BLOCKCHAIN SYSTEMS & PRIVACY

Aggelos Kiayias

University of Edinburgh & IOHK

project PANORAMIX

ABOUT ME

- Chair in Cyber Security & Privacy at U. of Edinburgh.
- Coordinator of H2020 Panoramix Consortium.
- Director of the Blockchain Technology Laboratory @ UEDIN.
  - conducting research on blockchain systems.
- Chief Scientist of IOHK, a blockchain tech R&D company.
  - we are developing scalable blockchain systems based on state of the art security engineering principles.

https://iohk.io

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

GDPR & RIGHTS OF DATA SUBJECTS

- Right of access
- Right of rectification
- Right to basic information
- Right to erasure
- Right to object/restrict processing
- Right of data portability
- …

RIGHT OF ACCESS

- Article 15. GDPR:
  - The data subject has the right of access .. to the following information

a) the purposes of the processing
b) the categories of personal data concerned
c) the recipients …the personal data have been … disclosed
…
h) the existence of automated decision-making… meaningful information about the logic involved

RIGHT TO ERASURE

- Article 17. GDPR:
  - data subject shall have the right to obtain … the erasure of personal data.

MOTIVATION

MOTIVATION

- Many recent privacy related discussions about blockchain systems deal with the privacy implications of using a particular blockchain application (namely cryptocurrencies such as bitcoin).

MOTIVATION

- Many recent privacy related discussions about blockchain systems deal with the privacy implications of using a particular blockchain application (namely cryptocurrencies such as bitcoin).

- Our main goal : using DLT and additional cryptographic techniques in a constructive fashion to rethink & improve GDPR compliance.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

Understanding DLT

Understanding DLT

- DLT has blockchain protocols as a primary reference point.

Understanding DLT

- DLT has blockchain protocols as a primary reference point.
- The blockchain is a distributed database that satisfies a unique set of safety and liveness properties.

Understanding DLT

- DLT has blockchain protocols as a primary reference point.

- The blockchain is a distributed database that satisfies a unique set of safety and liveness properties.

- To understand it, we can focus to its first (and so far most successful) application.

Case study : Money

- What is money?

Properties of Money

- a medium of exchange
- a unit of account
- a store of value

Properties of Money

- a medium of exchange

can be used as medium
for the exchange
of goods - no barter

- a unit of account

- a store of value

Properties of Money

- a medium of exchange

can be used as medium for the exchange of goods - no barter

- a unit of account

can be used for pricing of all goods and services, for accounting purposes and debt recording.

- a store of value

Properties of Money

- a medium of exchange

can be used as medium for the exchange of goods - no barter

- a unit of account

can be used for pricing of all goods and services, for accounting purposes and debt recording.

- a store of value

storing and retrieving it at a point in the future maintains its value.

Creating Money

Money 1.0 : using a trusted object

Analysis of Money 1.0

- a medium of exchange
- a unit of account
- a store of value

Analysis of Money 1.0

mediocre
[ok for face to face transactions]

- a medium of exchange
- a unit of account
- a store of value

Analysis of Money 1.0

- a medium of exchange
- a unit of account
- a store of value

mediocre
[ok for face to face transactions]

mediocre fungible, but not divisible well. typically forgeable.

Analysis of Money 1.0

- a medium of exchange
- a unit of account
- a store of value

mediocre
[ok for face to face transactions]

mediocre fungible, but not divisible well. typically forgeable.

bad. some objects may deteriorate, others may have unknown hidden quantities.

Creating Money

Money 2.0 : using a trusted entity

Trusted entity issues “IOU”s

Analysis of Money 2.0

- a medium of exchange
- a unit of account
- a store of value

Analysis of Money 2.0

- a medium of exchange
- a unit of account
- a store of value

good
[for transactions within the domain of the trusted entity]

Analysis of Money 2.0

- a medium of exchange
- a unit of account
- a store of value

good
[for transactions within the domain of the trusted entity]

great!
fungible & divisible.

Analysis of Money 2.0

- a medium of exchange
- a unit of account
- a store of value

good
[for transactions within the domain of the trusted entity]

great!
fungible & divisible.

mediocre
[tied to the availability & reputation of the issuing entity]

