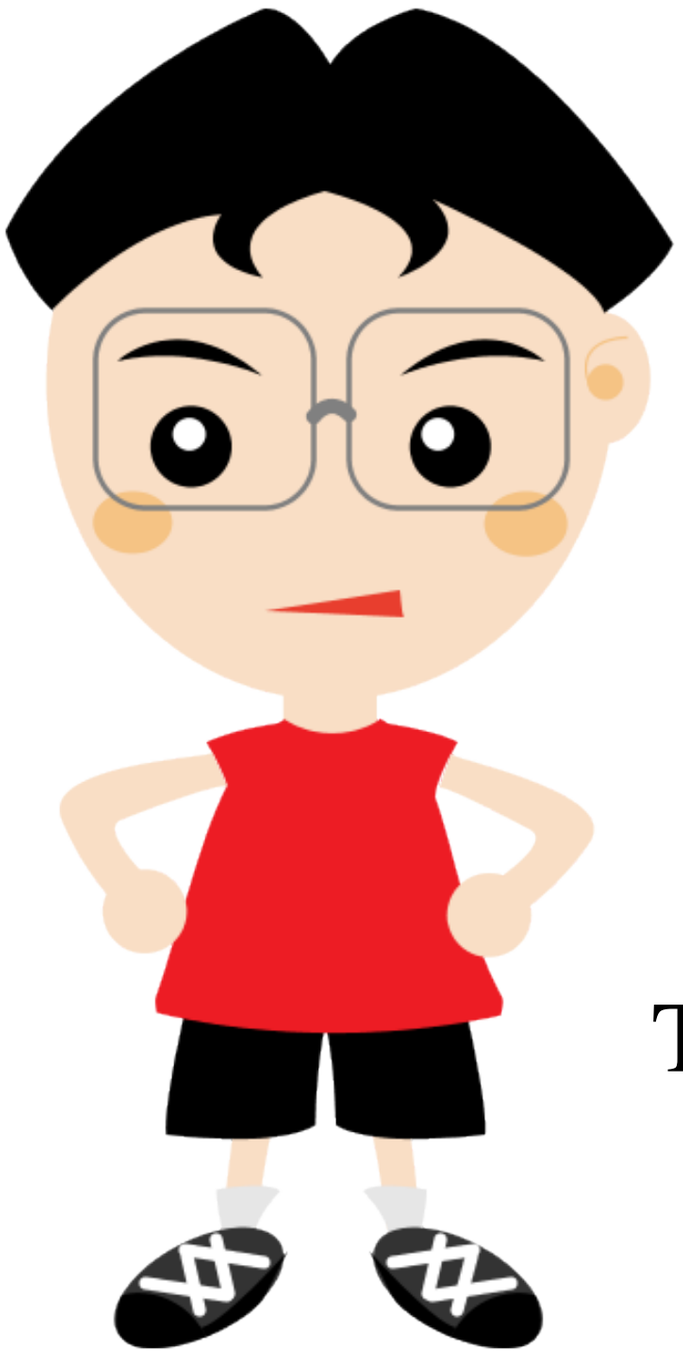

Data Storage
FAST

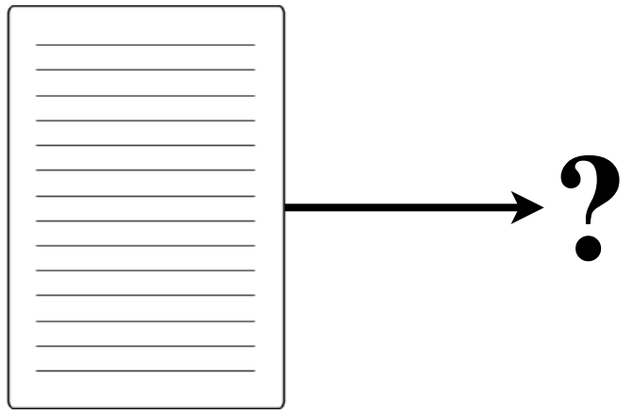
What to store?





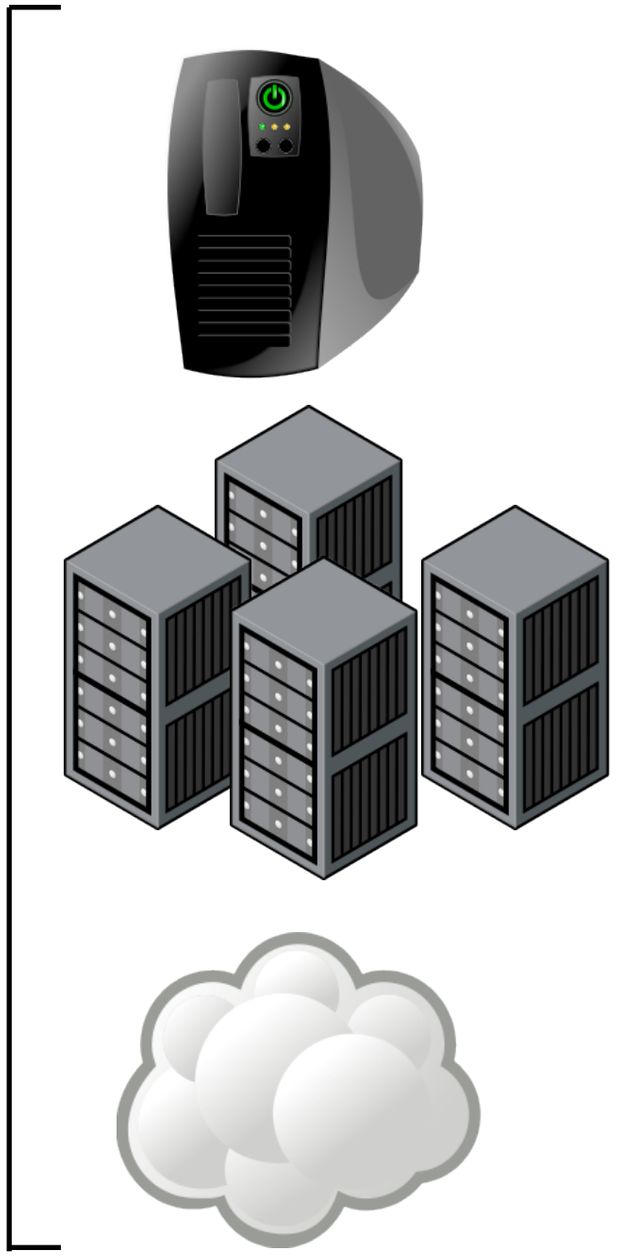
Where to store?



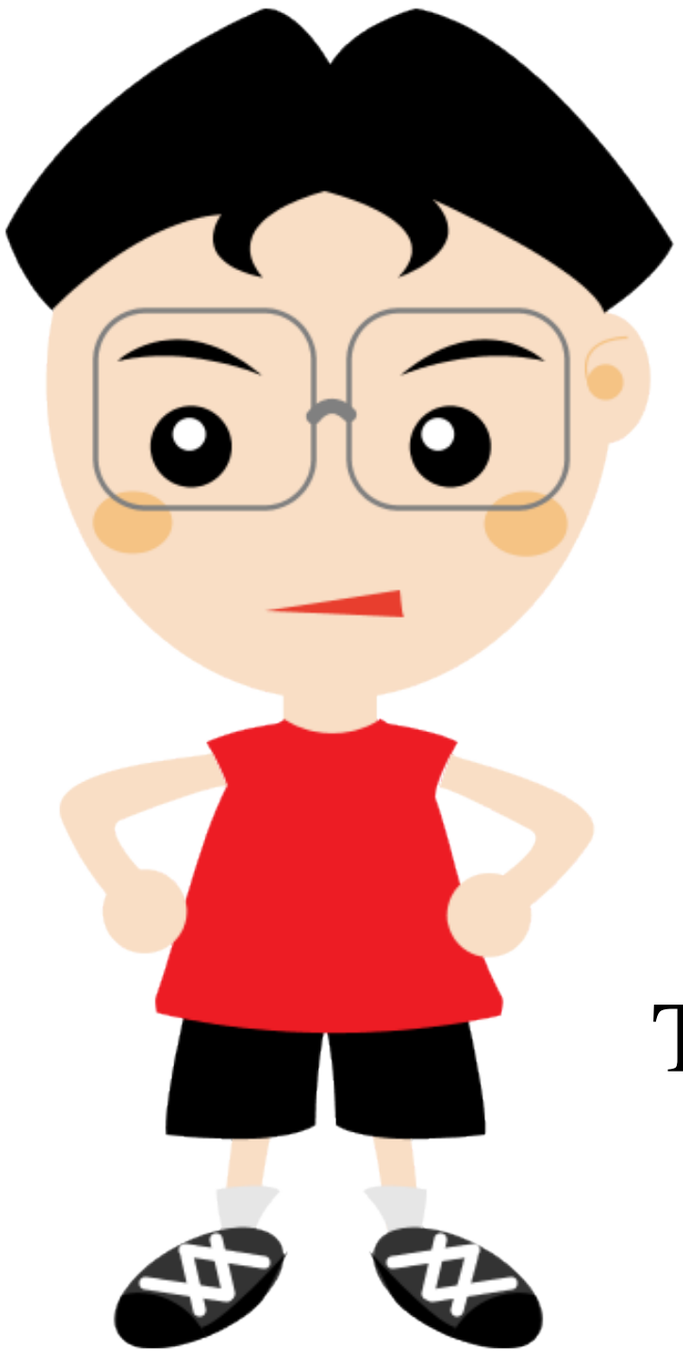


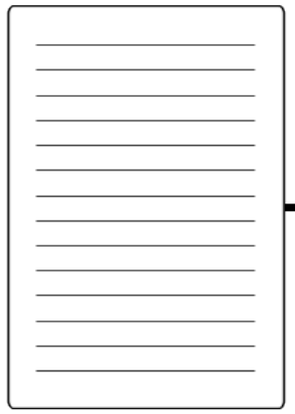
TPC-H Lineitem

?



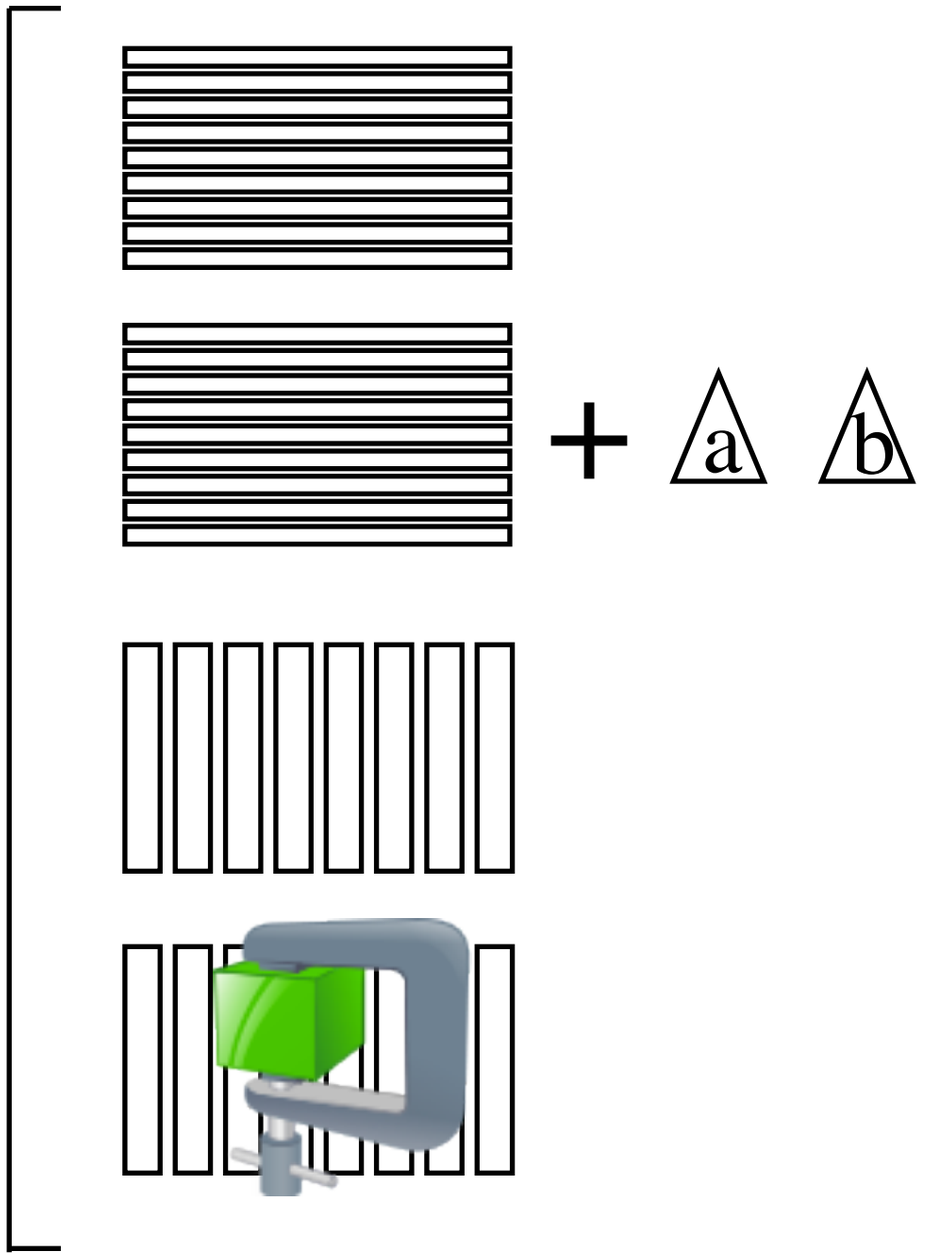
How to store?



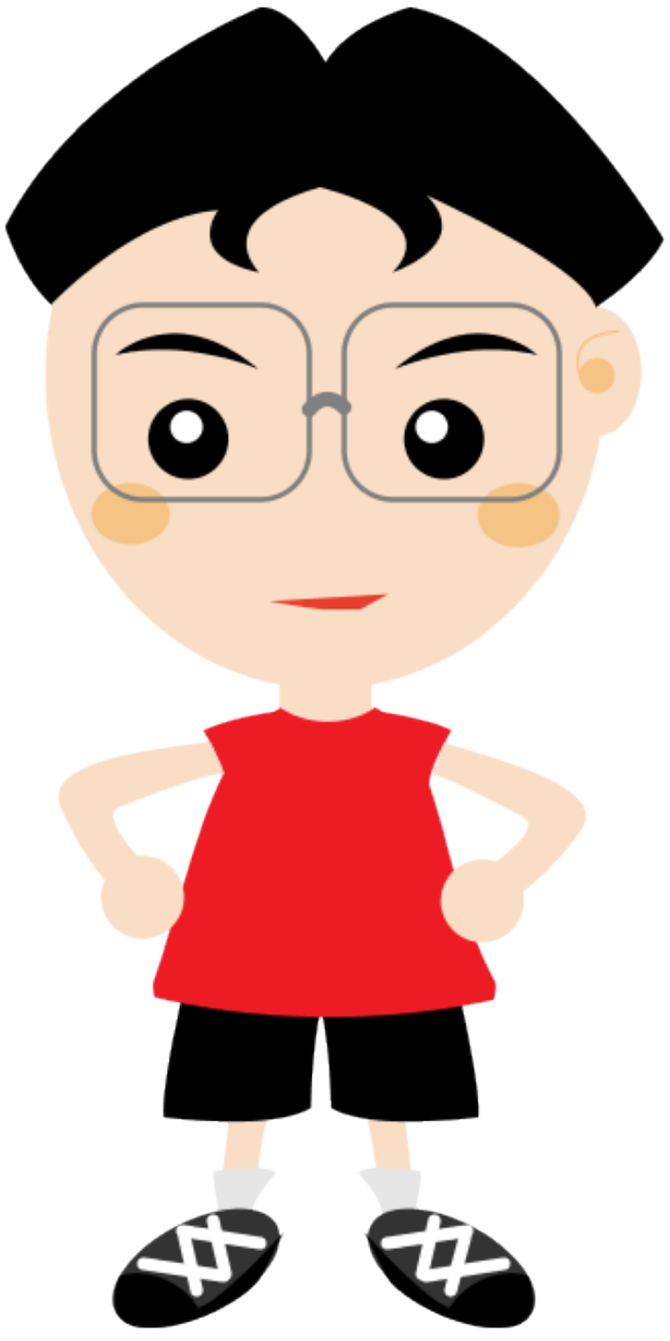


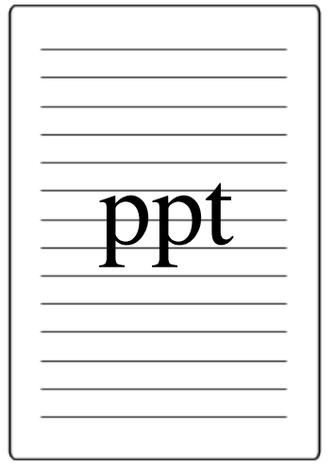
TPC-H Lineitem

?