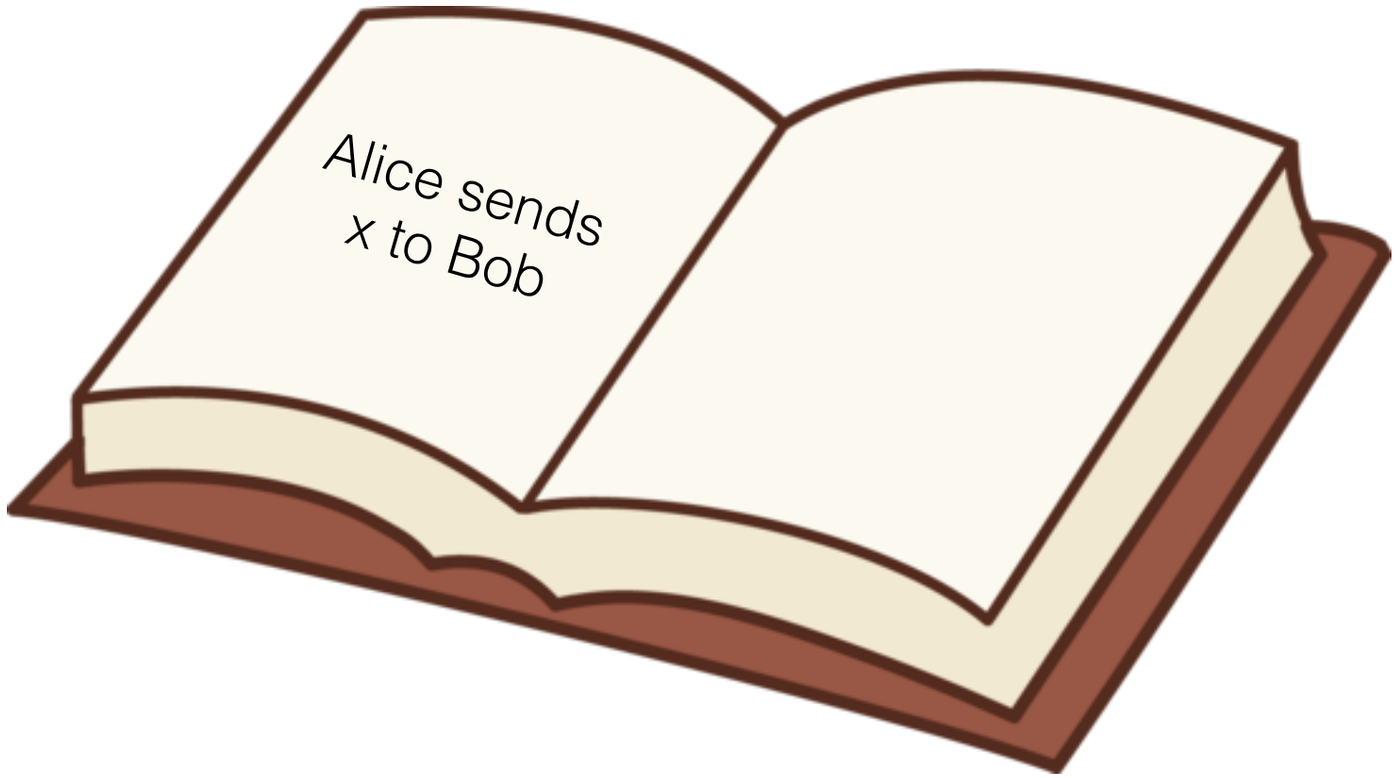
Creating Money

Money 3.0 : Bitcoin

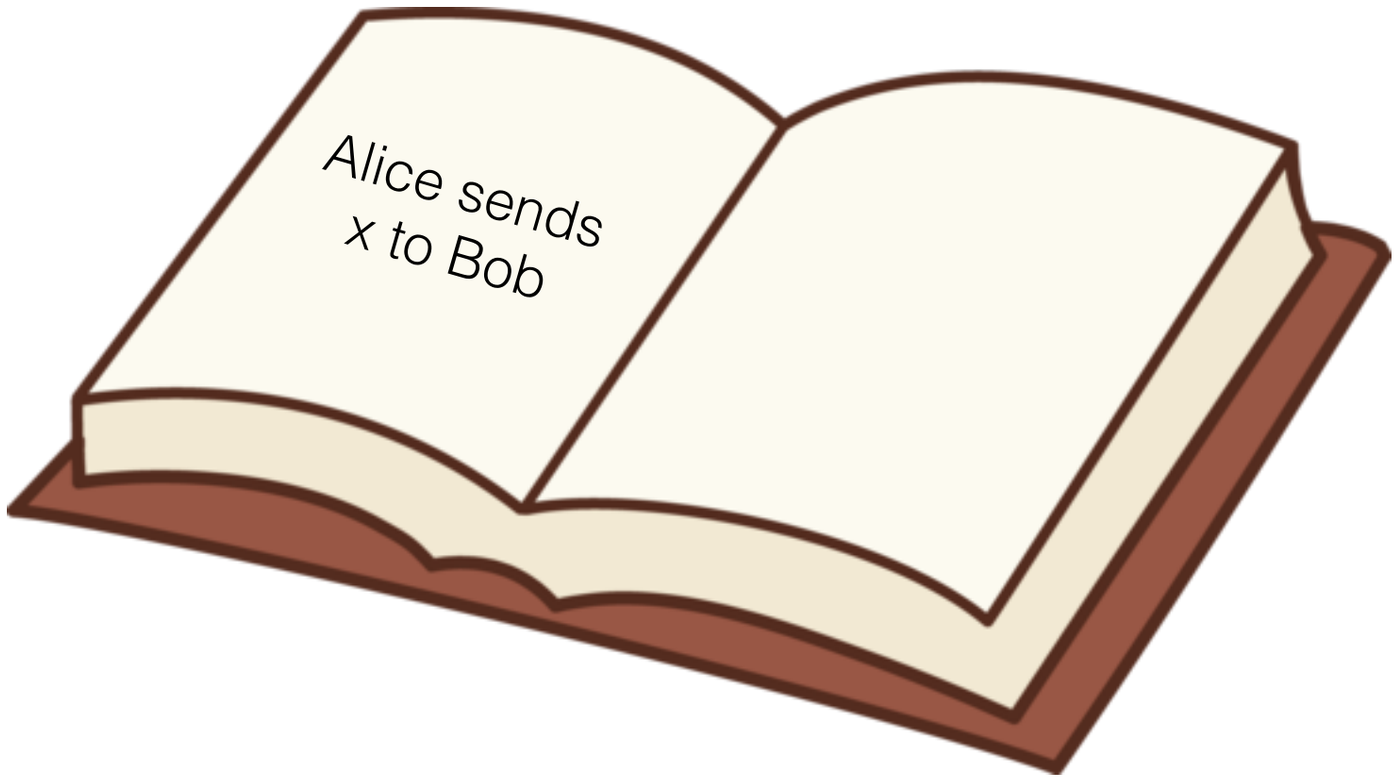
Enter Blockchain & distributed Ledgers

The never-ending book parable

A “book” of transactions

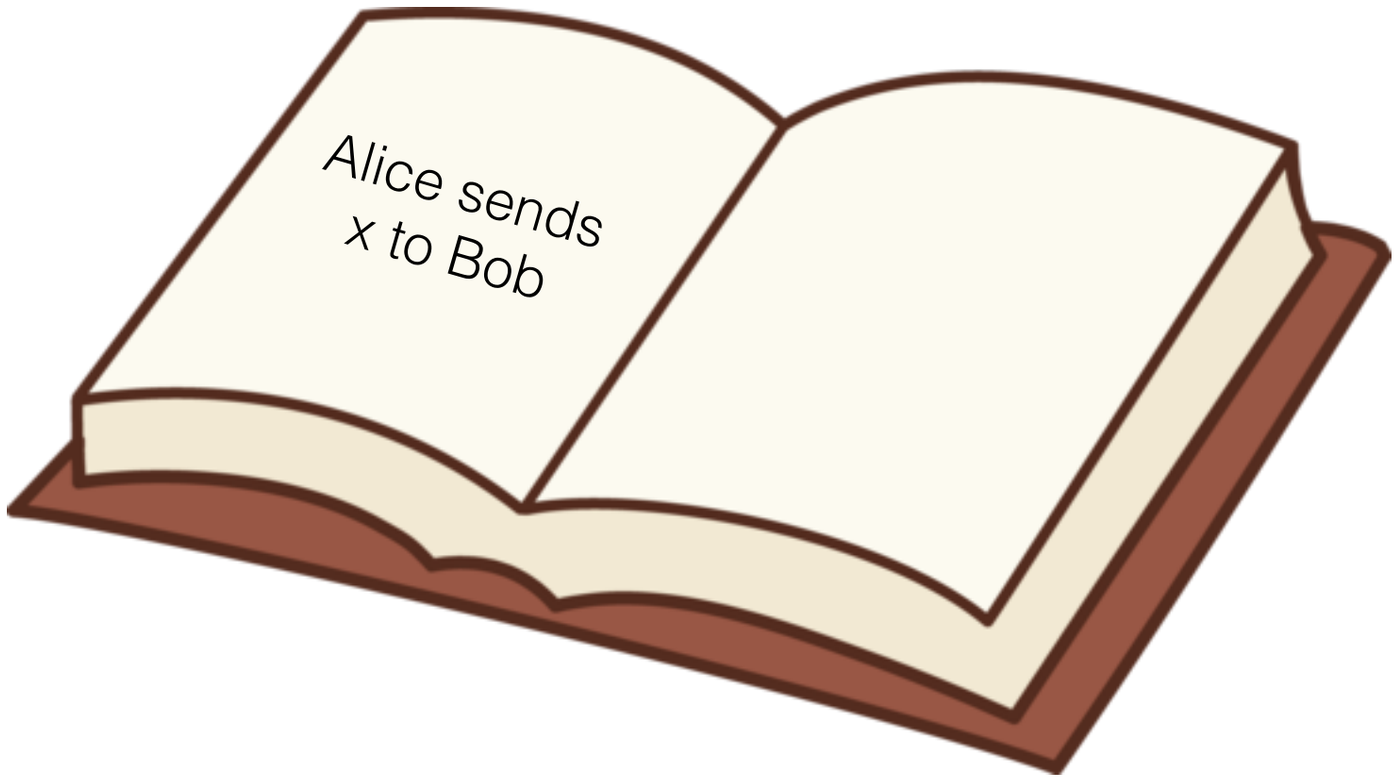


A “book” of transactions



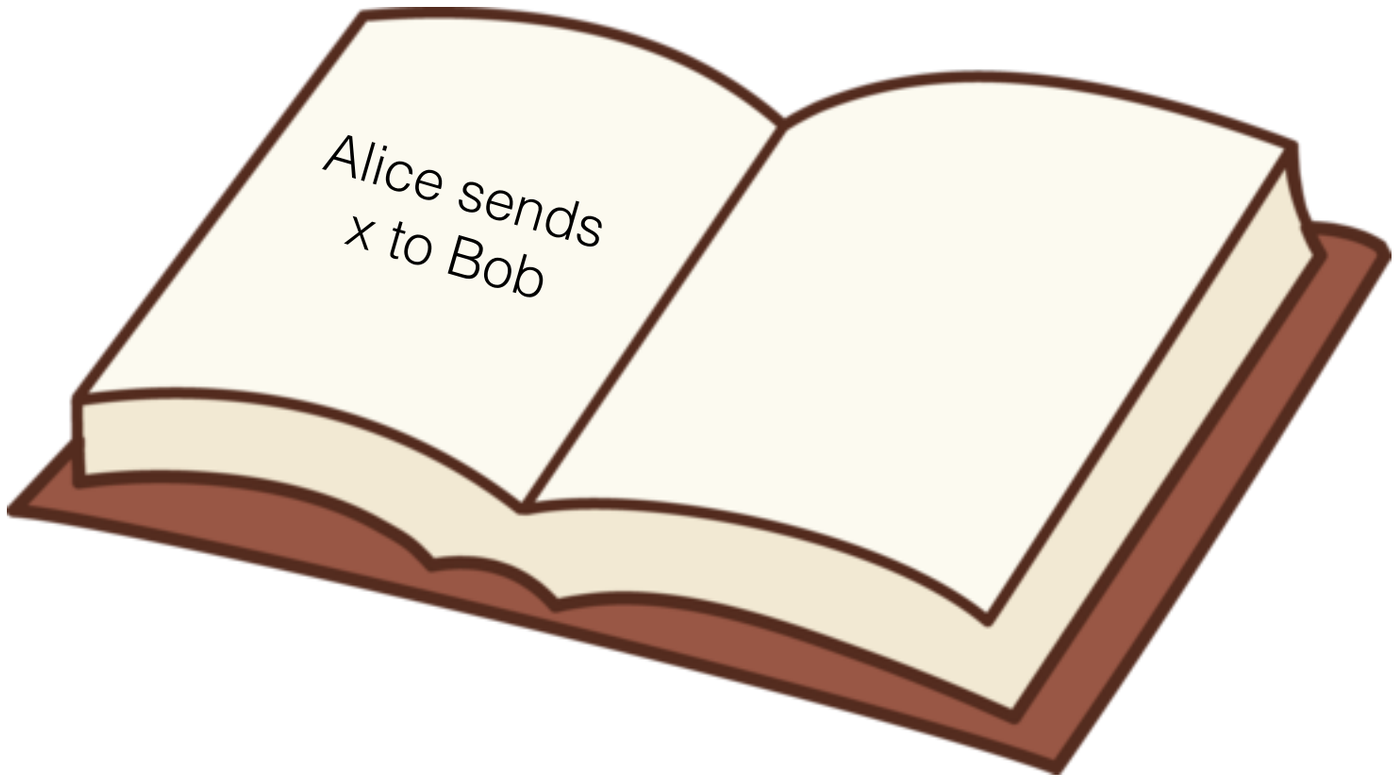
- Each new page requires some effort to produce.

A “book” of transactions



- Each new page requires some effort to produce.
- Anyone can be a scribe and produce a page.

A “book” of transactions

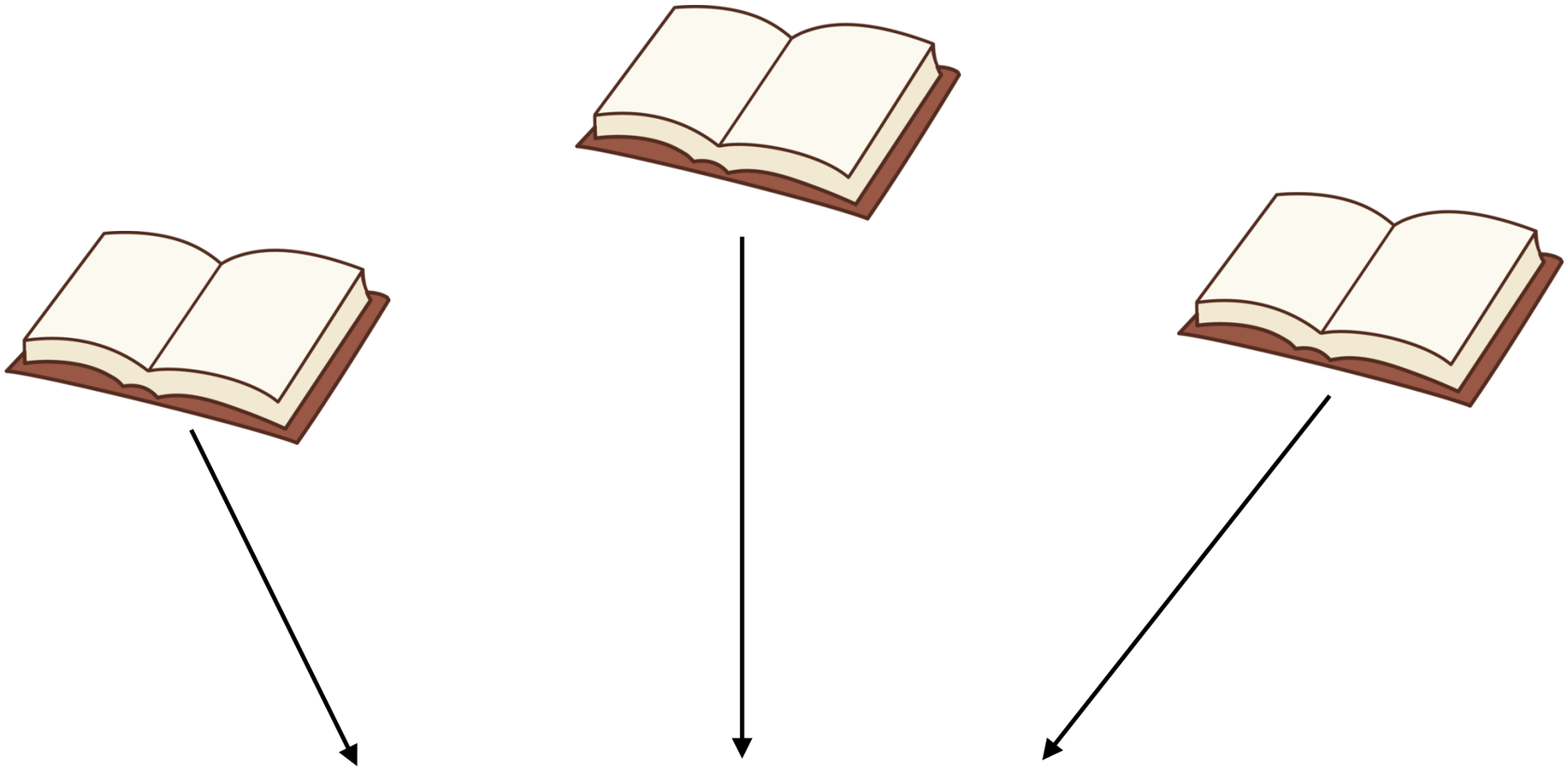


- Each new page requires some effort to produce.
- Anyone can be a scribe and produce a page.
- New pages are produced indefinitely as long as scribes are interested in doing so.