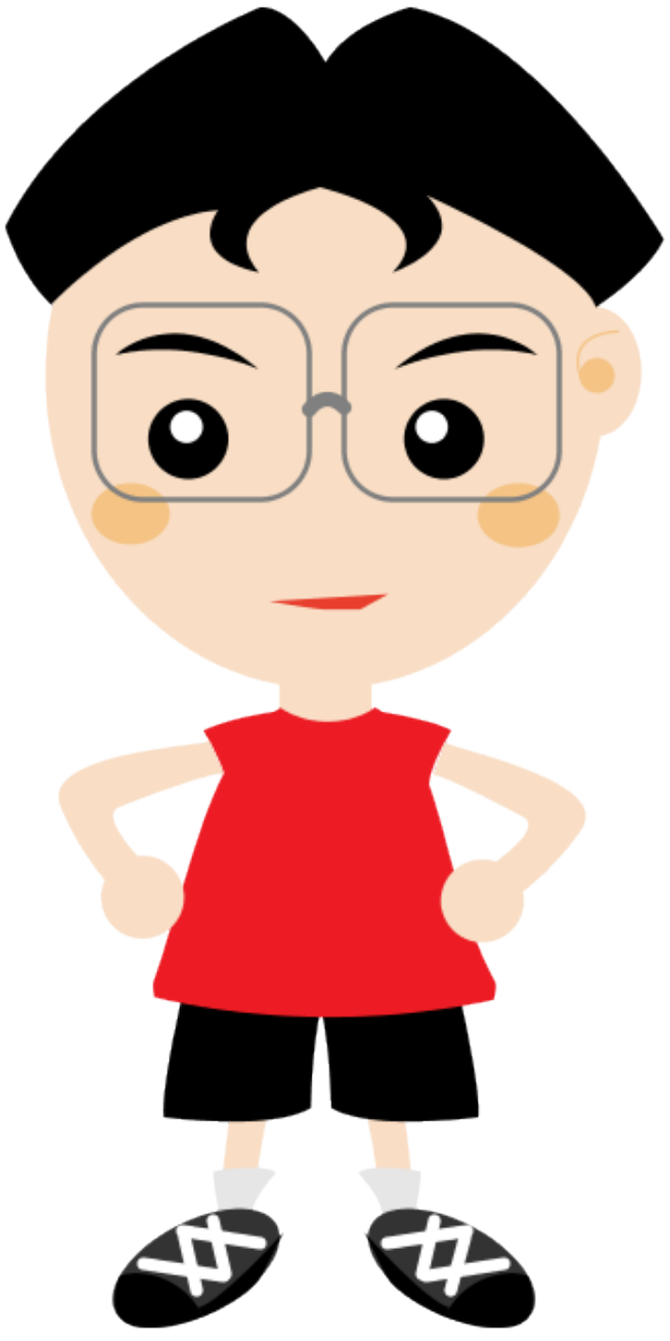


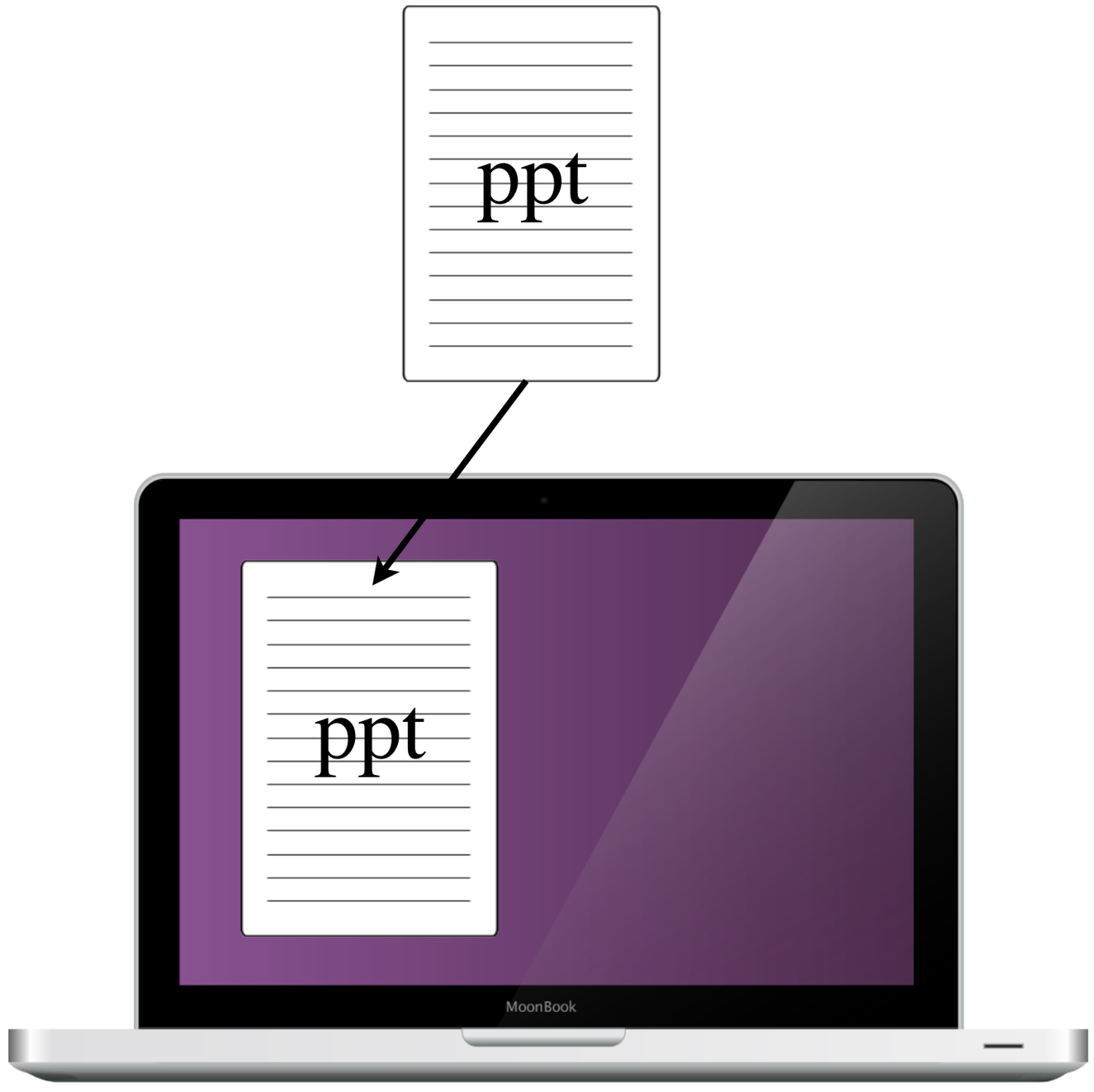
What, Where, and How in our everyday life

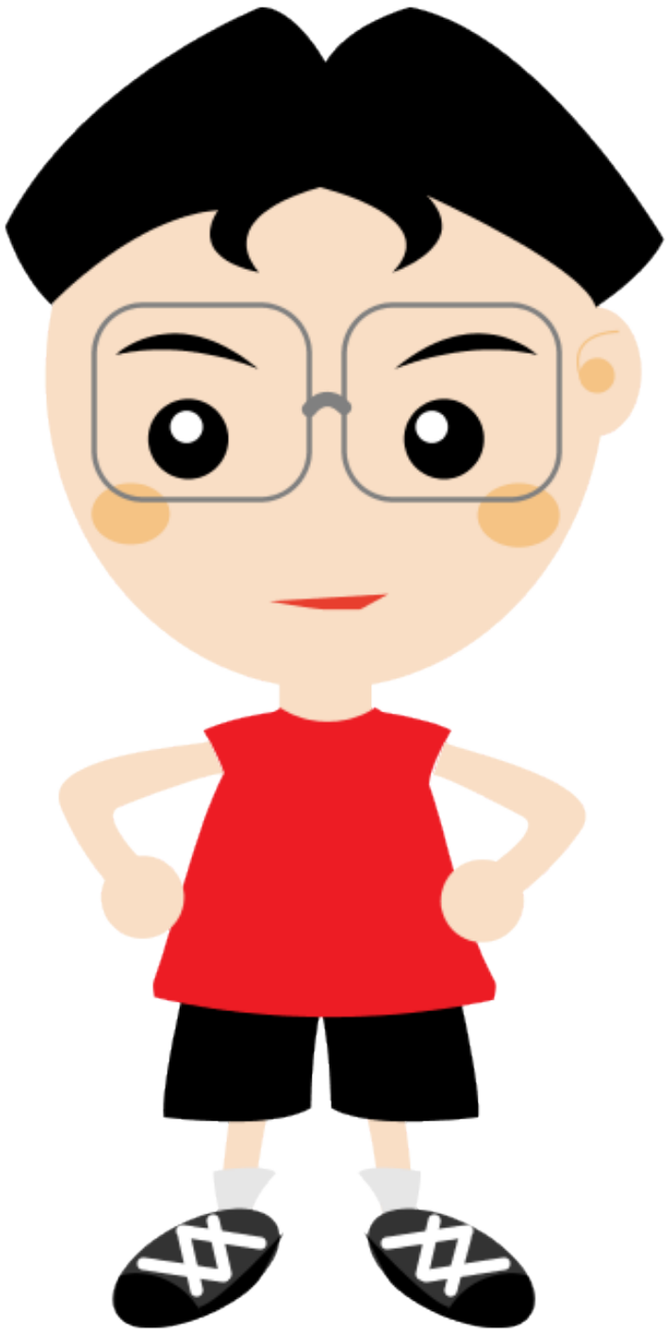


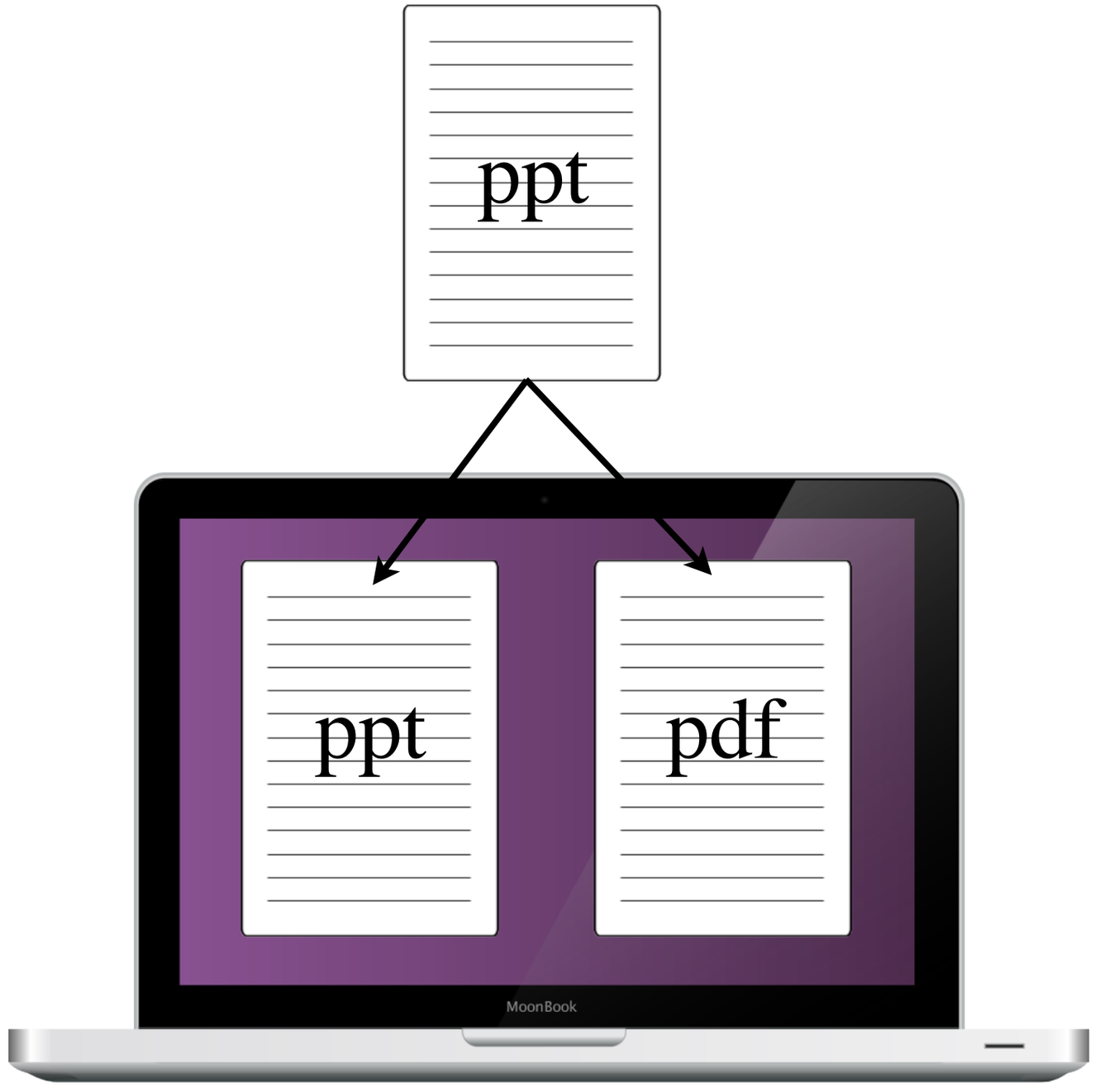

ppt


MoonBook




ppt
ppt
MoonBook

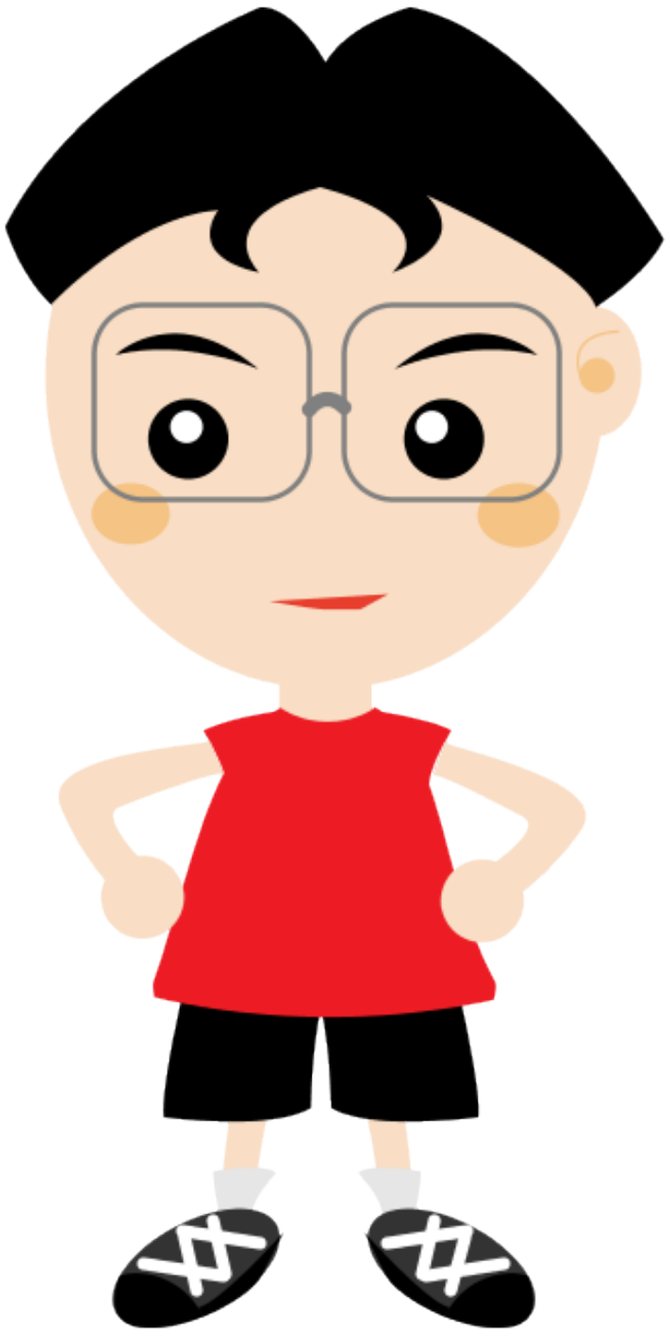


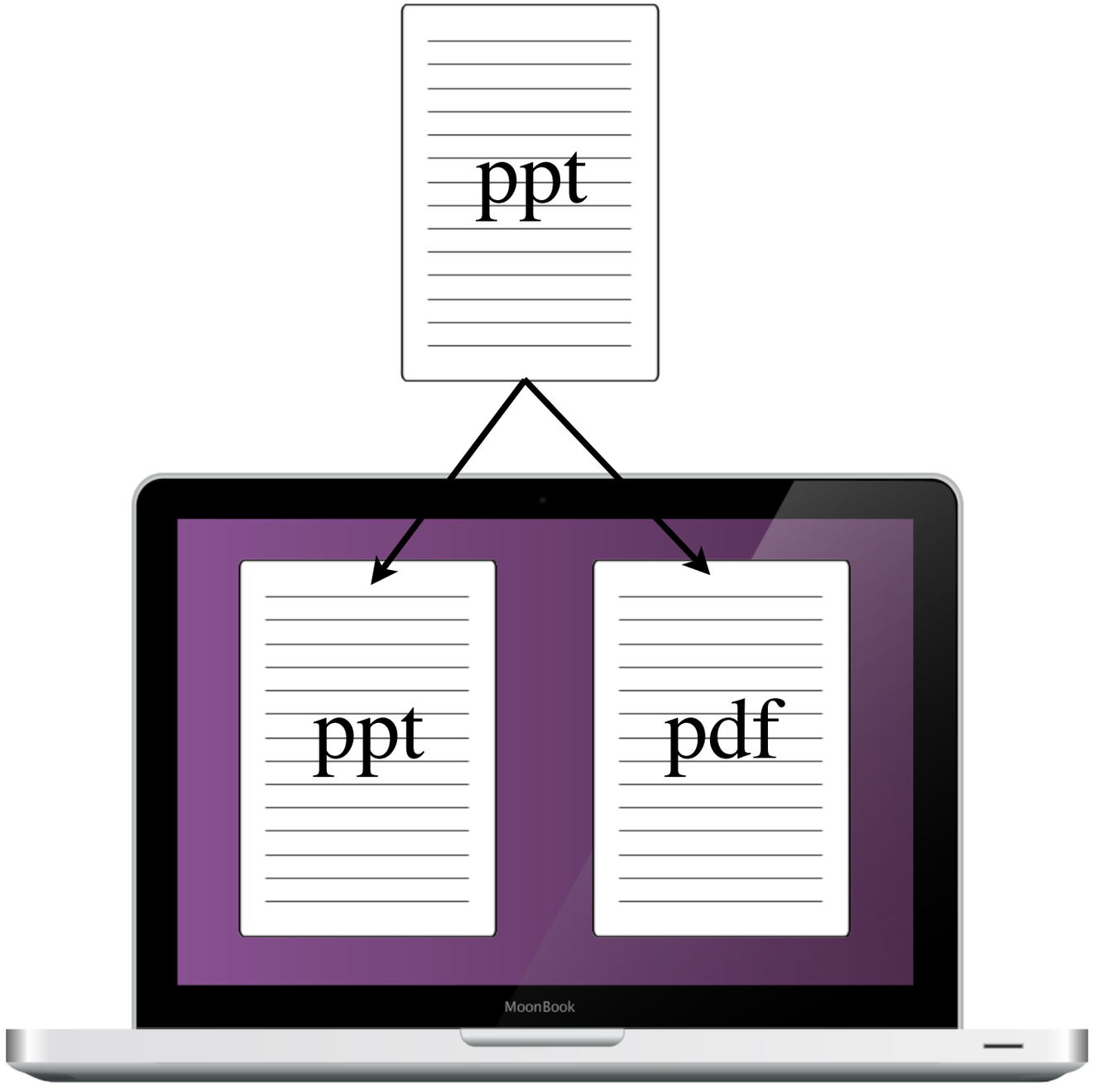

ppt
ppt
pdf
MoonBook

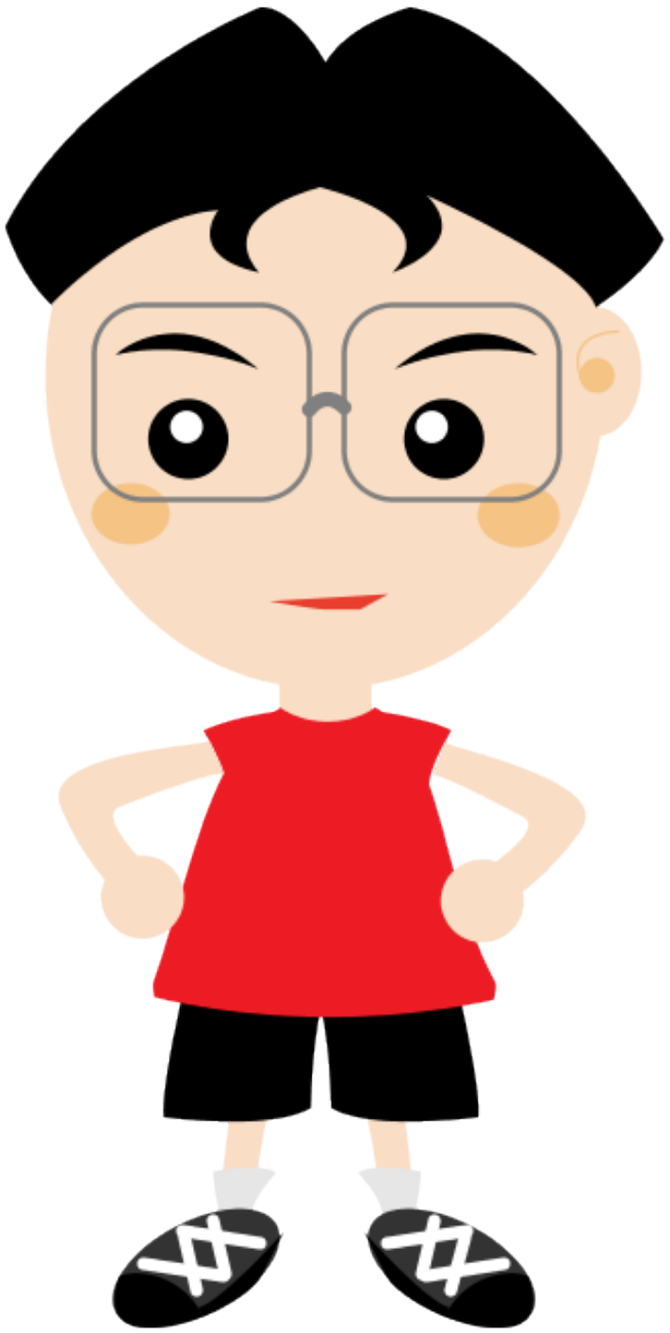

ppt
U.S. DEPARTMENT OF HOMELAND SECURITY

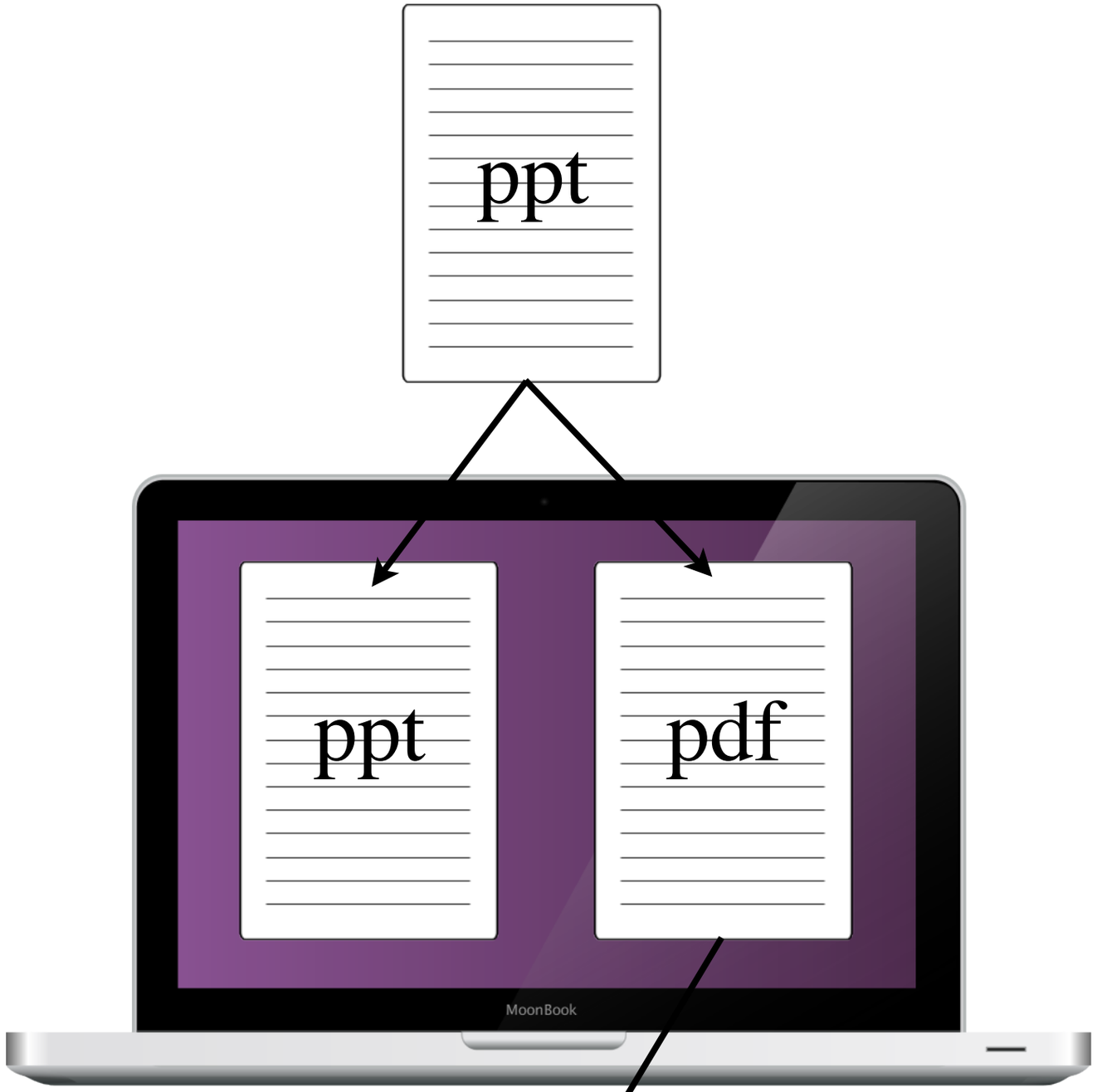

ppt
U.S. DEPARTMENT OF HOMELAND SECURITY


U.S. DEPARTMENT OF HOMELAND SECURITY
ppt
MoonBook