Importance of Consensus

- If multiple conflicting books exist, which is the “right one”?

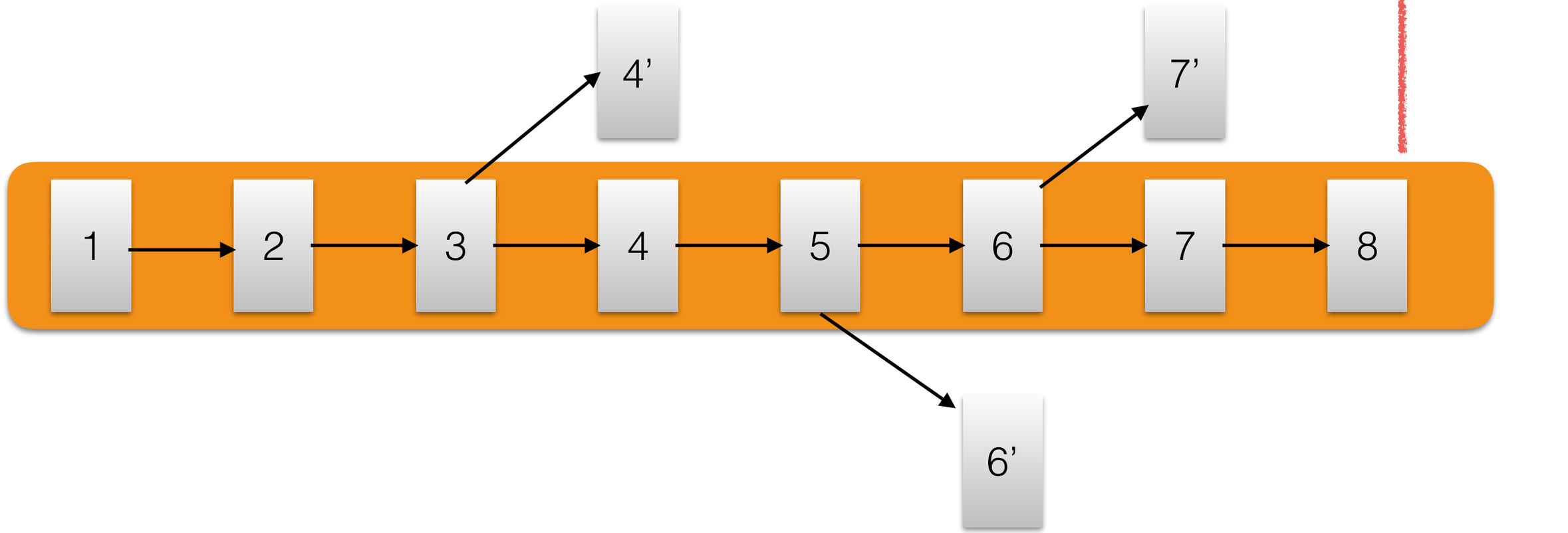
Choosing the correct book



The **current book** to work on & refer to is the book with the most pages. if multiple exist, just pick one at random.

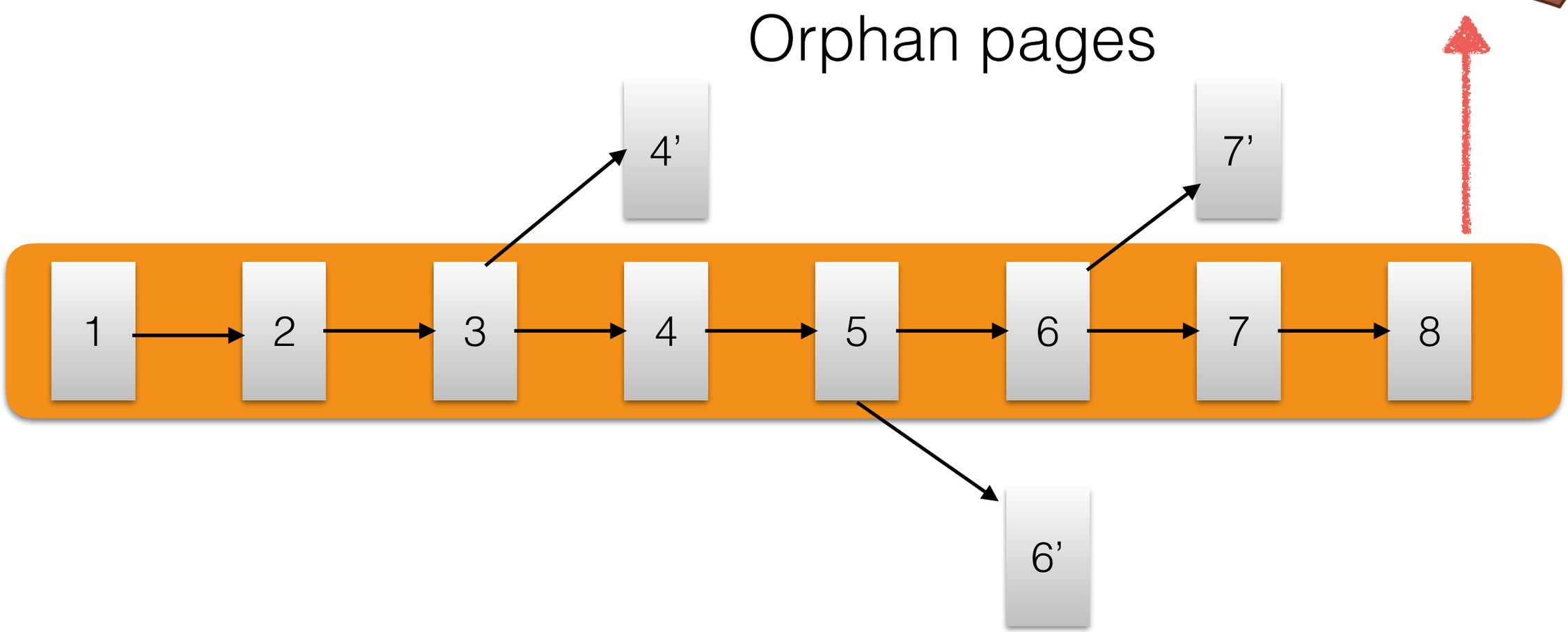
Assembling the current book

- each page refers only to the previous page
- current is assembled by stringing together the longest sequence of pages.



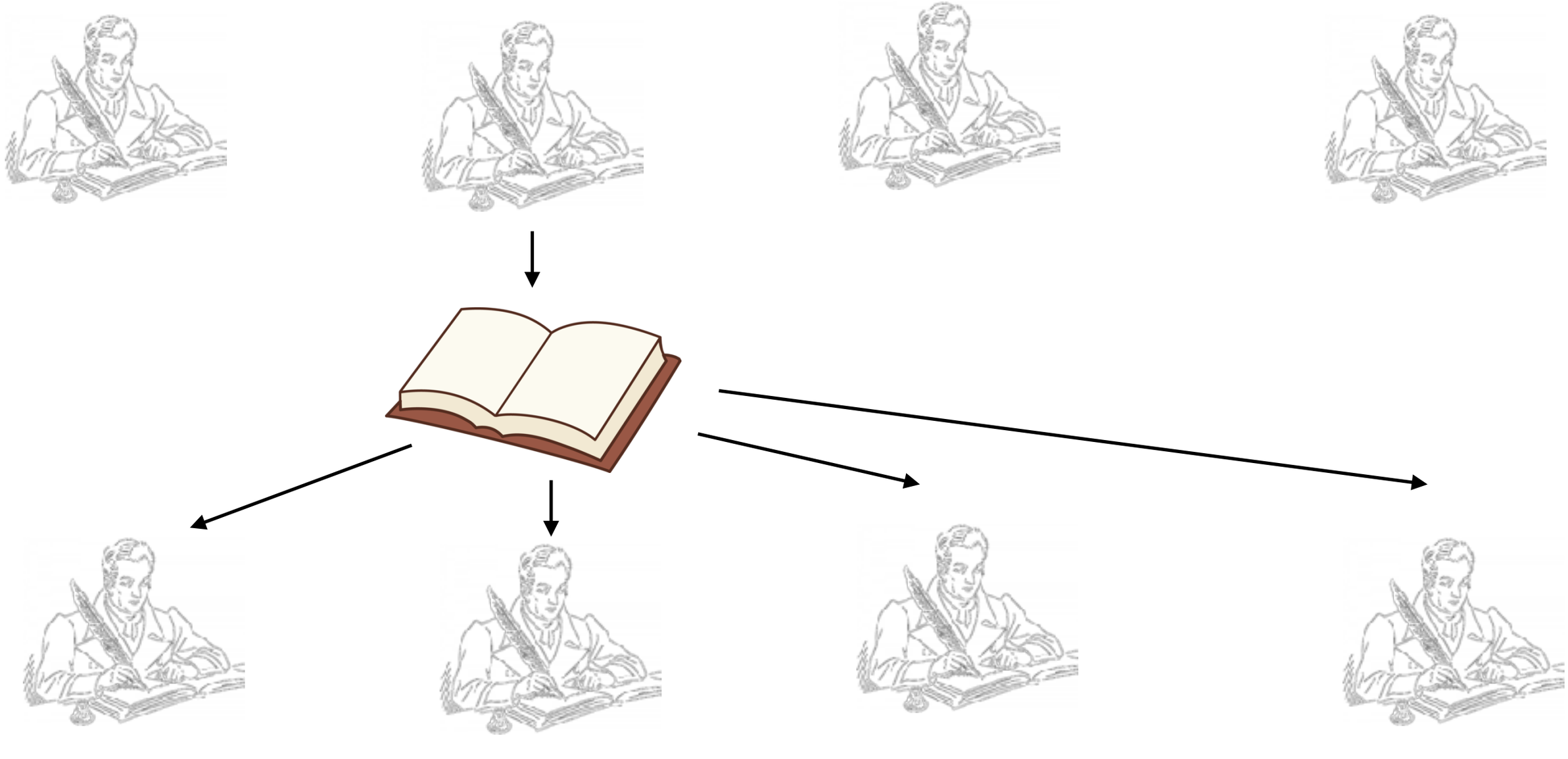
Assembling the current book

- each page refers only to the previous page
- current is assembled by stringing together the longest sequence of pages.



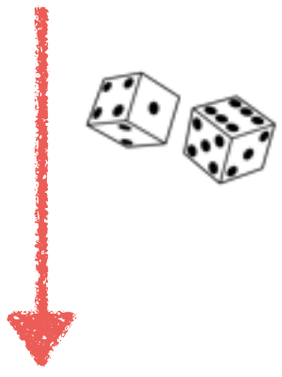
Rules of extending the book

The first scribe that discovers
a page announces it to everyone else



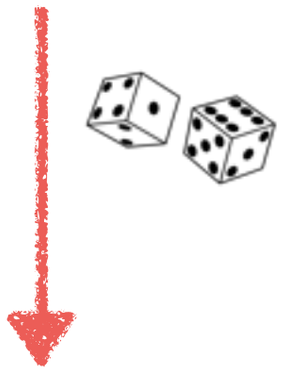
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



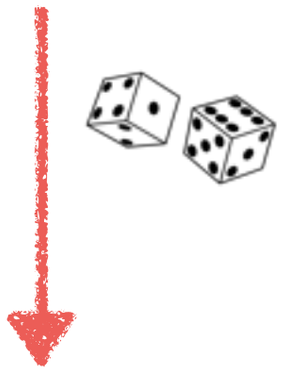
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



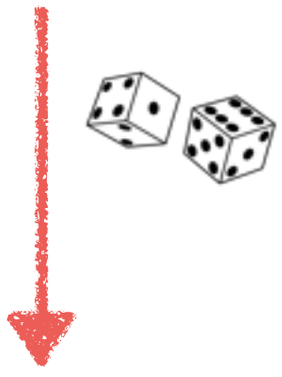
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



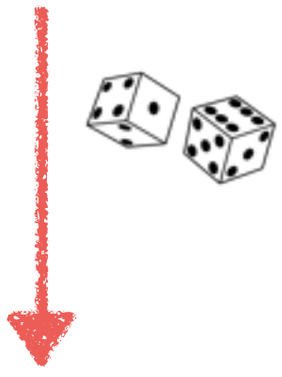
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



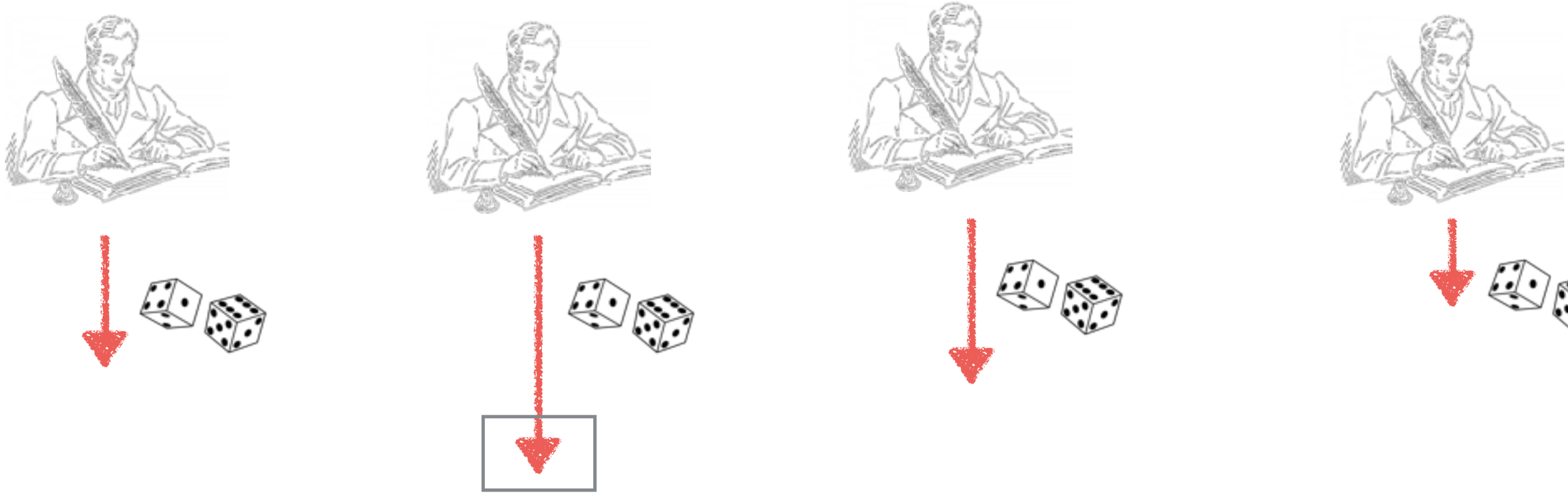
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



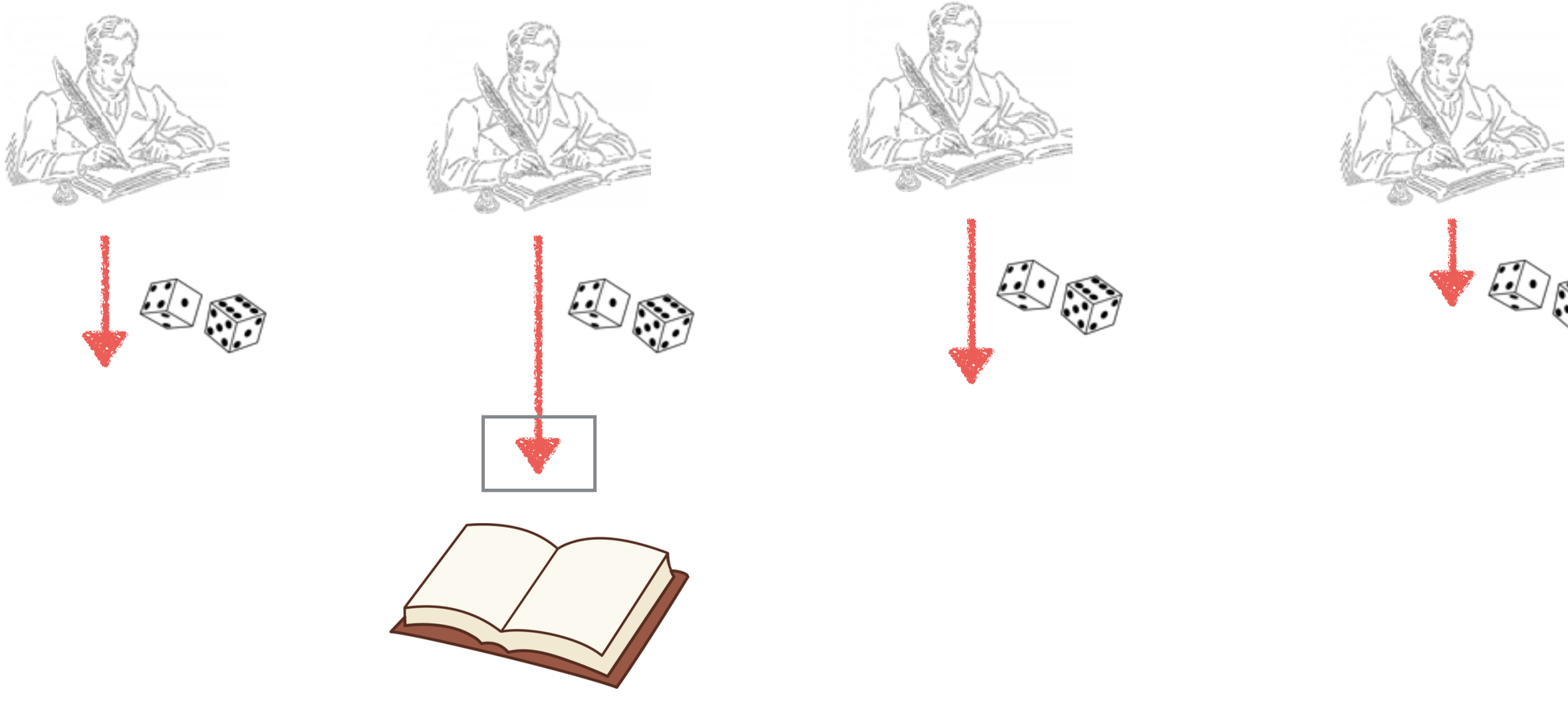
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



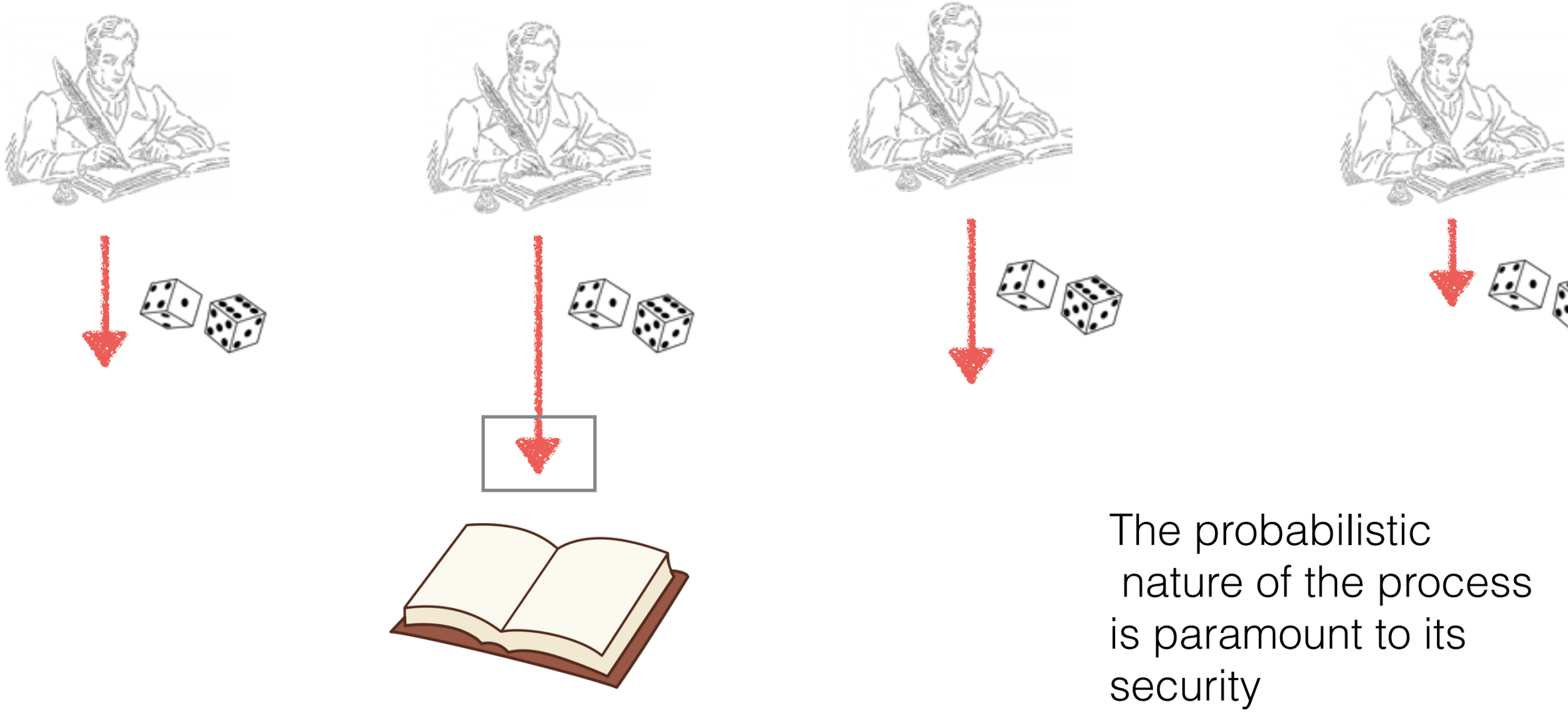
Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



Effort is needed to produce a page

equivalent to : each page needs a special combination from a set of dice to be rolled.