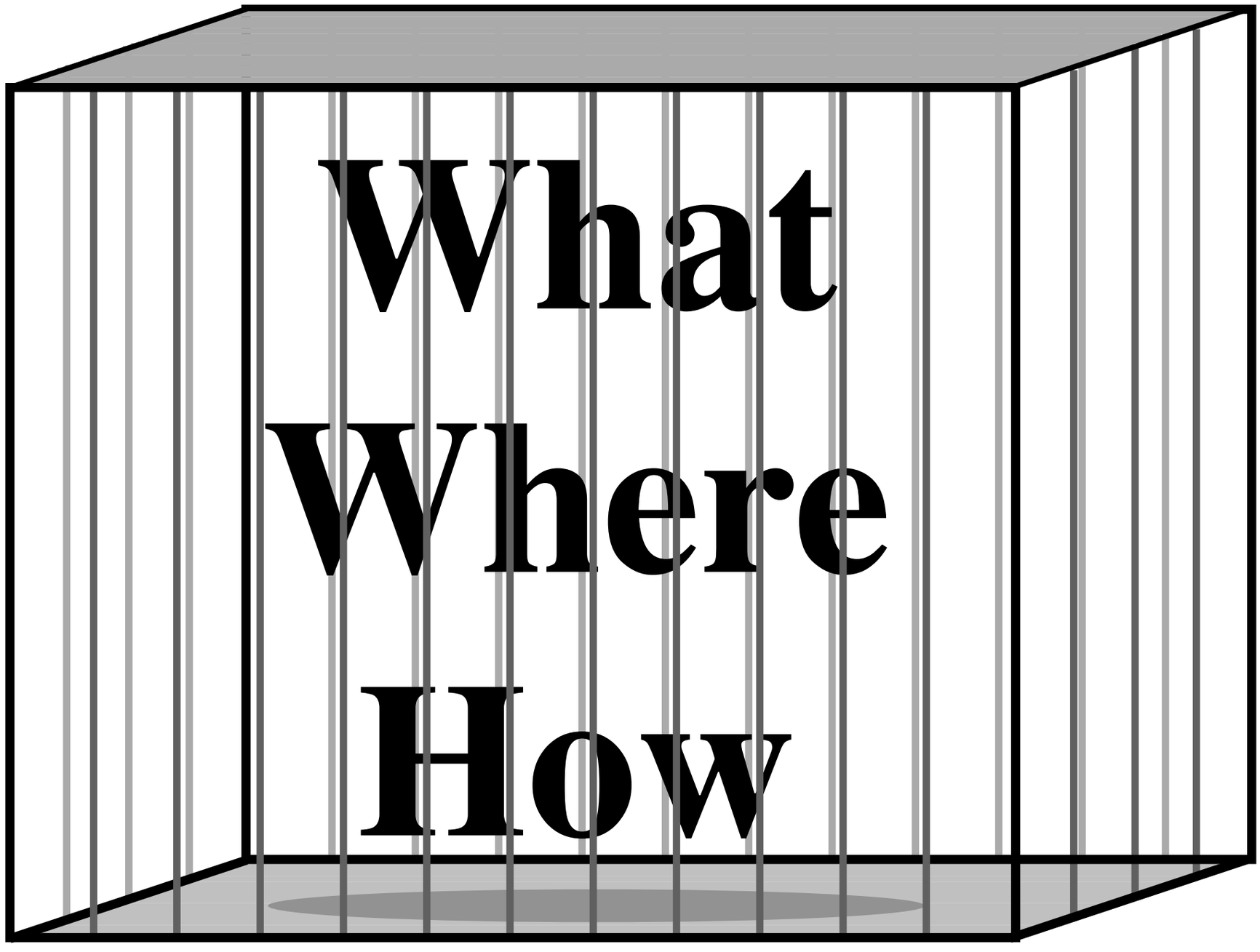

ppt
ppt
pdf
MoonBook

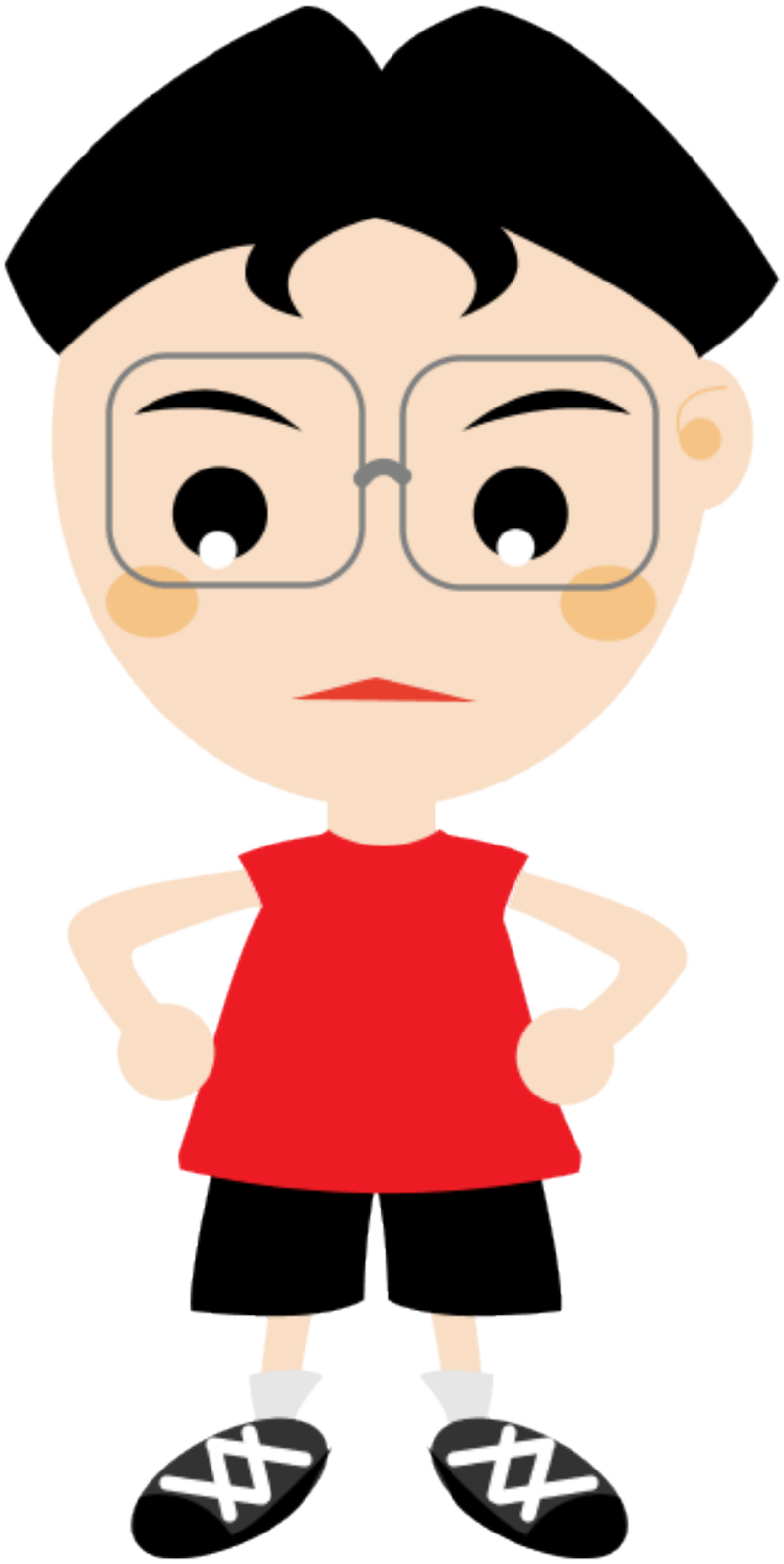



ppt
ppt
pdf
MoonBook

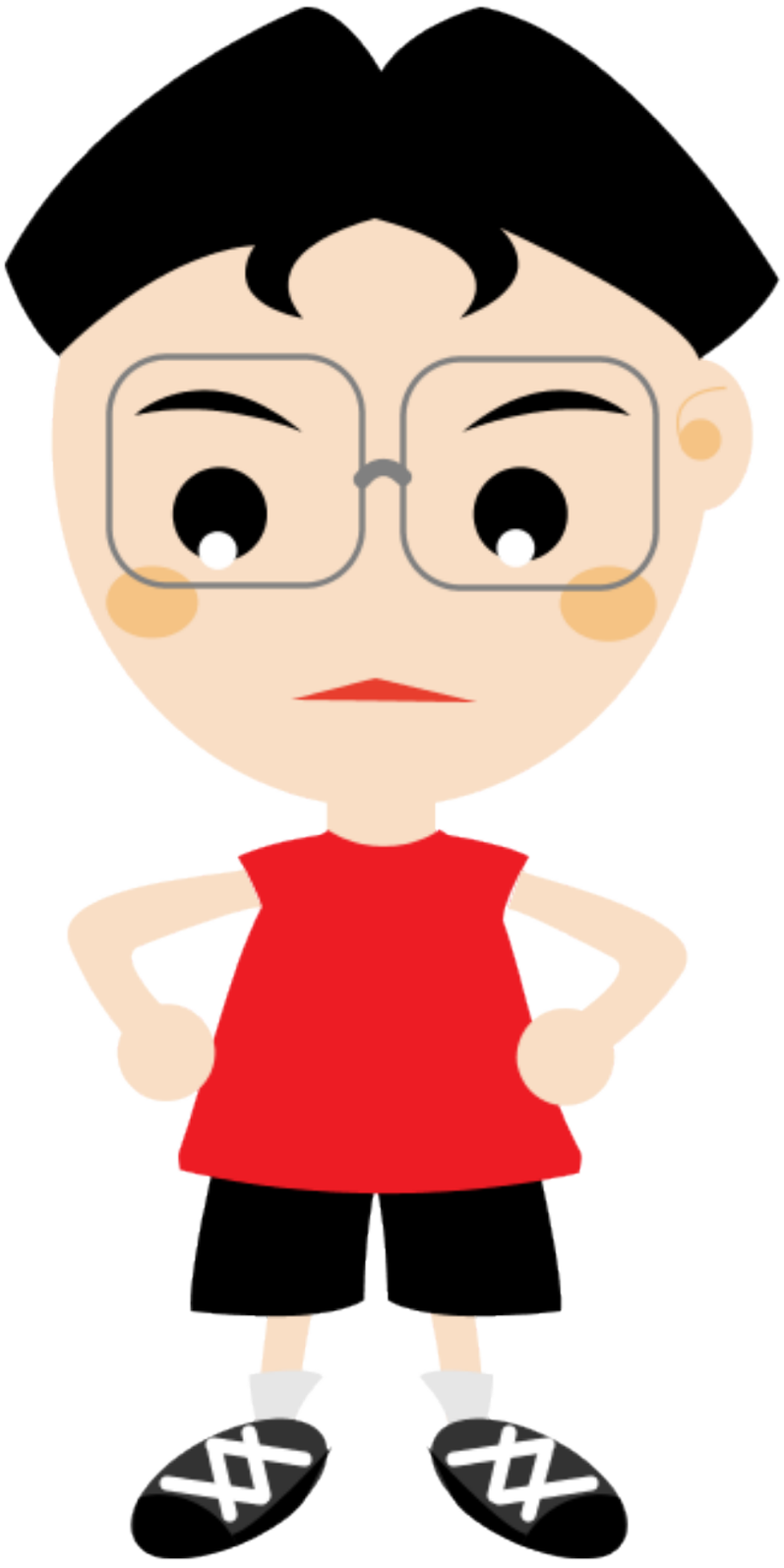

pdf

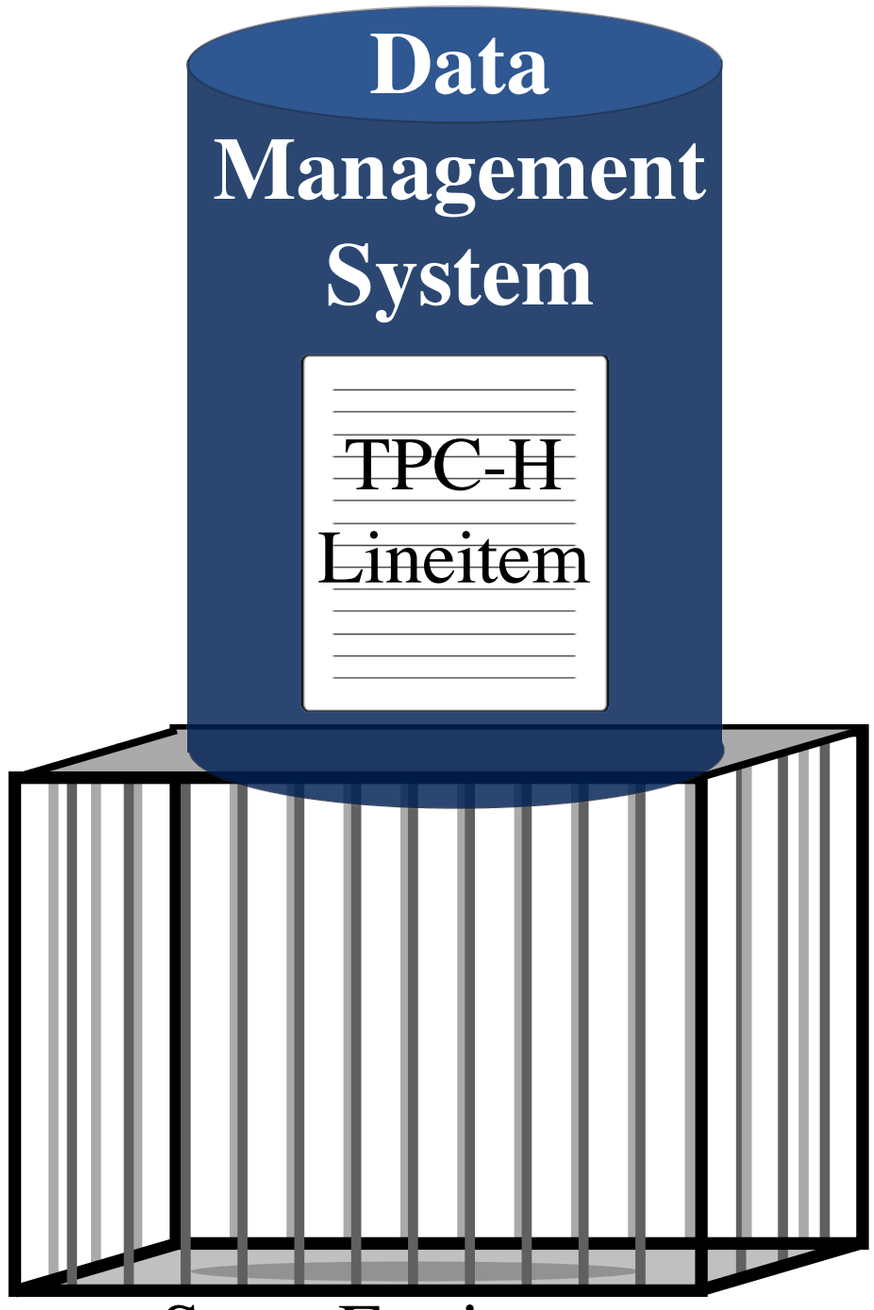
However

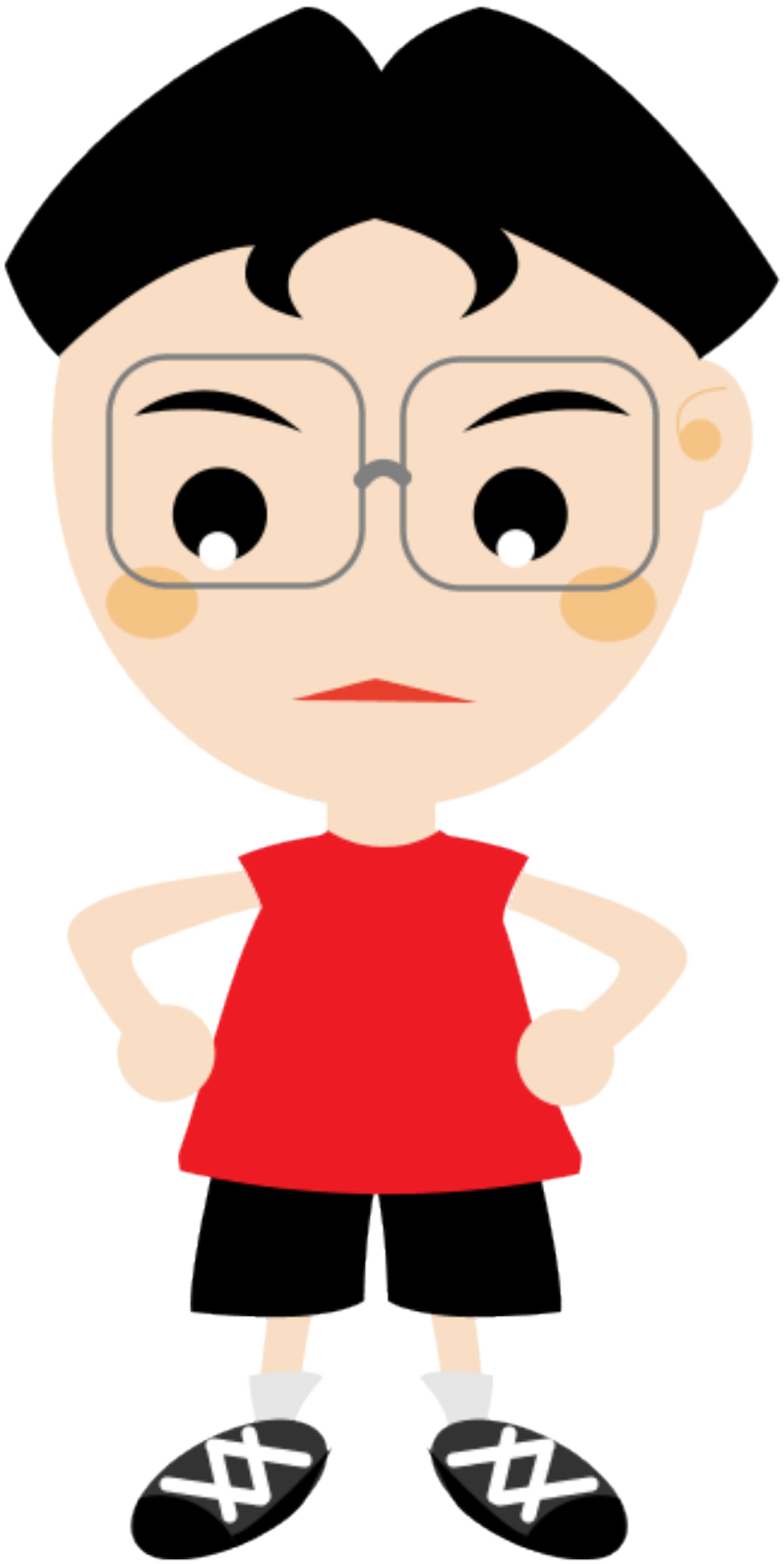

What
Where
How

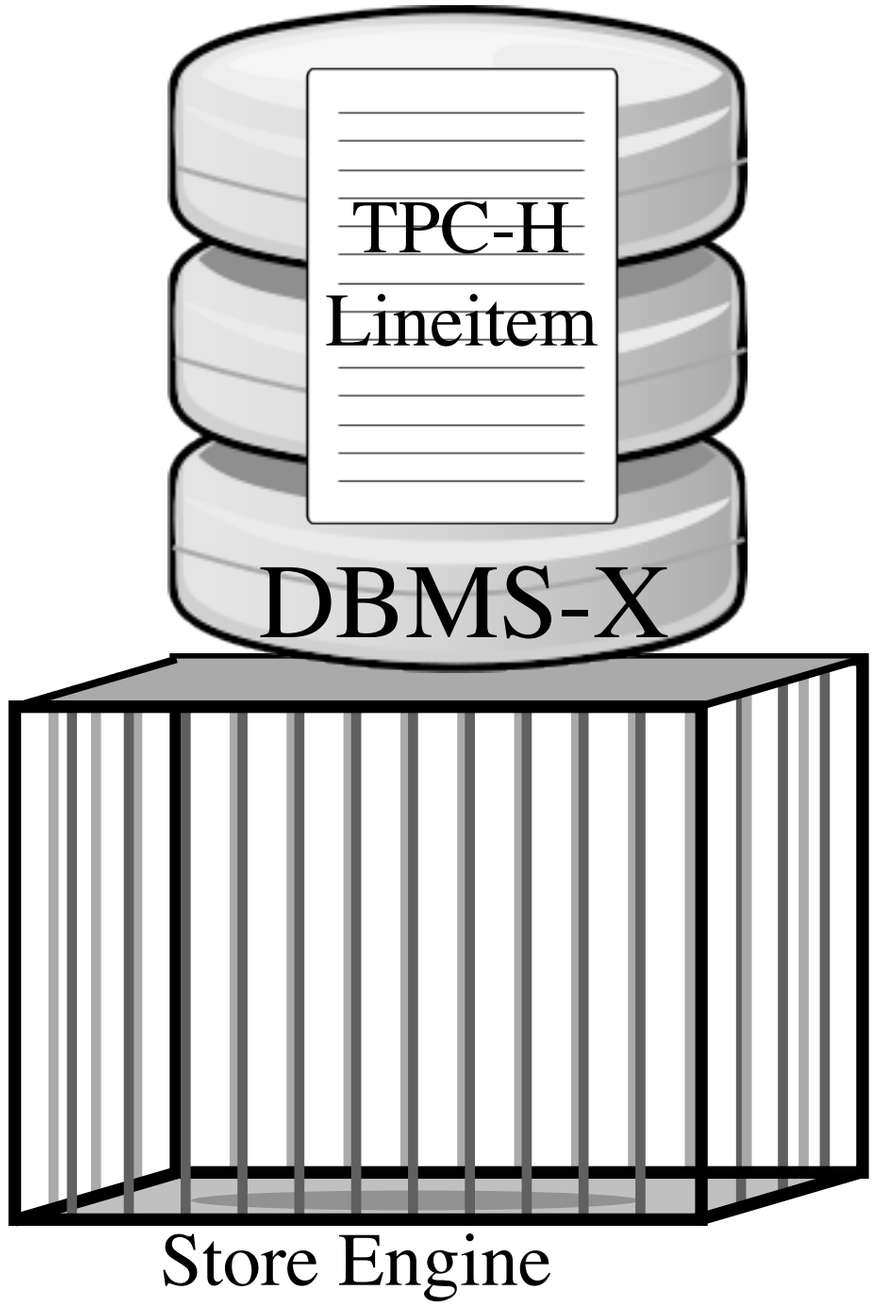


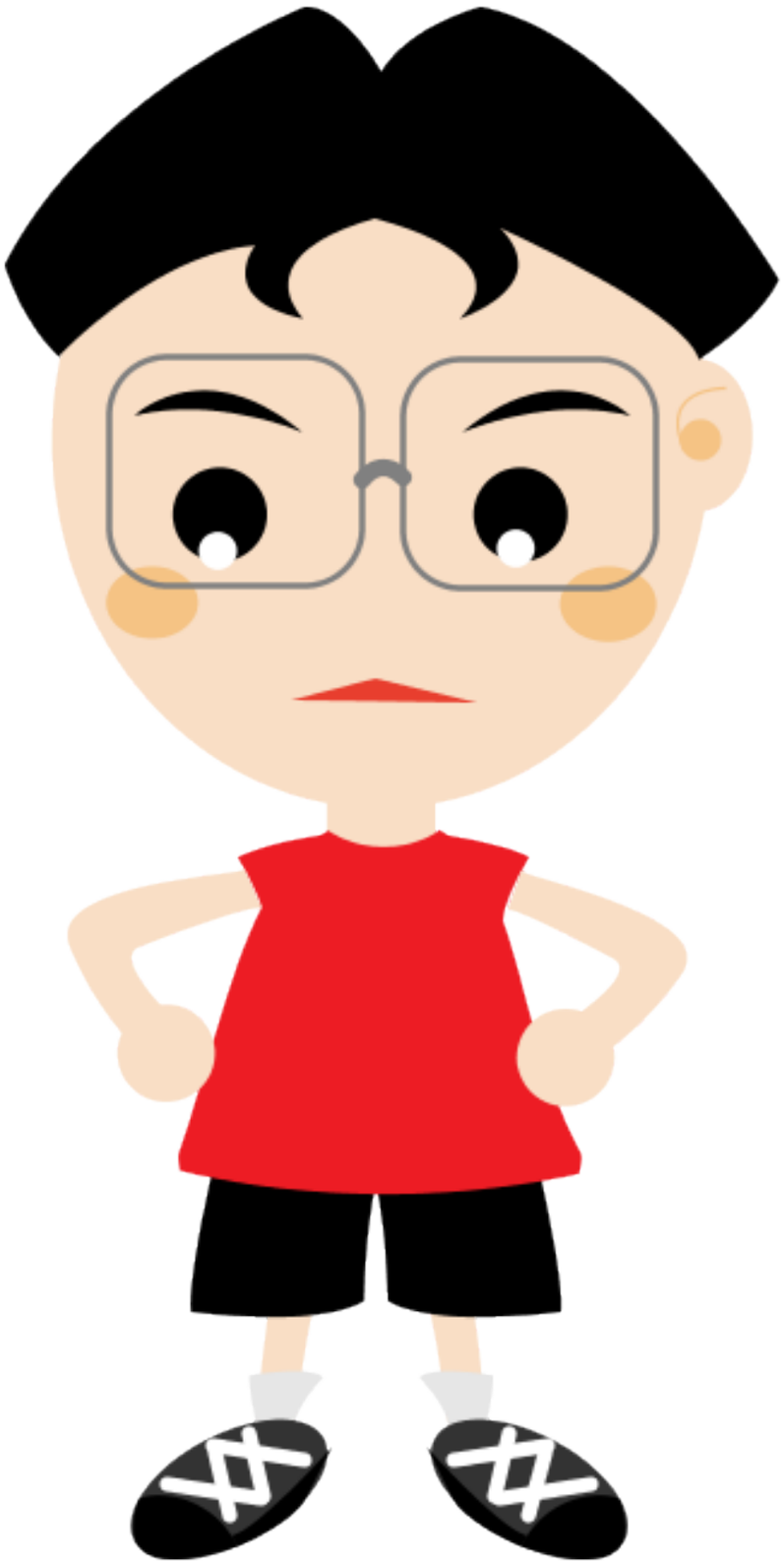

Data
Management
System
TPC-H
Lineitem

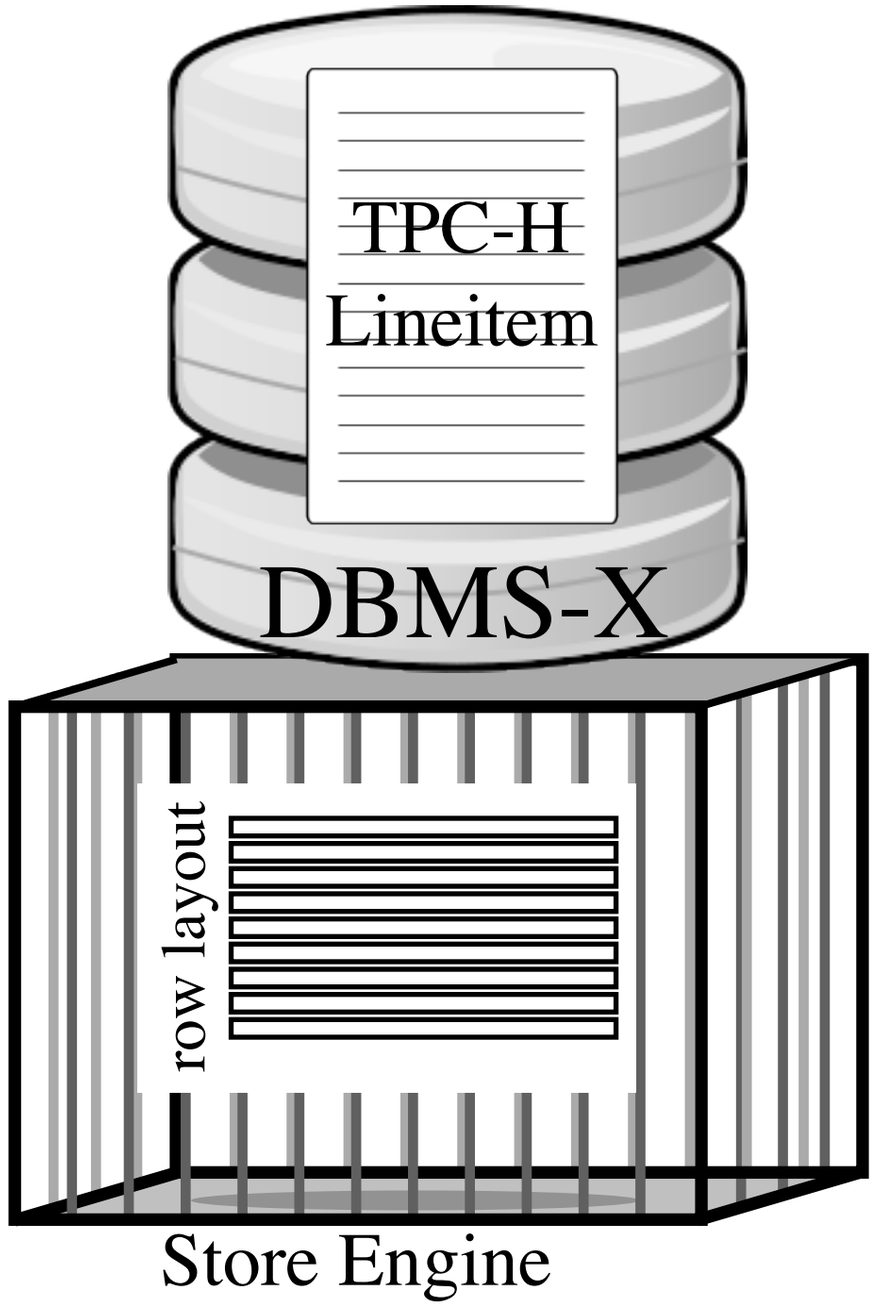