The probabilistic nature of the process is paramount to its security

The benefits of randomness

Imagine two scribes  working together

The benefits of randomness

Imagine two scribes  working together

Unlikely
to continuously
be lucky
together

eventually
one book
will be longer
and be adopted
by both of them

The benefits of randomness

Imagine two scribes  working together

Unlikely
to continuously
be lucky
together

eventually
one book
will be longer
and be adopted
by both of them

Symmetry Breaking

Being a scribe

Being a scribe

- Anyone can be a scribe for the book.

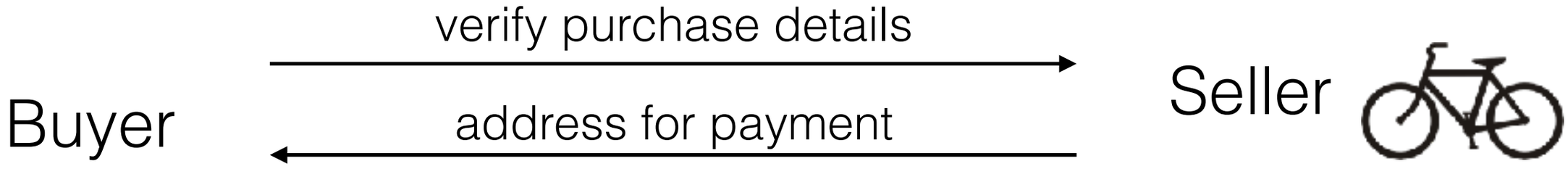
Being a scribe

- Anyone can be a scribe for the book.
- As long as you have a set of dice.

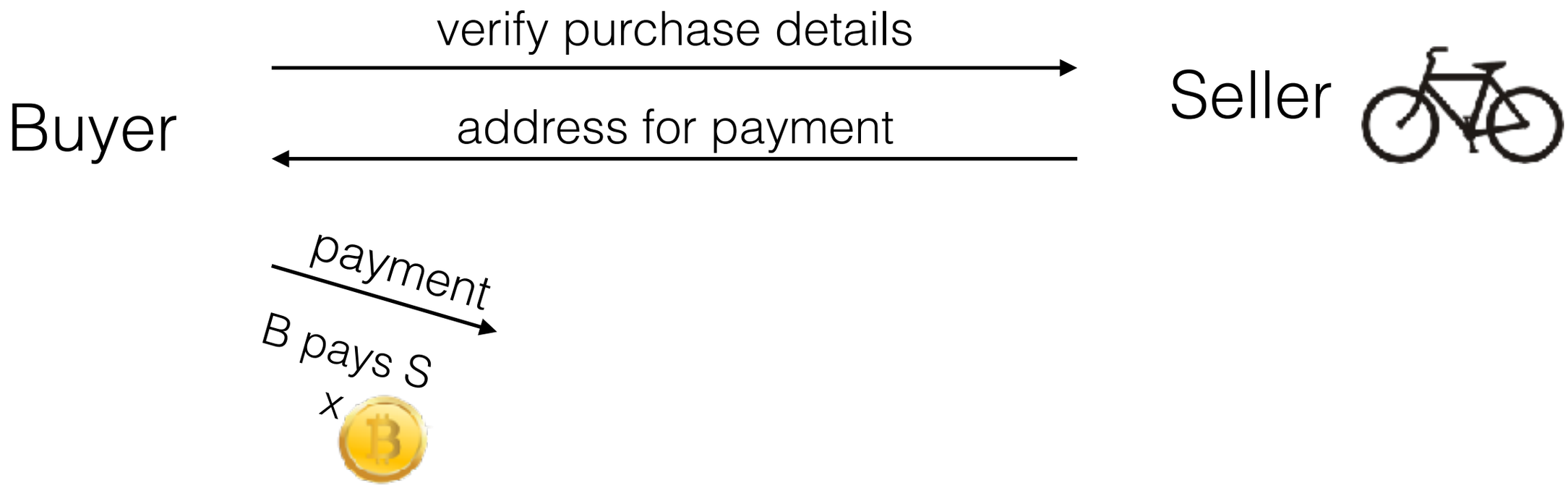
Being a scribe

- Anyone can be a scribe for the book.
- As long as you have a set of dice.
- The more dice one has, the higher the likelihood to produce the winning combination to make a page.