Data
Management
System
TPC-H
Lineitem
Store Engine




TPC-H
Lineitem
DBMS-X
Store Engine

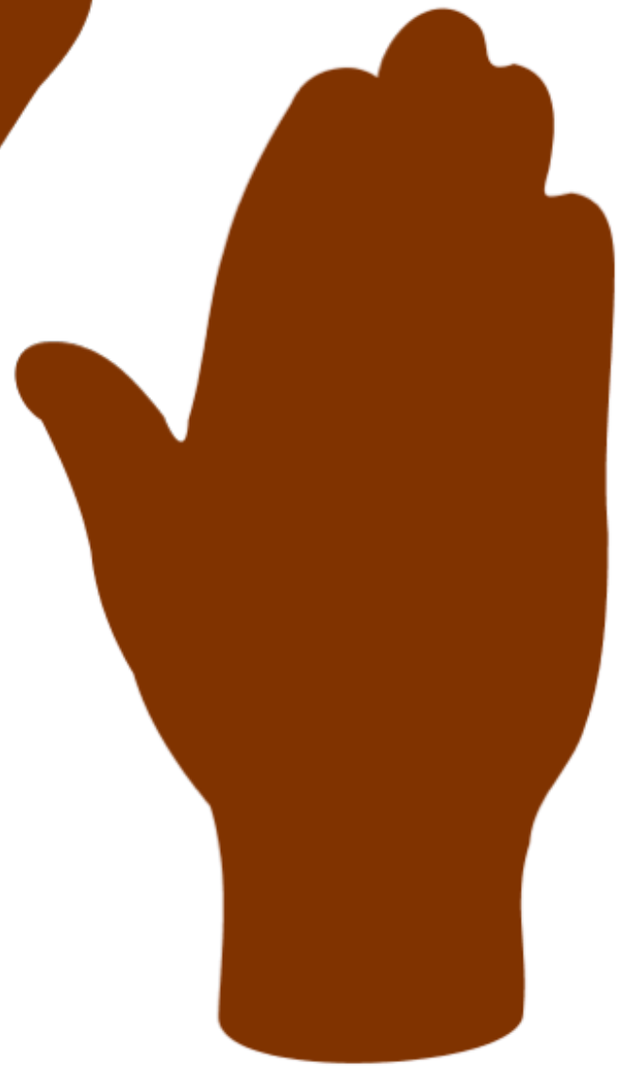


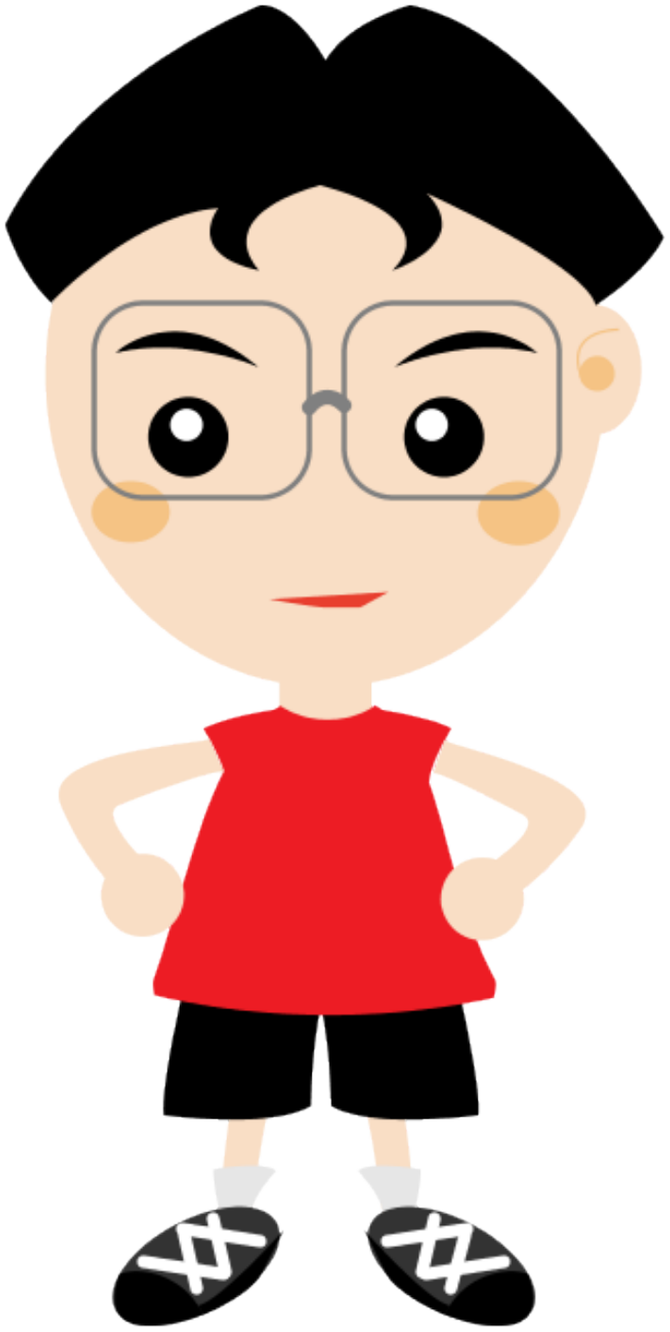

TPC-H
Lineitem
DBMS-X
row layout
Store Engine

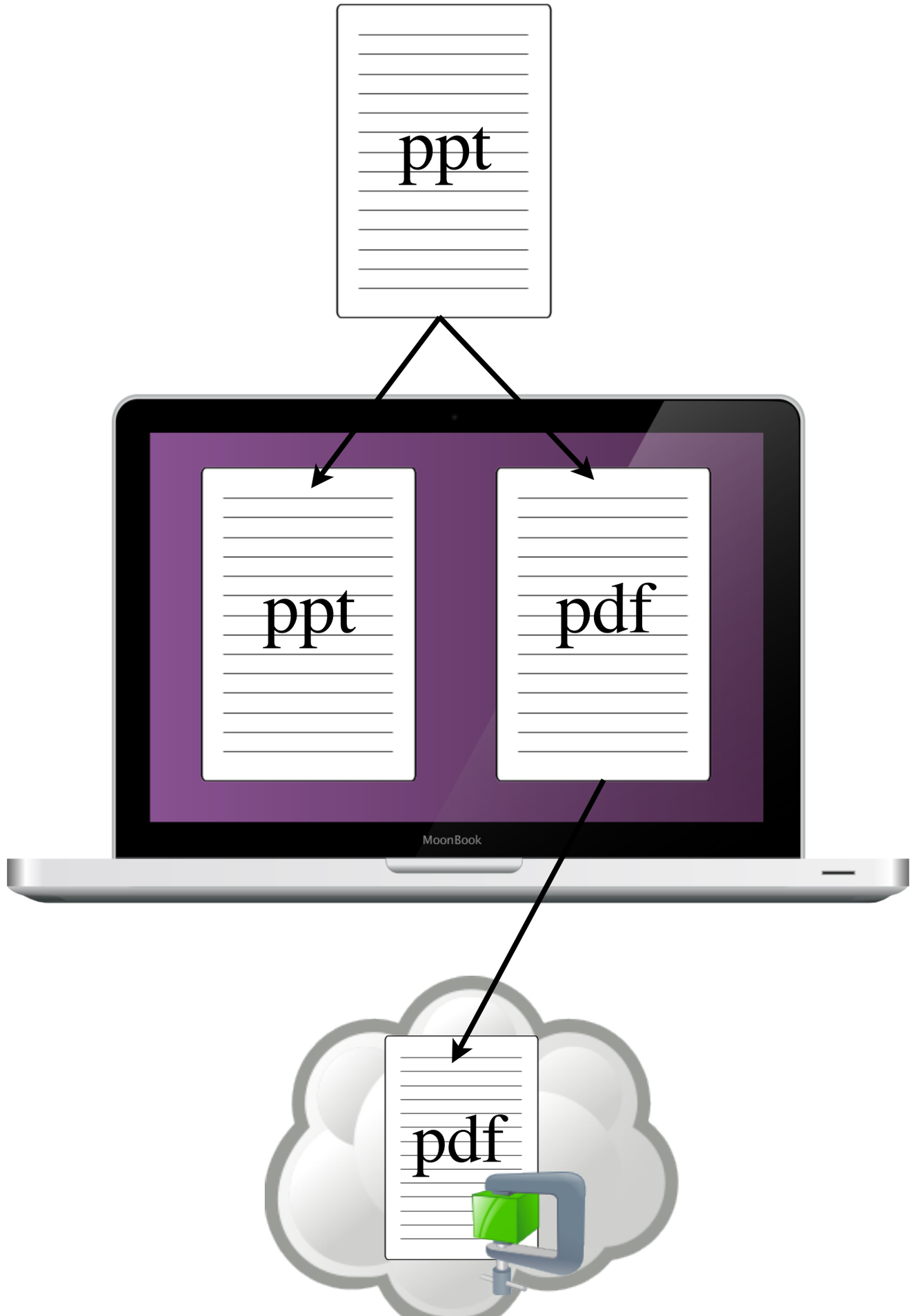
Where

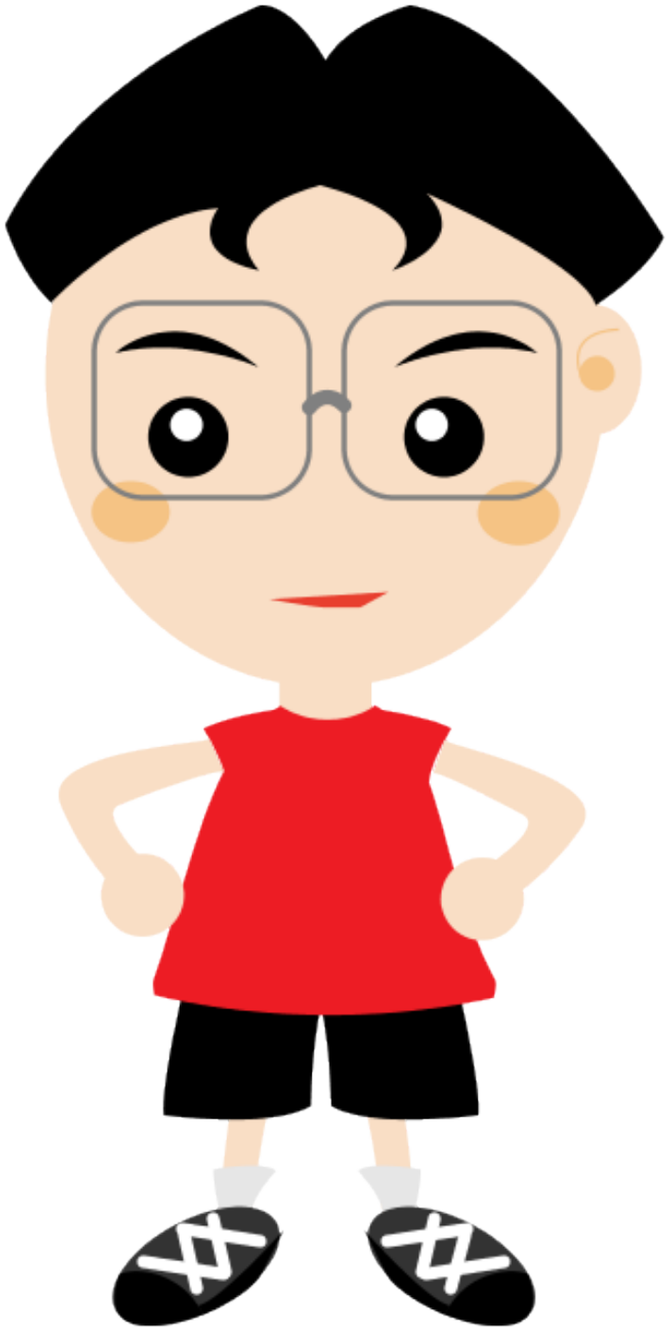
What

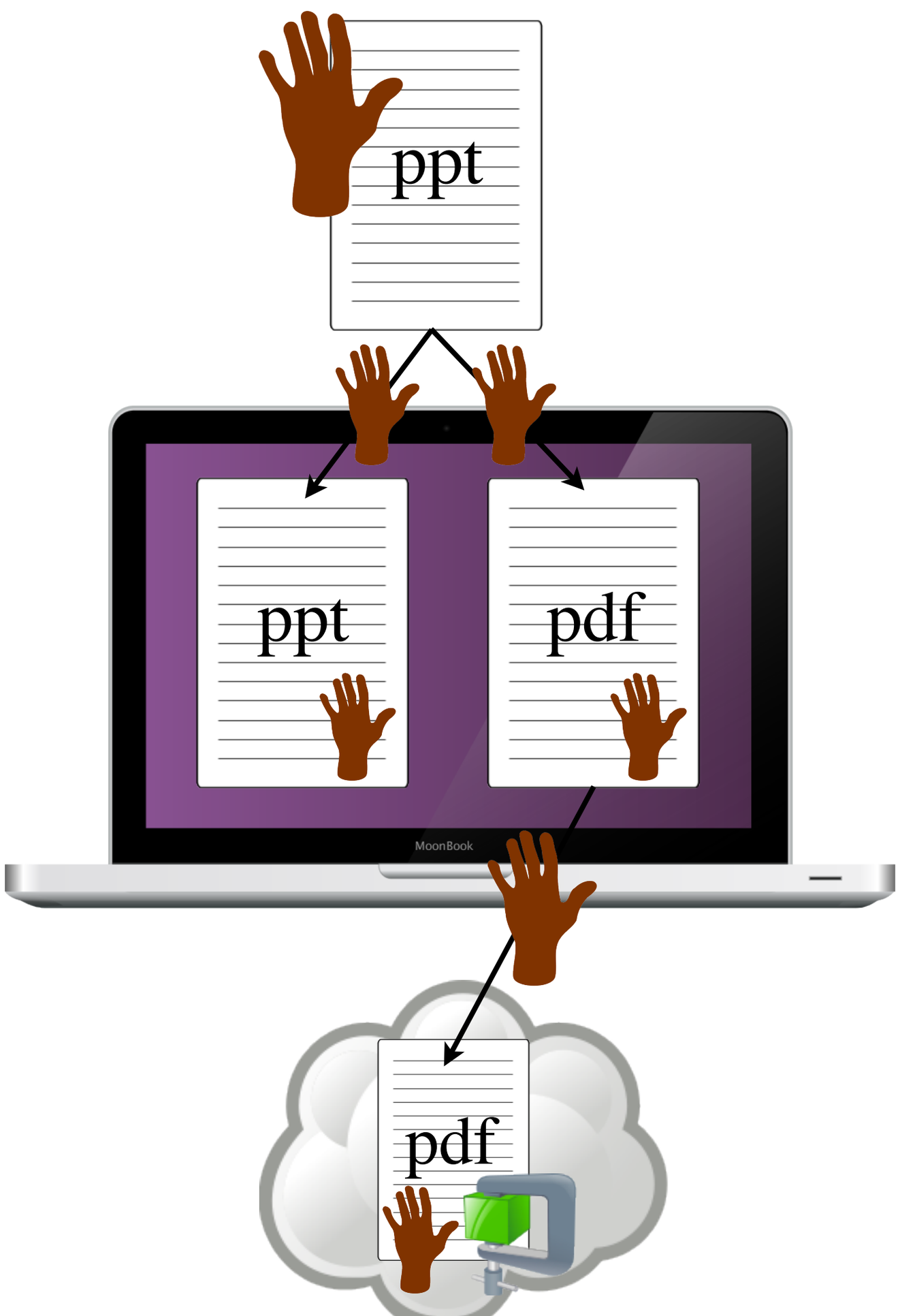
How






ppt
ppt
pdf
MoonBook
pdf




ppt
ppt
pdf
MoonBook
pdf

WWHow!

WWHow!

h h

a e

t r

e

WWHow! Vision

1 Flexible on *What*, *Where*, and *How*

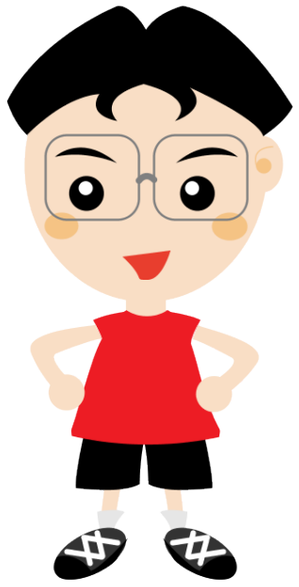
WWHow! Vision

1. **Flexible on *What*, *Where*, and *How***
2. **Declarative Data Storage Language**

WWHow! Vision

1. **Flexible on *What*, *Where*, and *How***
2. **Declarative Data Storage Language**
3. **Data Storage Optimizer**

Logical
Data View



Data
Management
System

DSL

DSL

WWHow! Language

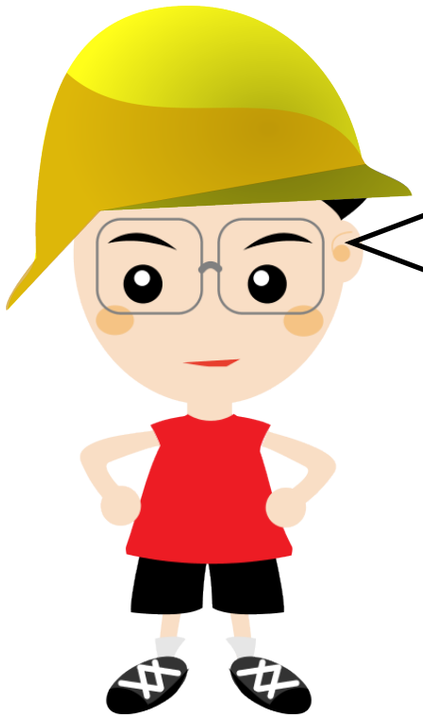
WWHow! Layer

Physical Storage Interface

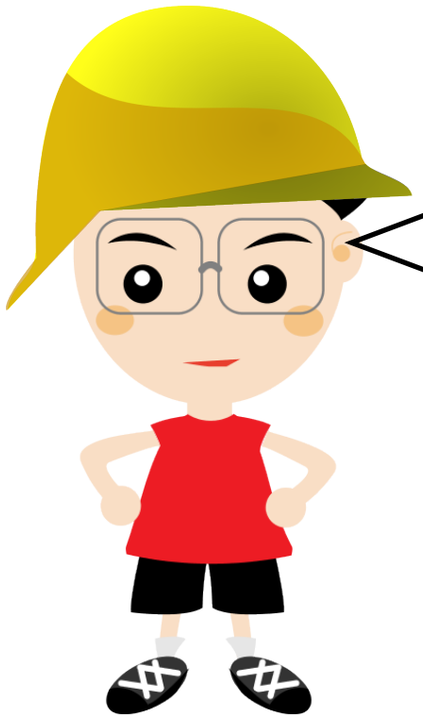
Physical
Data View

WWHow!

hat can we do?

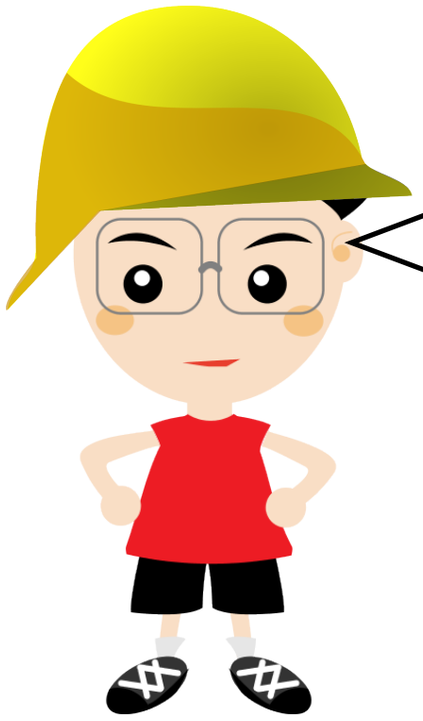


I want to store **3 replicas** of the weblogs
on the cloud with a **different index** each replica



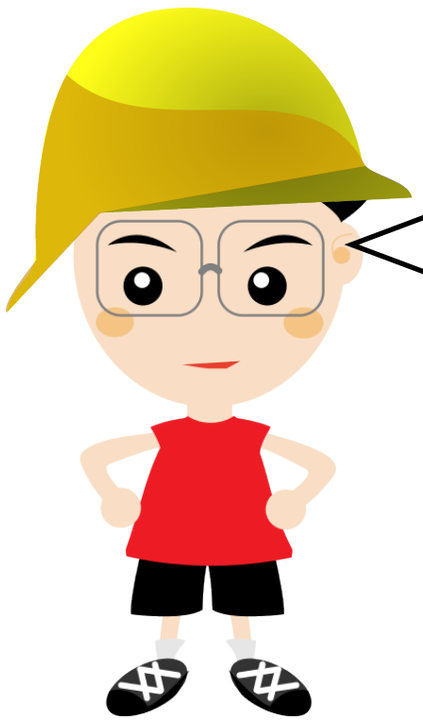
I want to store **3 replicas** of the weblogs
on the cloud with a **different index** each replica

```
STORE ‘/webApp/logs/UserVisits.log’
```



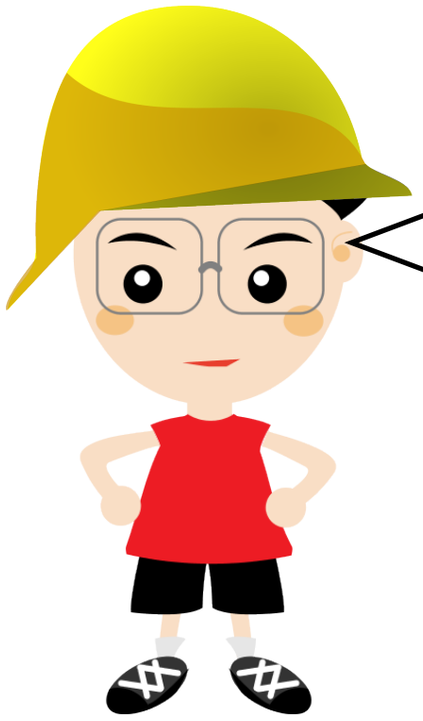
I want to store **3 replicas** of the weblogs
on the cloud with a **different index** each replica

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
```



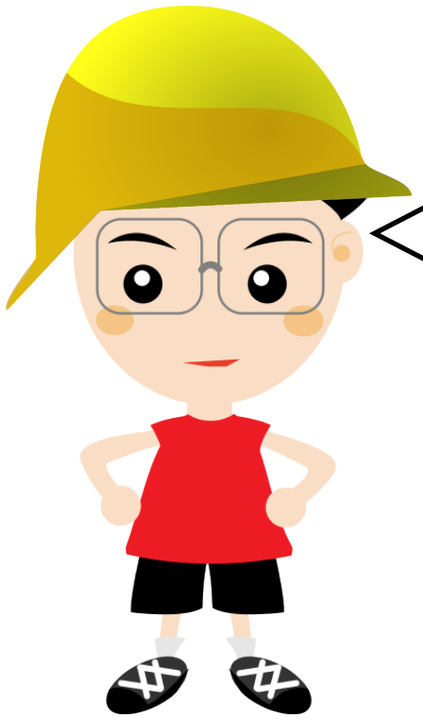
I want to store **3 replicas** of the weblogs
on the cloud with a **different index** each replica

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
WHERE ec2-007-23-compute-1-amazonaws.com
```

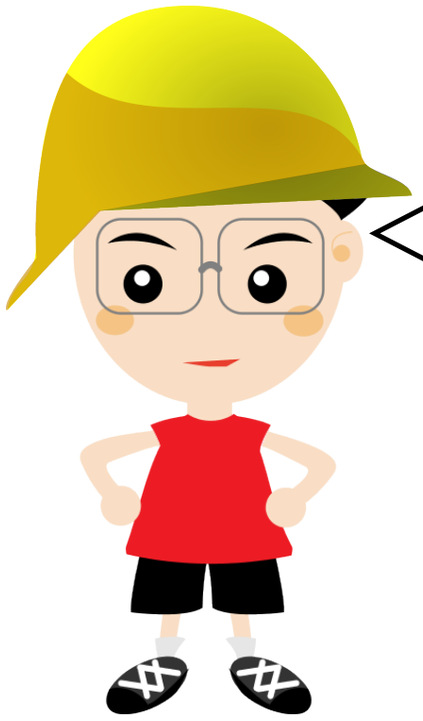


I want to store **3 replicas** of the weblogs
on the cloud with a **different index** each replica

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
WHERE ec2-007-23-compute-1-amazonaws.com
HOW idx(url) FOR rep-1,
    idx(sourceIP) FOR rep-2,
    idx(visitDate) FOR rep-3;
```

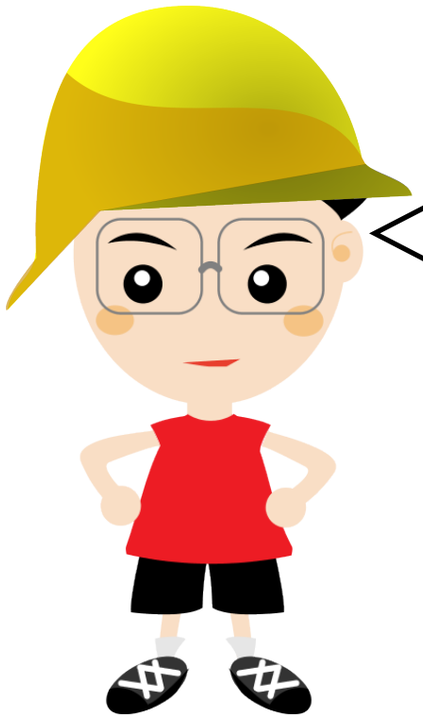


I want to store **3 replicas** of the weblogs with a **different index** each replica and with **high privacy**



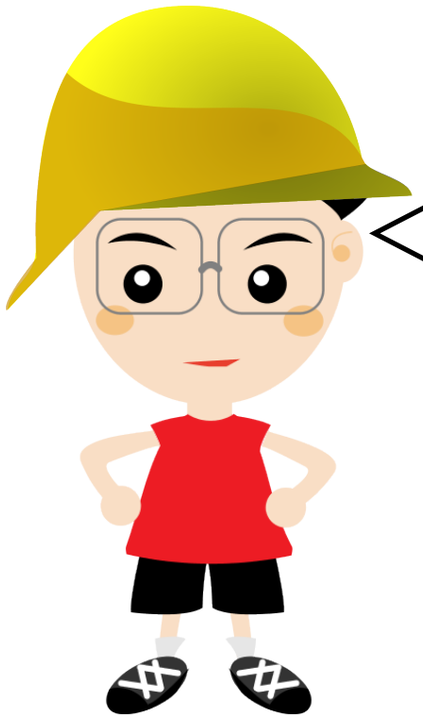
I want to store **3 replicas** of the weblogs with a **different index** each replica and with **high privacy**

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
HOW idx(url) FOR rep-1,
      idx(sourceIP) FOR rep-2,
      idx(visitDate) FOR rep-3;
```



I want to store **3 replicas** of the weblogs with a **different index** each replica and with **high privacy**

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
HOW idx(url) FOR rep-1,
      idx(sourceIP) FOR rep-2,
      idx(visitDate) FOR rep-3;
CONSTRAINT Privacy = ‘high’
```



I want to store **3 replicas** of the weblogs with a **different index** each replica and with **high privacy**

```
STORE ‘/webApp/logs/UserVisits.log’
WHAT * AS rep-1, * AS rep-2, * AS rep-3
HOW idx(url) FOR rep-1,
      idx(sourceIP) FOR rep-2,
      idx(visitDate) FOR rep-3;
CONSTRAINT Privacy = ‘high’
```

job for the
WWhow! data storage optimizer


I want to store the weblogs with **high privacy** and if possible with **high availability**


```
STORE ‘/webApp/logs/UserVisits.log’
```


```
STORE ‘/webApp/logs/UserVisits.log’
PREFERENCE Availability = ‘high’
```


```
STORE ‘/webApp/logs/UserVisits.log’
PREFERENCE Availability = ‘high’
CONSTRAINT Privacy = ‘high’
```