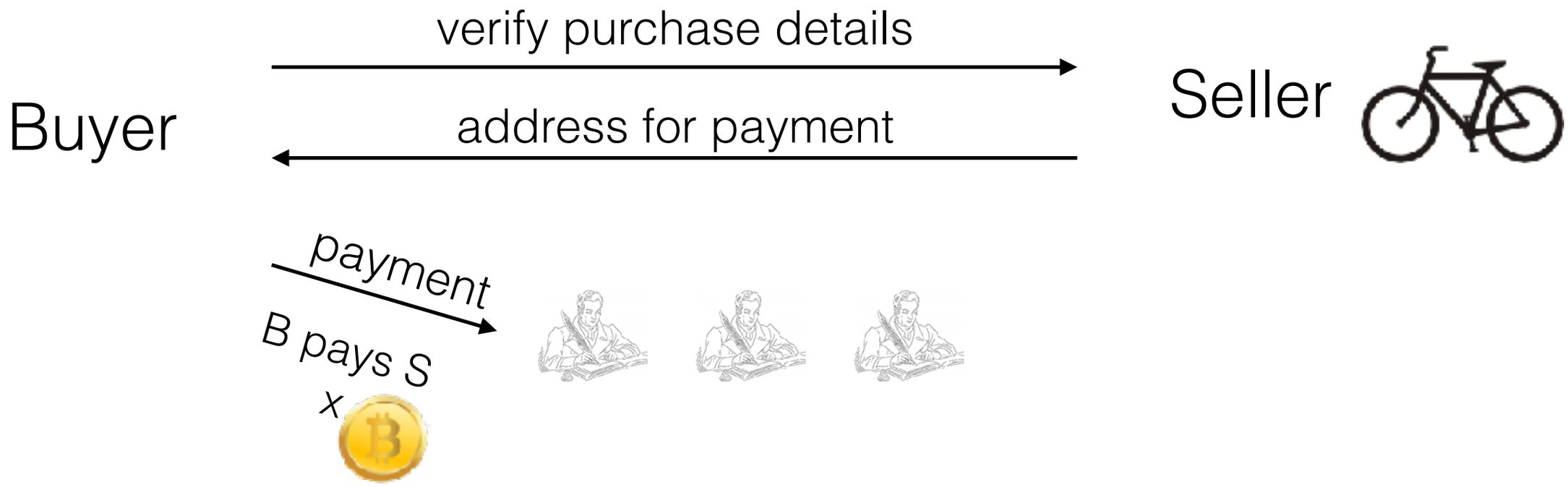
Using the book - Money 3.0



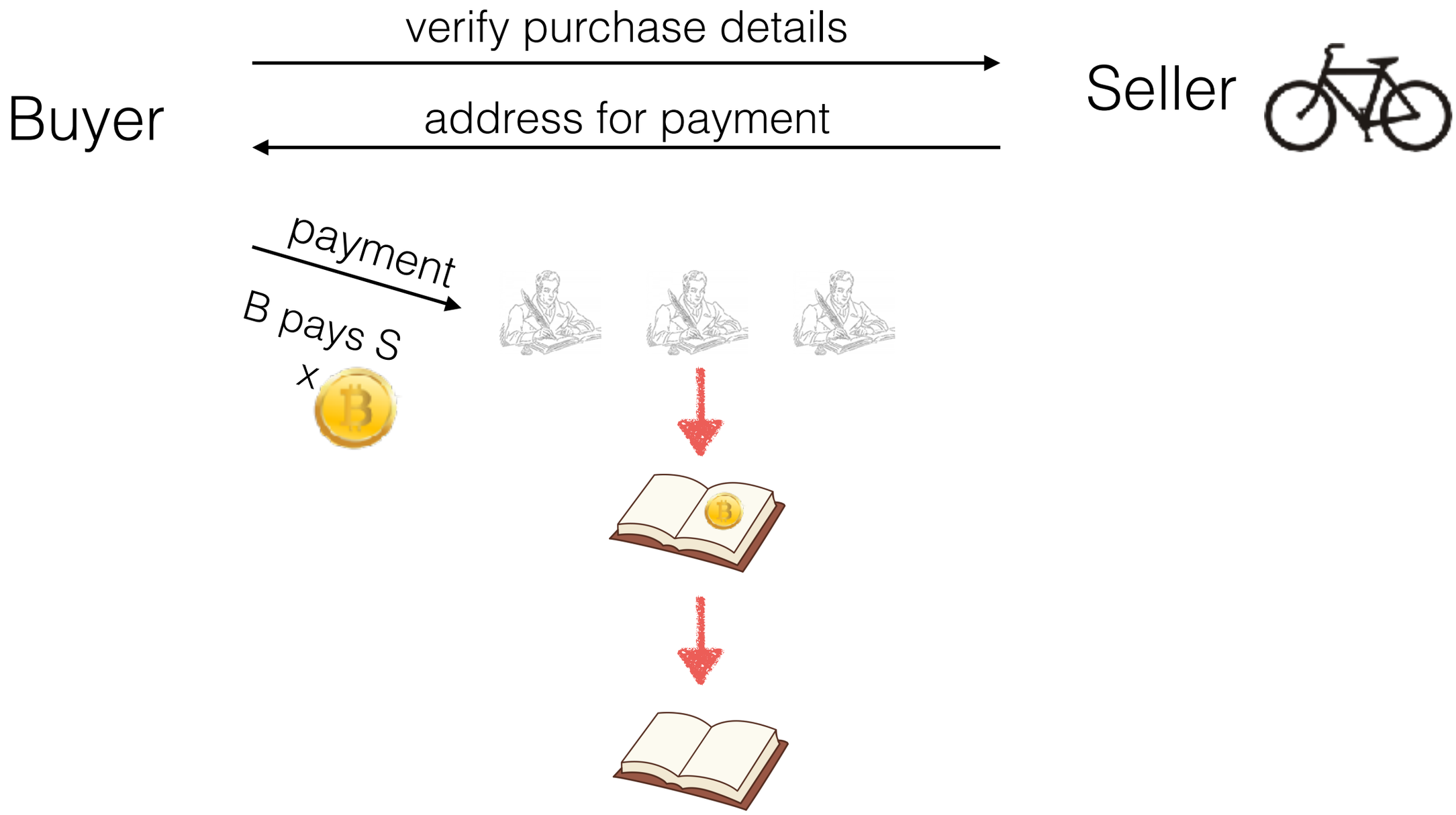
Using the book - Money 3.0



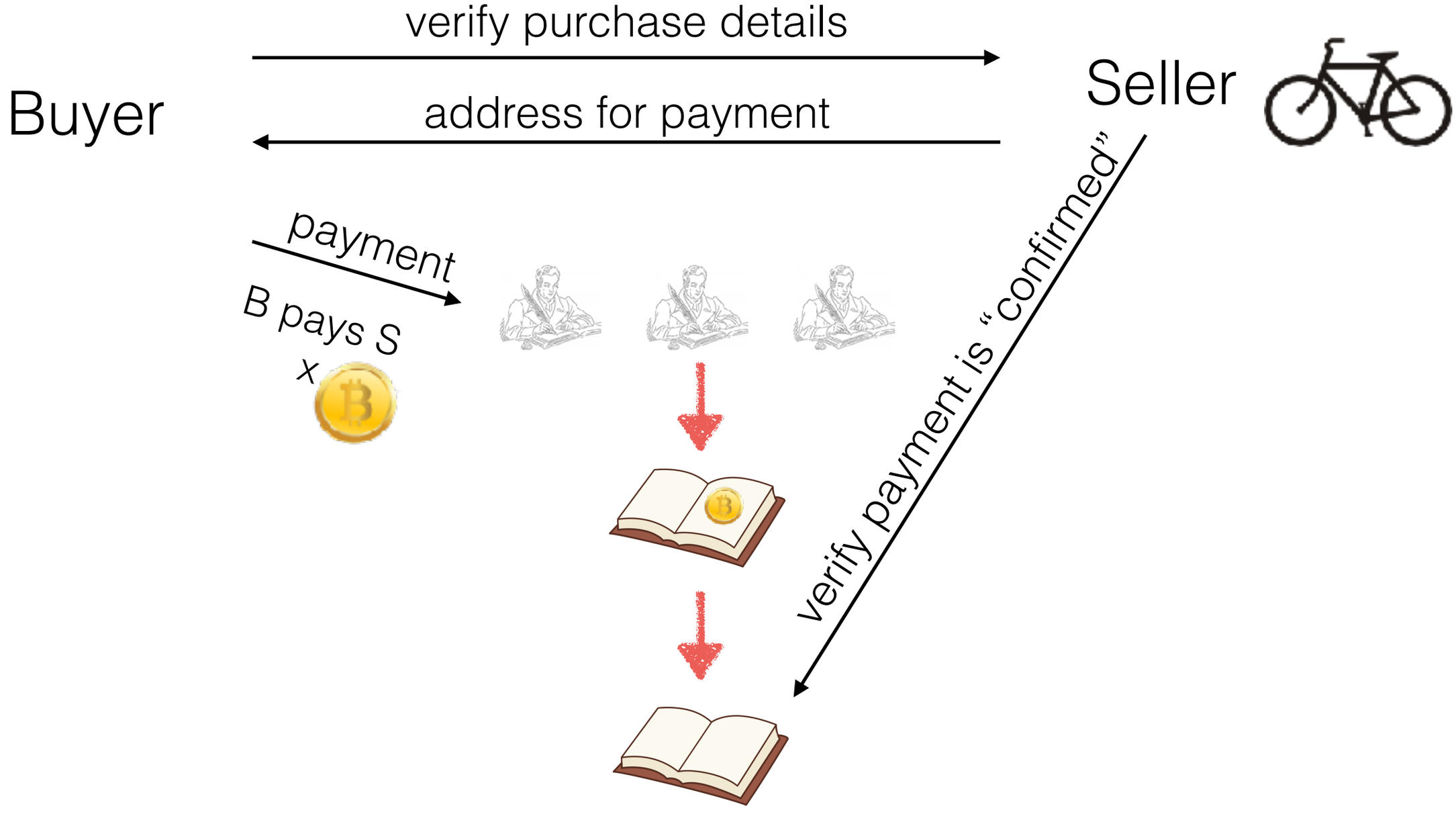
Using the book - Money 3.0



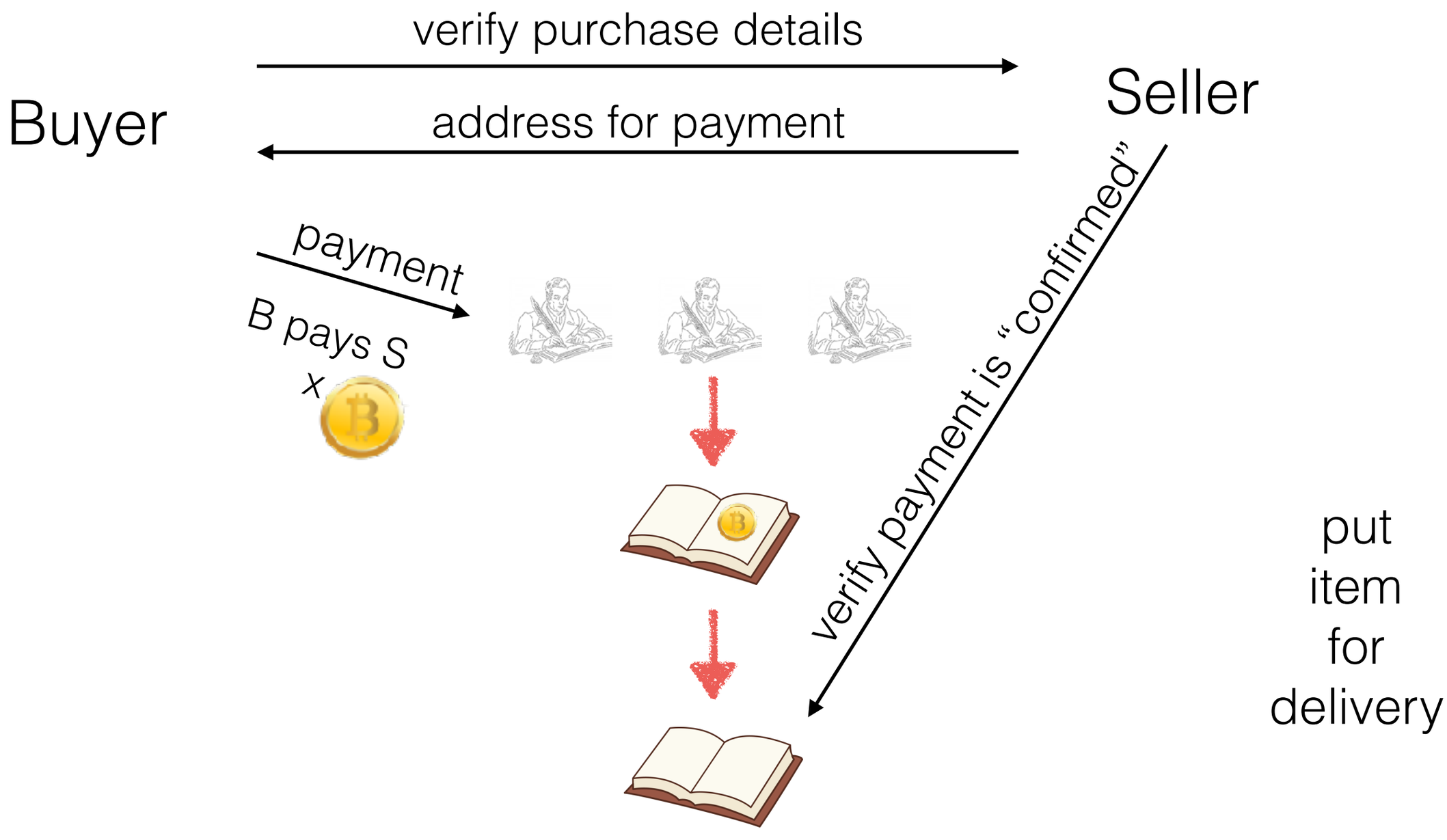
Using the book - Money 3.0



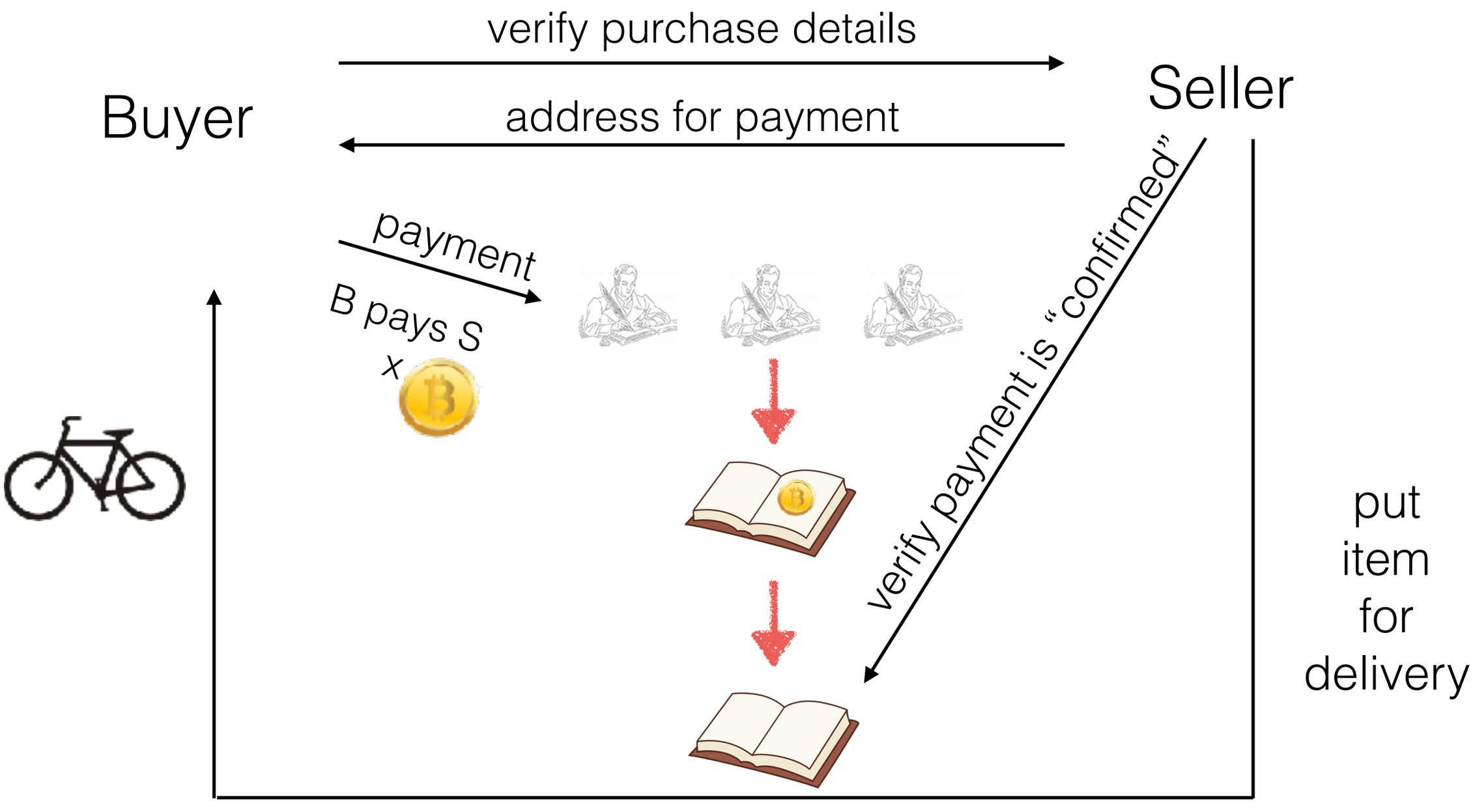
Using the book - Money 3.0



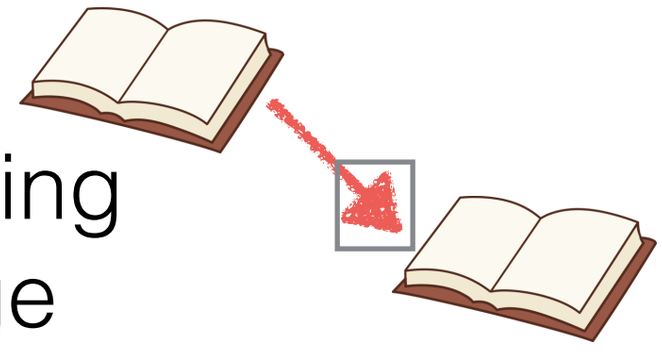
Using the book - Money 3.0



Using the book - Money 3.0



Parable & Reality

book 	the “blockchain”
scribes 	“Miners” / Computer systems that organize transactions in blocks
producing a page 	Solving a **cryptographic puzzle** that is **moderately hard** to solve
rolling a set of dice 	Using a computer to test for a solution from a large space of candidate solutions

Analysis of Money 3.0

- a medium of exchange
- a unit of account
- a store of value

Analysis of Money 3.0

improving
[assuming internet connectivity / adoption]

- a medium of exchange
- a unit of account
- a store of value

Analysis of Money 3.0

- a medium of exchange
- a unit of account
- a store of value

improving
[assuming internet connectivity / adoption]

great!
fungible & divisible.

Analysis of Money 3.0

- a medium of exchange
- a unit of account
- a store of value

improving
[assuming internet connectivity / adoption]

great!
fungible & divisible.

good
[no trusted parties - no natural deterioration]

From Money to Smart Contracts

From Money to Smart Contracts

- Since we have created **the book**, why stop at recording monetary transactions?

From Money to Smart Contracts

- Since we have created **the book**, why stop at recording monetary transactions?
- We can encode in the book's pages **arbitrary relations** between persons.

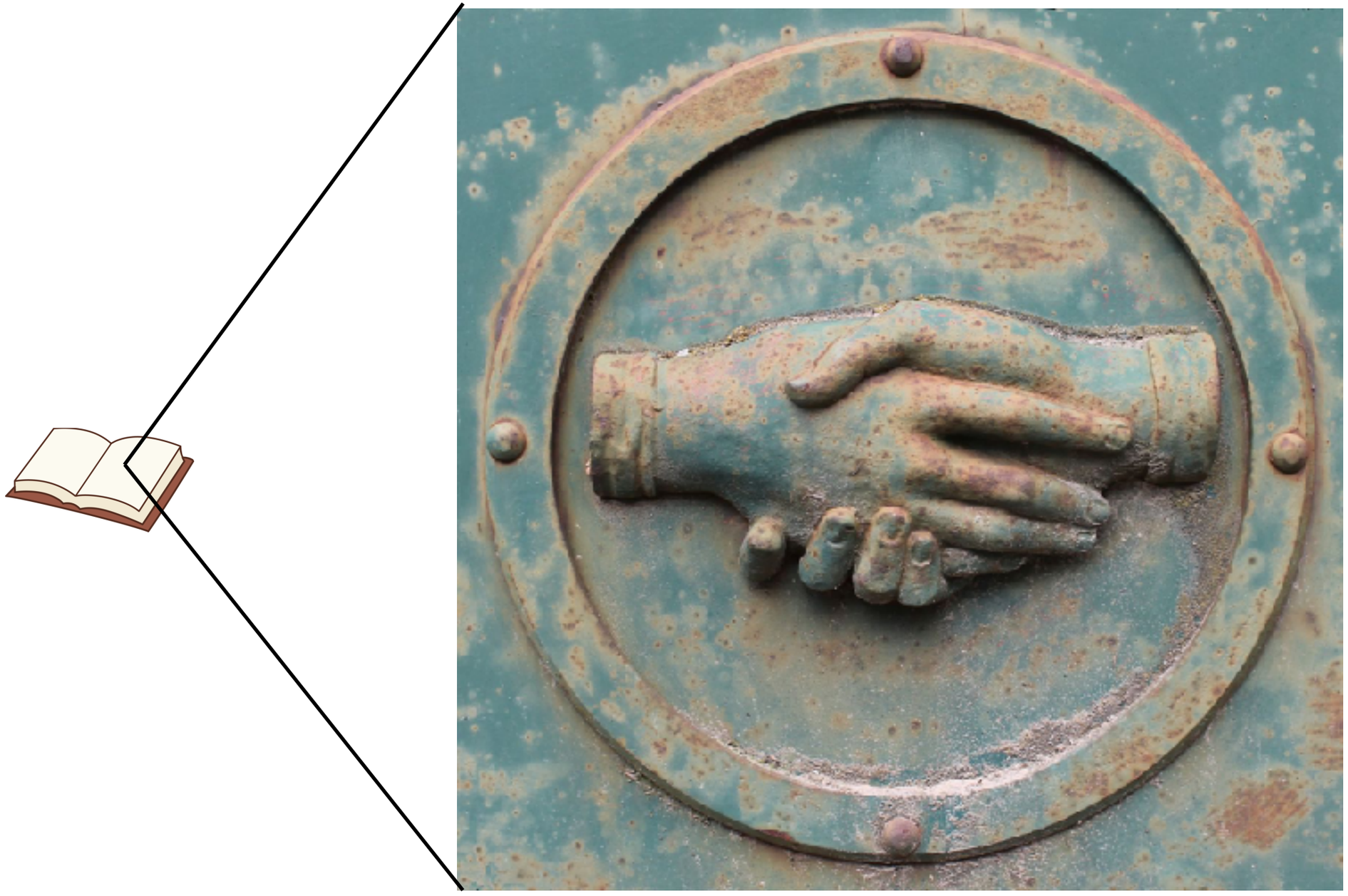
From Money to Smart Contracts

- Since we have created **the book**, why stop at recording monetary transactions?

- We can encode in the book's pages **arbitrary relations** between persons.

- Furthermore, scribes, can perform tasks such as verifying that stakeholders **comply** to contractual obligations … and **take action** if they do not.

Smart Contract



Smart Contract Operation

Smart Contract Operation

- A smart contract is a **piece of code** written in a formal language that records all terms for a certain engagement between a set of persons, “stakeholders.”

Smart Contract Operation

- A smart contract is a **piece of code** written in a formal language that records all terms for a certain engagement between a set of persons, “stakeholders.”
- Stakeholders are identified by their **accounts**.

Smart Contract Operation

- A smart contract is a **piece of code** written in a formal language that records all terms for a certain engagement between a set of persons, “stakeholders.”
- Stakeholders are identified by their **accounts**.
- The smart contract has a **public state**.

Smart Contract Operation

- A smart contract is a **piece of code** written in a formal language that records all terms for a certain engagement between a set of persons, “stakeholders.”
- Stakeholders are identified by their **accounts**.
- The smart contract has a **public state**.
- The smart contract **self executes** each time a certain trigger condition is fulfilled.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

. . .

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

. . .

deposits request for processing

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

...

deposits request for processing

Data processor

data processing engine

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

...

processing

data processing engine

deposits request for processing

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

...

processing

receives output

deposits request for processing

data processing engine

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

auditor

...

processing

data processing engine

deposits request for processing

receives output

Data processor

Application to Privacy Preserving Data Processing

Procedure: Data Subject produces data

data controller creates smart contract to maintain and manage data

...

GDPR compliance

auditor

processing

data processing engine

deposits request for processing

receives output

Data processor

DLT FOR PERSONAL DATA

- Personal data managed by smart contract.
  - Actions that are permitted include updating and effective erasure.
- Requests for processing are also smart contract based.
  - Actions that are permitted include responding to the request.
- Auditing (right of access) can be achieved by parsing the ledger.

A CHALLENGE

- How to **encode** the personal data ?

- How to implement the processor so to comply with **minimum information disclosure**.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

CRYPTOGRAPHY : SECURE MPC

[Goldreich Micali Wigderson 1987]

- (Secure) Multiparty Computation (MPC)
  - Parameterized by function *f(.)*
  - A set of *n* parties contribute inputs *x1, x2, …, xn*
  - At the end of the protocol they compute *f(x1, x2, …, xn)*

MPC CONSTRUCTION IDEA, I

- Consider three roles:
  - Input-providers, Processors, Output-receivers
- Input providers secret-share their input to processors
  - Secret-sharing:

Additive Secret Sharing

$$s_1 + s_2 + \ldots + s_m = x \bmod P$$

MPC CONSTRUCTION IDEA, II

Represent function *f* as Boolean circuit, e.g., XOR, AND, NOT and arithmetize it!

Gates

Addition

Input	a, b
Output	$(a + b) \bmod 2$

Multiplication

Input	a, b
Output	$(a \cdot b) \bmod 2$

NOT

Input	a
Output	$(1 + a) \bmod 2$

(any function can be implemented using these gates)

MPC CONSTRUCTION IDEA, III

XOR GATE

- Suppose *m* parties hold shares of two inputs to an XOR gate.

$$[a], [b] = \langle a_1, \ldots, a_m \rangle, \langle b_1, \ldots, b_m \rangle$$

- How do they calculate shares of the output of the XOR gate?

$$[a] + [b] \bmod 2$$

MPC CONSTRUCTION IDEA, IV

NOT GATE

- Suppose *m* parties hold shares of two inputs to a NOT gate.

$$[a] = \langle a_1, \ldots, a_m \rangle$$

- How do they calculate shares of the output of the NOT gate?

$$[\overline{a}] = \langle 1 + a_1 \bmod 2, a_2, \ldots, a_m \rangle$$

MPC CONSTRUCTION IDEA, V

AND GATE

- Suppose *m* parties hold shares of two inputs to an AND gate.

$$[a], [b] = \langle a_1, \ldots, a_m \rangle, \langle b_1, \ldots, b_m \rangle$$

- How do they calculate shares of the output of the AND gate?

$$[a] \cdot [b] = \langle a_1 b_1 \bmod 2, \ldots, a_m b_m \bmod 2 \rangle$$

but we want: $s_1 + \ldots + s_m = (\sum_{i=1}^{m} a_i)(\sum_{i=1}^{m} b_i)$

MPC CONSTRUCTION IDEA, VI

- Use interaction between parties.
  - Tool : additive homomorphic encryption:
    - it enables:

$$\mathcal{E}(x) \cdot \mathcal{E}(y) = \mathcal{E}(x + y \bmod 2)$$

$$a, b, \mathcal{E}(x) \Rightarrow \mathcal{E}(ax + b)$$

- e.g. Goldwasser-Micali Cryptosystem (Turing awardees 2012).