```
STORE ‘/webApp/logs/UserVisits.log’
PREFERENCE Availability = ‘high’
CONSTRAINT Privacy = ‘high’
```

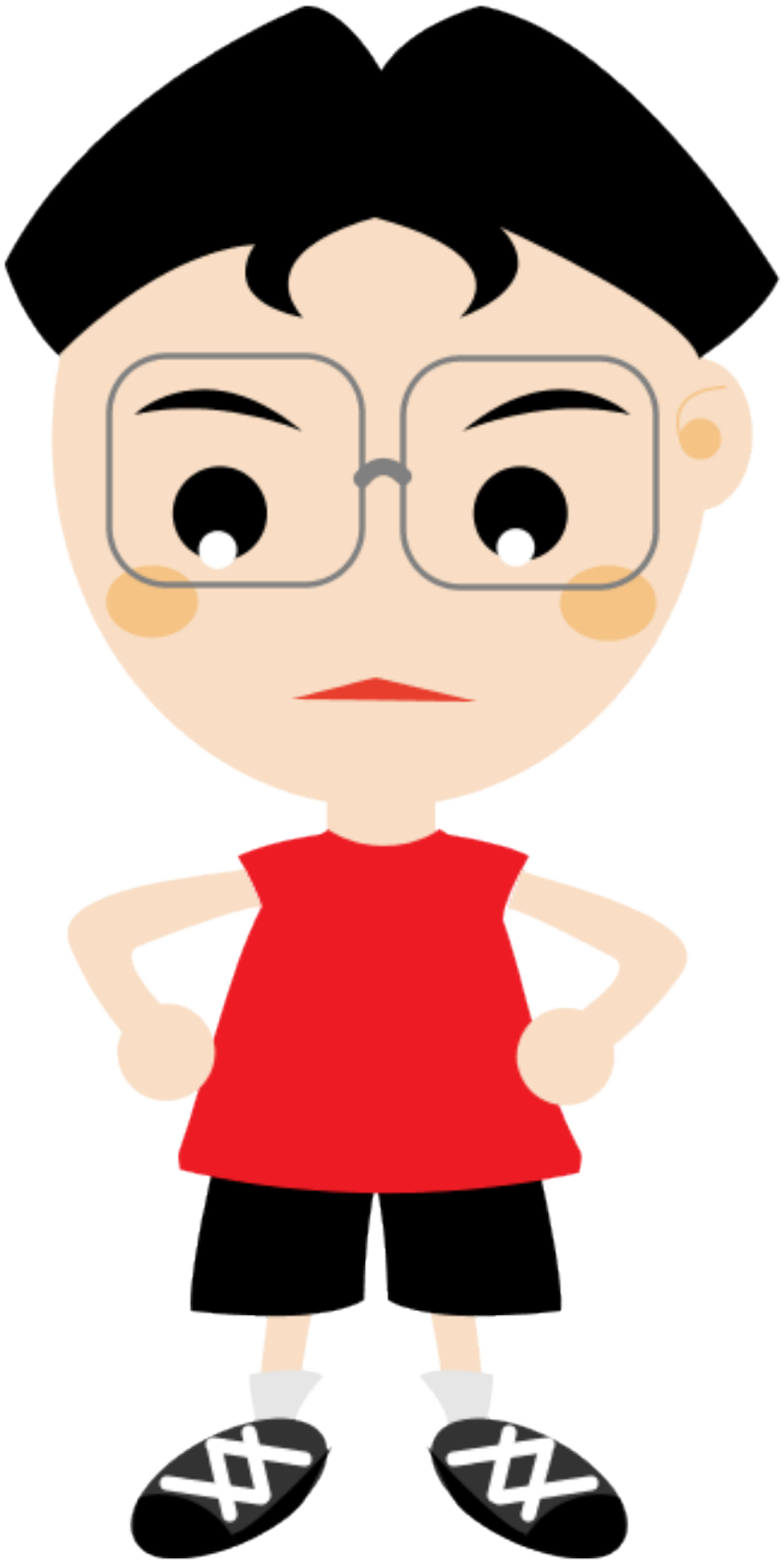
job for the
WWhow! data storage optimizer

Euh!

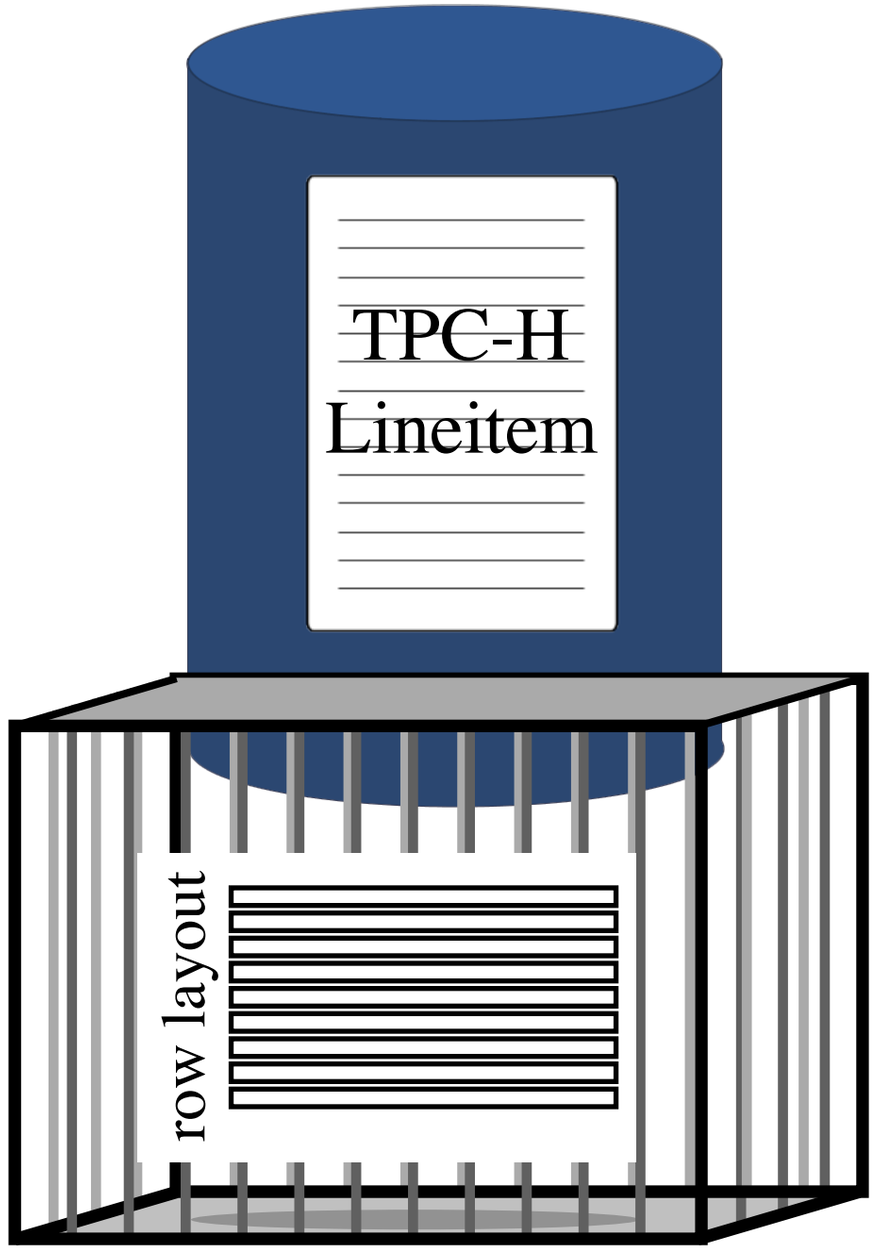
What about existing technology?

WWHow!

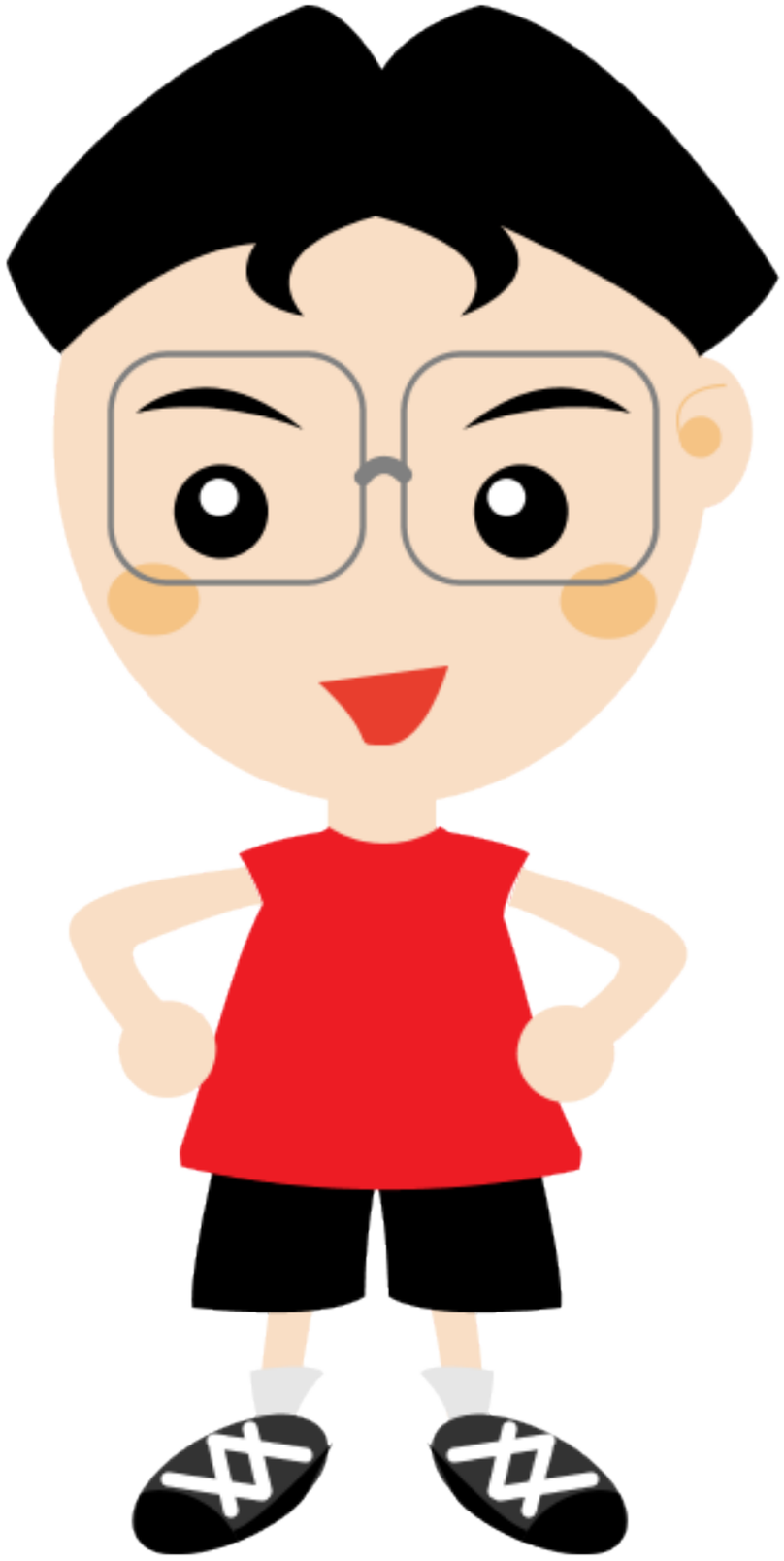
here do we go?



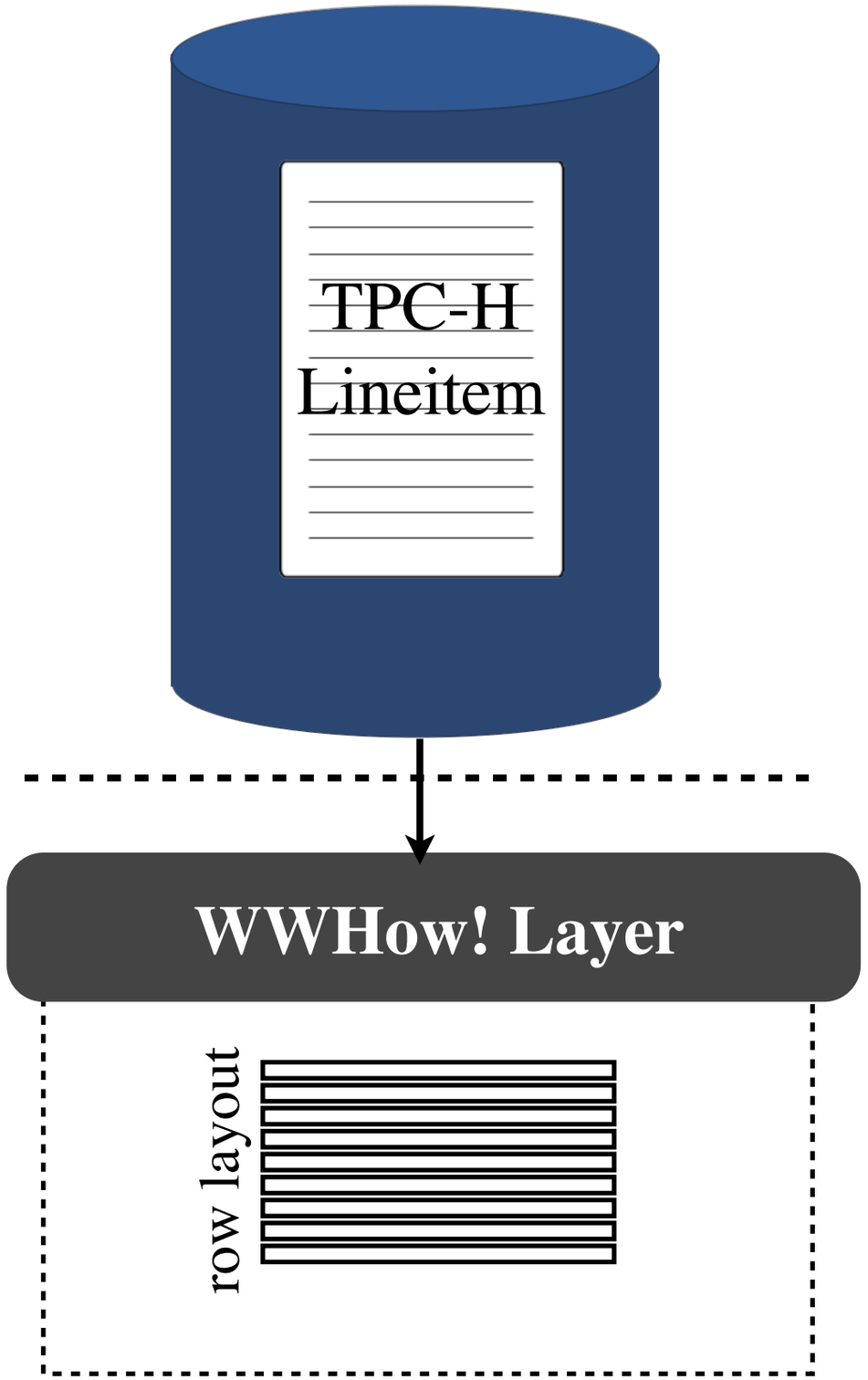
Data Management System

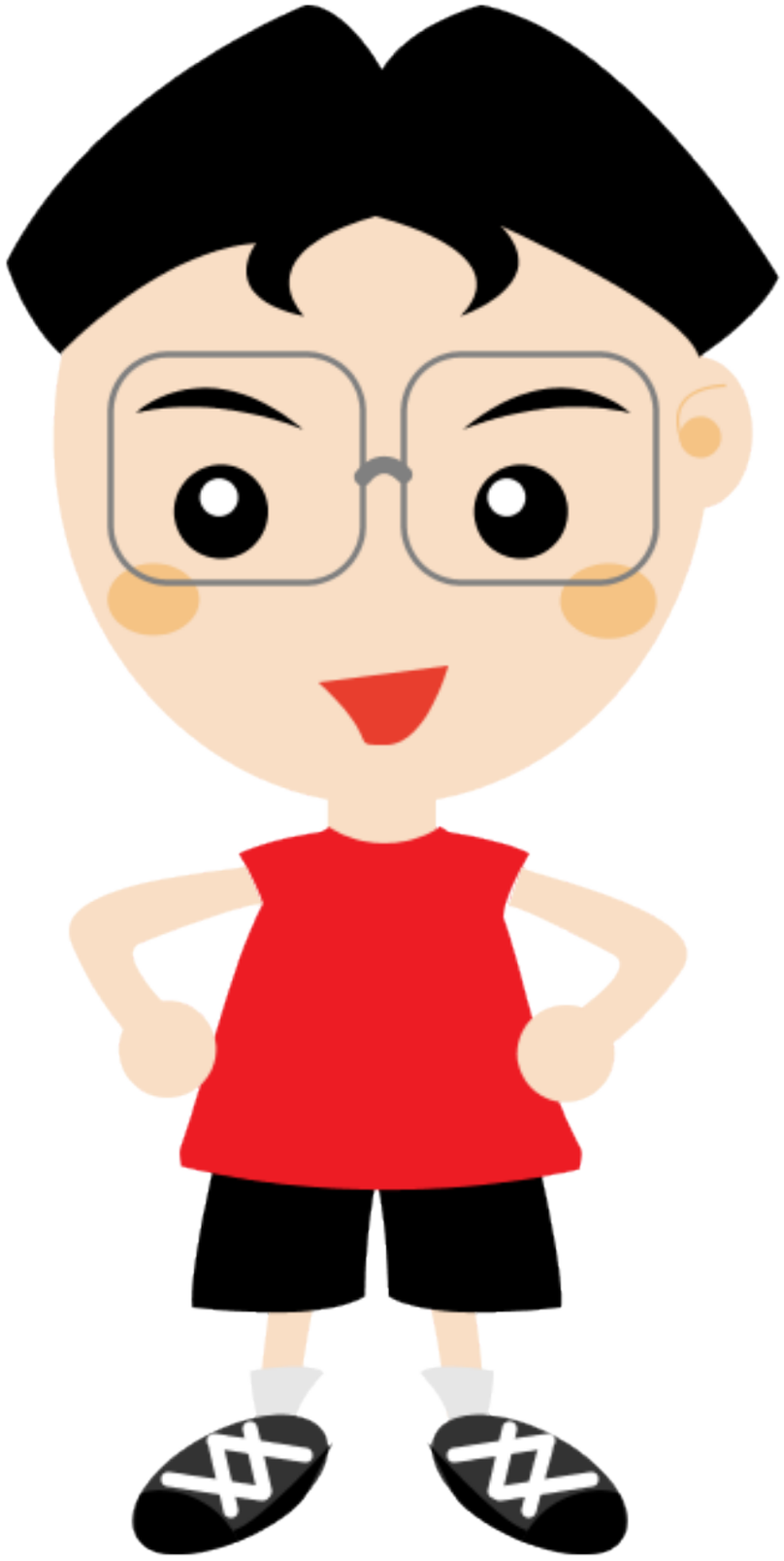


Store Engine

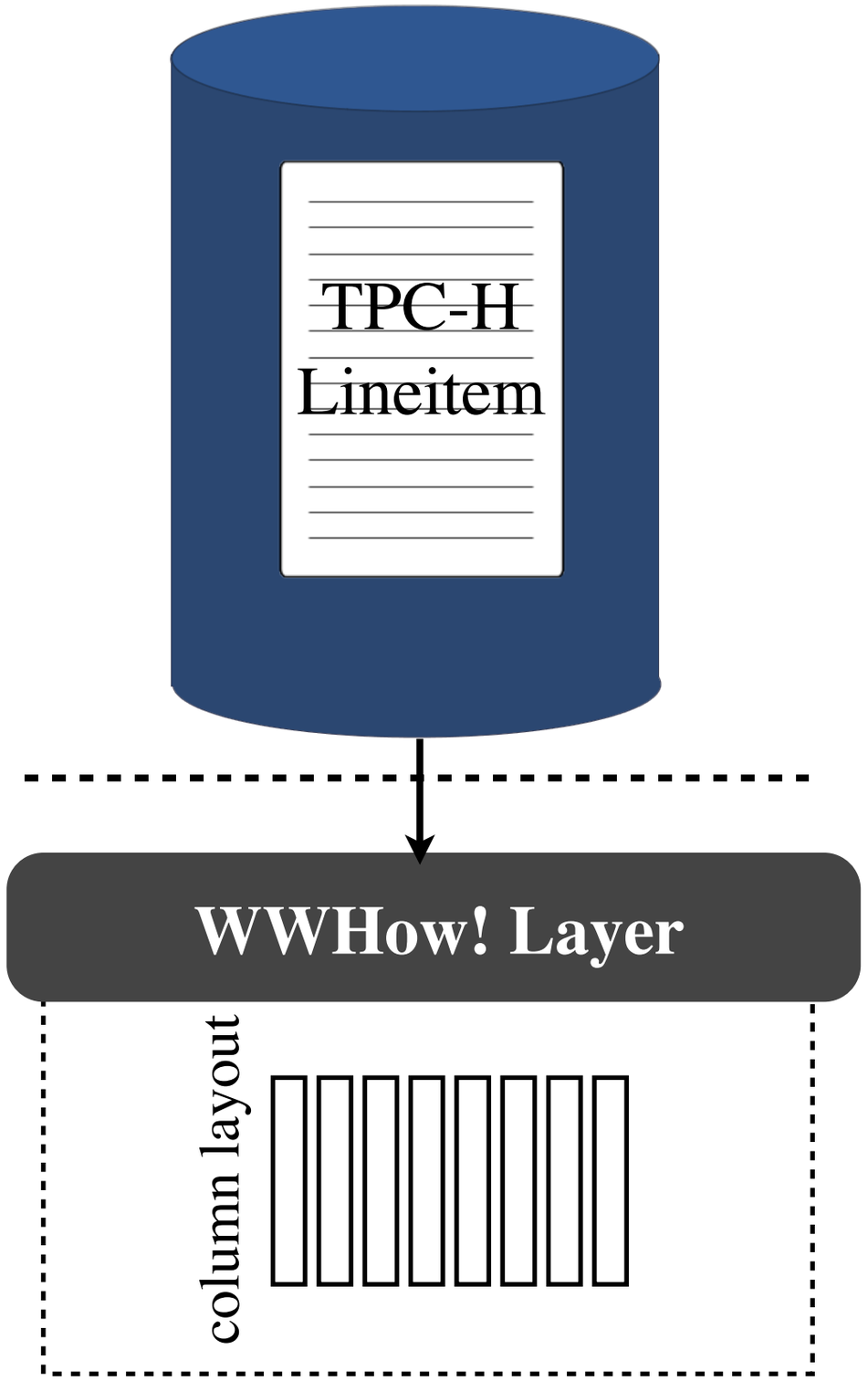


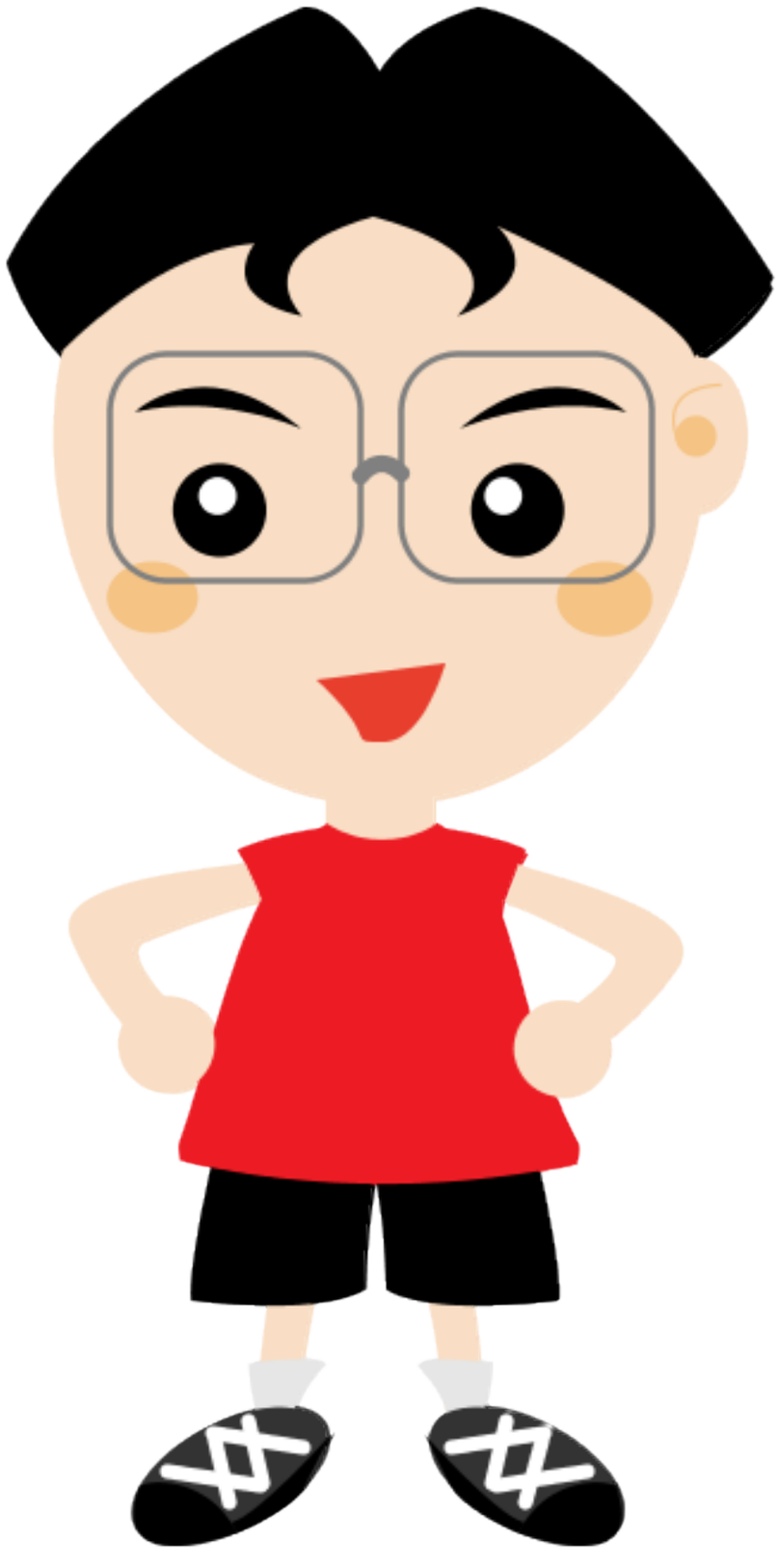
Data Management System



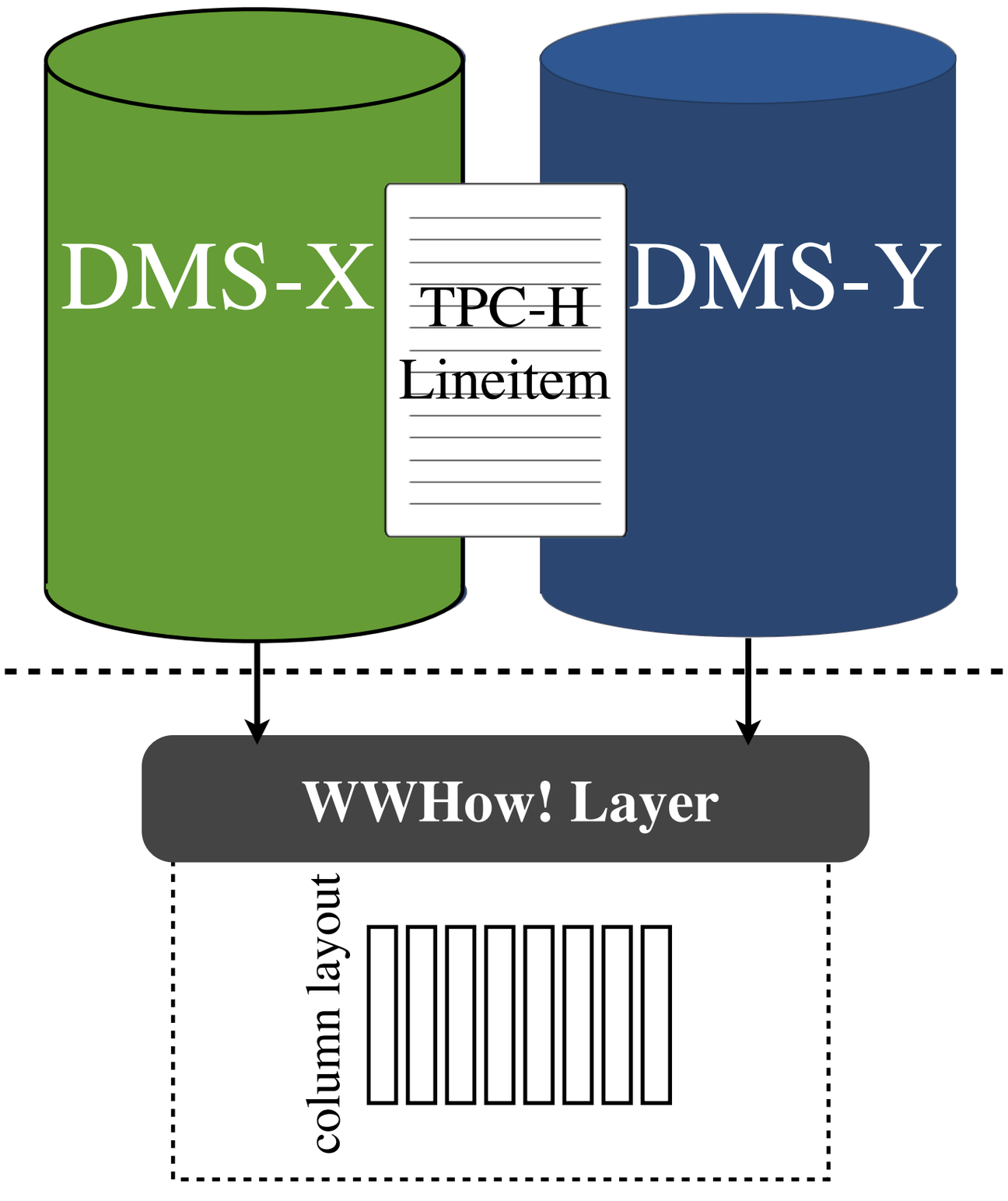


Data Management System

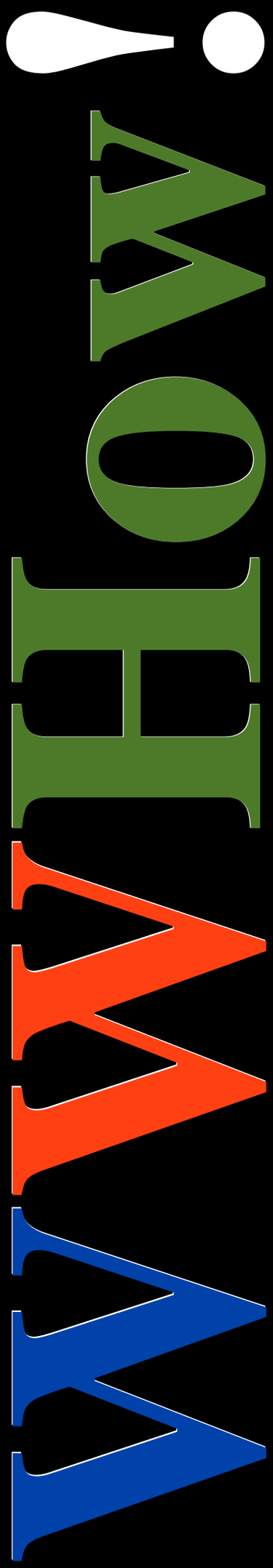




Data Management Systems



How to proceed?



WWHow!

Logical Data View

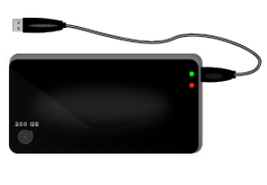
Data Management System

WWHow! Language

WWHow! Layer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

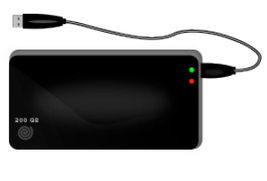
Data Management System

WWHow! Language

WWHow! Layer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

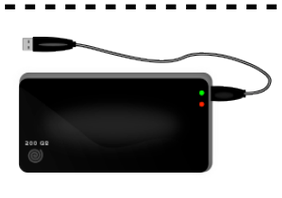
WWHow! Language

API (Logical View)

WWHow! Layer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

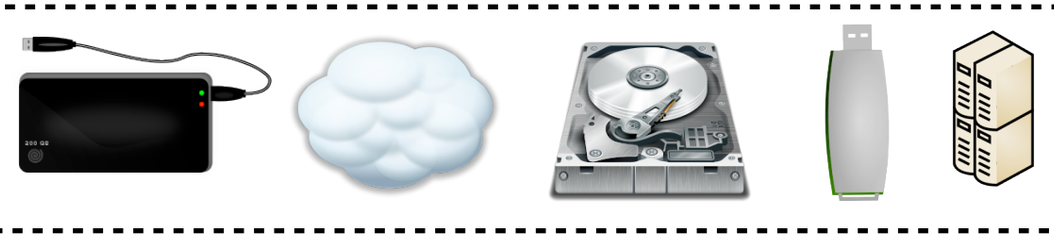
WWHow! Language

API (Logical View)

WWHow! Layer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Language

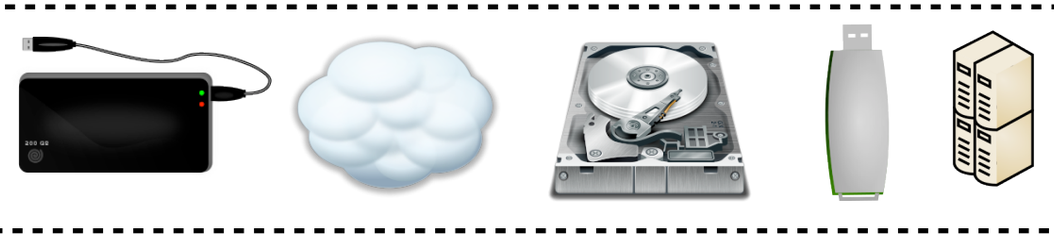
API (Logical View)

WWHow! Language Interpreter

WWHow! Layer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Language

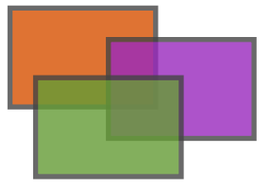
API (Logical View)

WWHow! Layer

WWHow! Language Interpreter

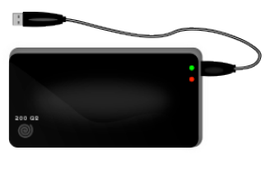
WWHow! Controller

Storage themes



Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Language

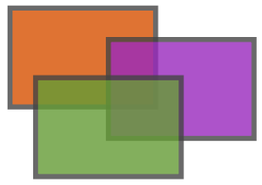
API (Logical View)

WWHow! Layer

WWHow! Language Interpreter

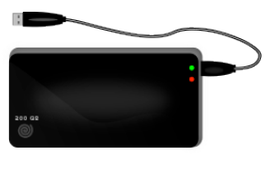
WWHow! Controller

Storage themes



Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Layer

WWHow! Language

API (Logical View)

WWHow! Language Interpreter

WWHow! Controller

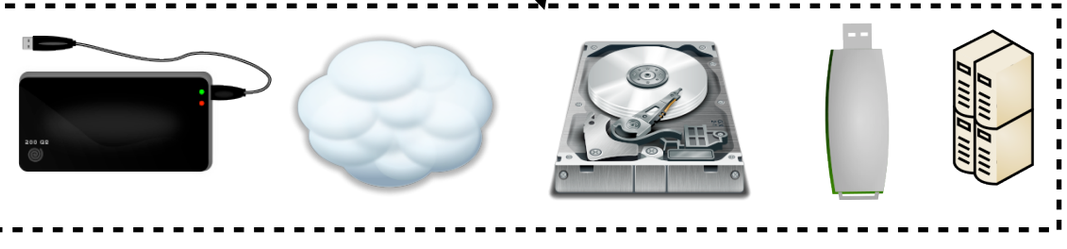
Storage themes

WWHow! Optimizer

Data Storage Optimizer

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Layer

WWHow! Language

API (Logical View)

WWHow! Language Interpreter

WWHow! Controller

Storage themes

WWHow! Optimizer

Data Storage Optimizer

What

Where ⟷ How

Physical Storage Interface

Physical Data View

WWHow!

Logical Data View

Data Management System

WWHow! Layer

WWHow! Language

API (Logical View)

WWHow! Language Interpreter

WWHow! Controller

Storage themes

WWHow! Optimizer

Data Storage Optimizer

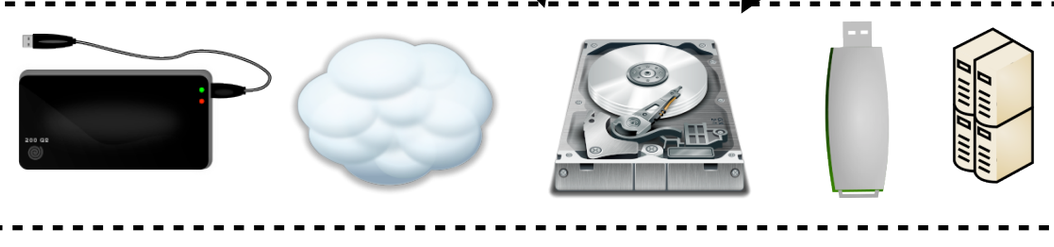
What

Data Access Optimizer

Where ↔ How

Physical Storage Interface

Physical Data View



WWHow!

Logical Data View

Data Management System

WWHow! Layer

WWHow! Language

API (Logical View)

WWHow! Language Interpreter

WWHow! Controller

Storage themes

WWHow! Optimizer

Data Storage Optimizer

What

Data Access Optimizer

Where ↔ How

Physical Storage Interface

Physical Data View