public-key : N Blum - Integer $N = pq, p \equiv q \equiv 3 \bmod 4$

encryption : $(-1)^m y \bmod N \qquad y \in \mathsf{QR}(N)$

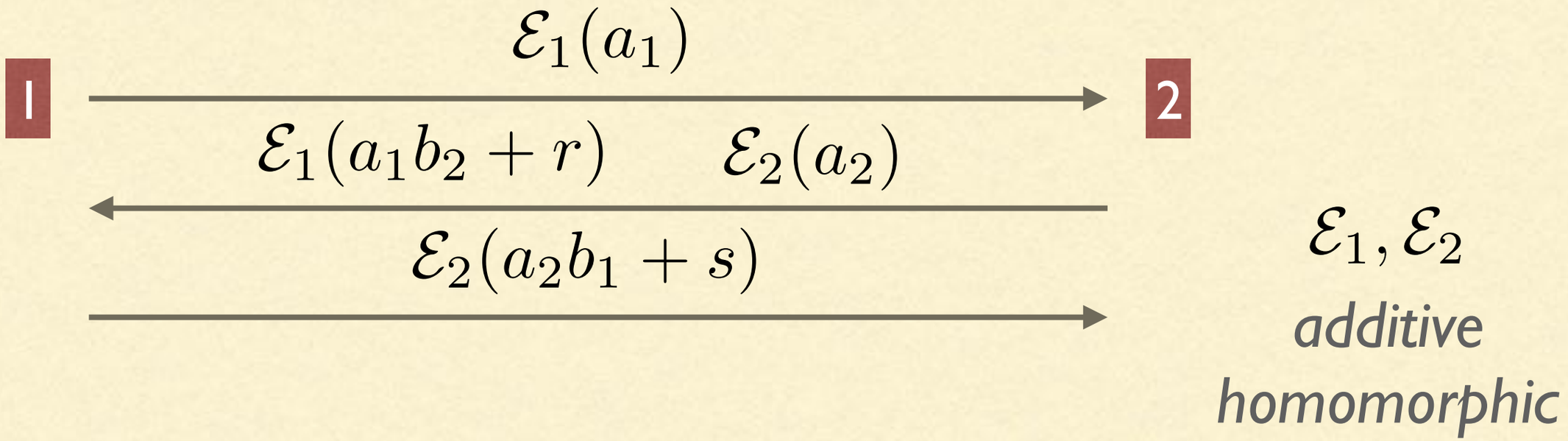
decryption : Test for quadratic residuosity $\psi^{\frac{p-1}{2}} \bmod p = 1$

$\psi^{\frac{q-1}{2}} \bmod q$

MPC CONSTRUCTION IDEA, VII

$$(\sum_{i=1}^{2} a_i)(\sum_{i=1}^{2} b_i) = a_1b_1 + a_2b_2 + a_1b_2 + a_2b_1$$

- *m=2*

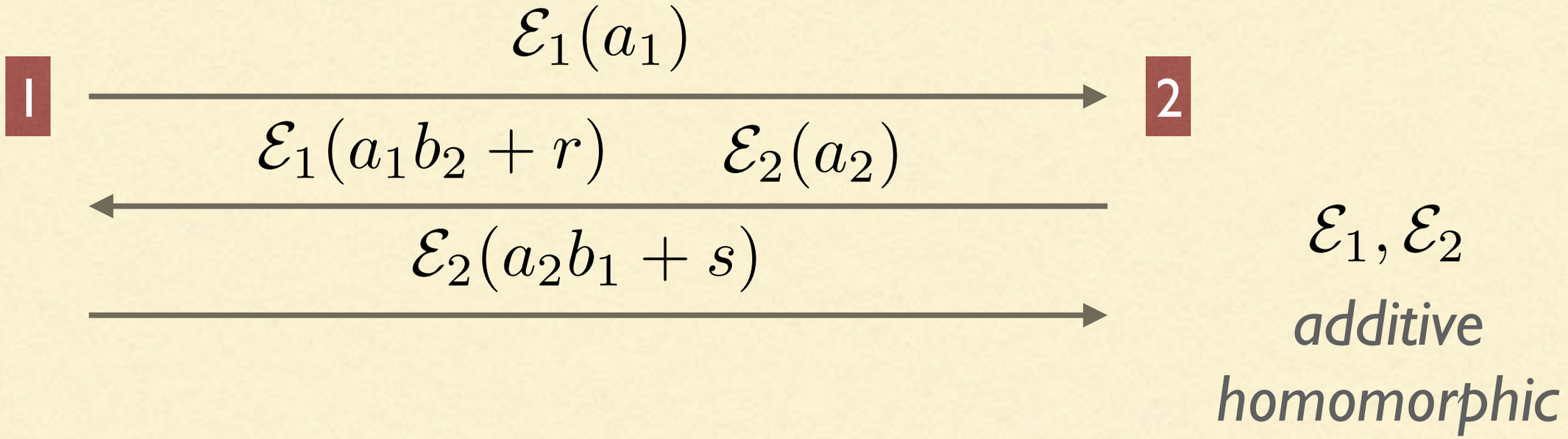


MPC CONSTRUCTION IDEA, VII

$$(\sum_{i=1}^{2} a_i)(\sum_{i=1}^{2} b_i) = a_1b_1 + a_2b_2 + a_1b_2 + a_2b_1$$

$$= (a_1b_1 + a_1b_2 + r - s) + (a_2b_2 + a_2b_1 + s - r)$$

- *m=2*



MPC CONSTRUCTION IDEA, VIII

- There are various cryptographic techniques that achieve simulation of multiplication gates.

- At the end, the processors posses shares of the output wires of the circuit.

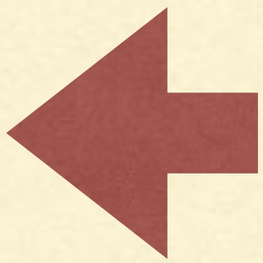
  - Such shares can be encrypted with the output-receivers' key and the result of the computation can be recovered.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.

TALK PLAN

- GDPR and motivation.
- Understanding Distributed Ledger Technology: implementing money.
- Privacy-Preserving Data Processing.
- Secure Multiparty Computation.
- Putting it all together.



PUTTING IT ALL TOGETHER, (1)

- Data gatekeepers, public entities which:
  - will provide a **public-key**.
  - sensitive data will be **locked** under their public-keys jointly.
  - able to **respond** to processing requests by data processors.
  - their existence will be **incentivized** by data processors.

PUTTING IT ALL TOGETHER, (2)

Procedure: Data Subject selects data gatekeepers
& encodes data into the smart contract.

$$\mathcal{E}_1(a_1), \ldots, \mathcal{E}_m(a_m)$$

$$\sum_{i=1}^{m} a_i = x$$

PUTTING IT ALL TOGETHER, (2)

Procedure: Data Subject selects data gatekeepers
& encodes data into the smart contract.

$$\mathcal{E}_1(a_1), \ldots, \mathcal{E}_m(a_m)$$

creates smart
contract containing
data encoding using
data gatekeeper's PK's

$$\sum_{i=1}^{m} a_i = x$$

PUTTING IT ALL TOGETHER, (2)

Procedure: Data Subject selects data gatekeepers
& encodes data into the smart contract.

$$\mathcal{E}_1(a_1), \ldots, \mathcal{E}_m(a_m)$$

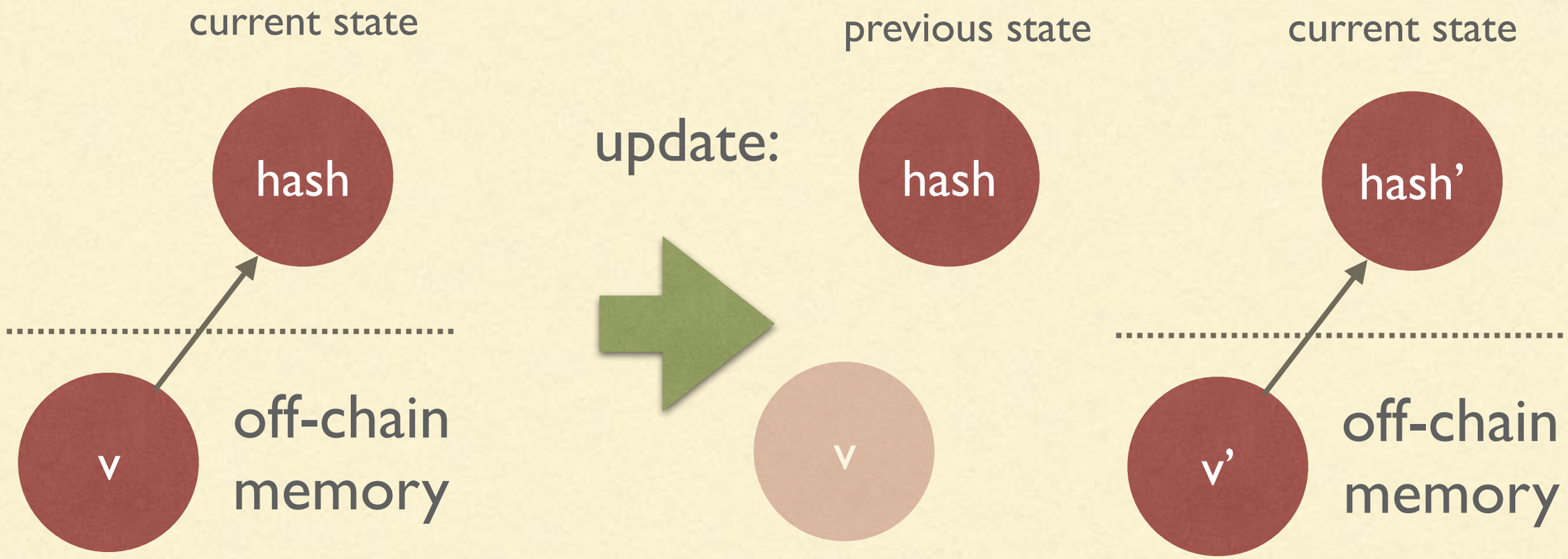
creates smart
contract containing
data encoding using
data gatekeeper's PK's

$$\sum_{i=1}^{m} a_i = x$$

Data subject
updates data,
or marks them as **erased**.

UPDATES, (1)

- Smart contract contains a Merkle-Tree:



UPDATES, (2)

- Updating a smart contract, requires a secret-key (to prove ownership)

- Either the data subject or the data controller can maintain secret-key (or event : **jointly**).

- In the case of joint key ownership, an update **requires** interaction between data subject and data controller.

PUTTING IT ALL TOGETHER, (3)

gatekeeper

gatekeeper

gatekeeper

Data
Processor

PUTTING IT ALL TOGETHER, (3)

gatekeeper

gatekeeper

gatekeeper

deposits
request
for processing
as a S.C.

Data
Processor

PUTTING IT ALL TOGETHER, (3)

gatekeeper

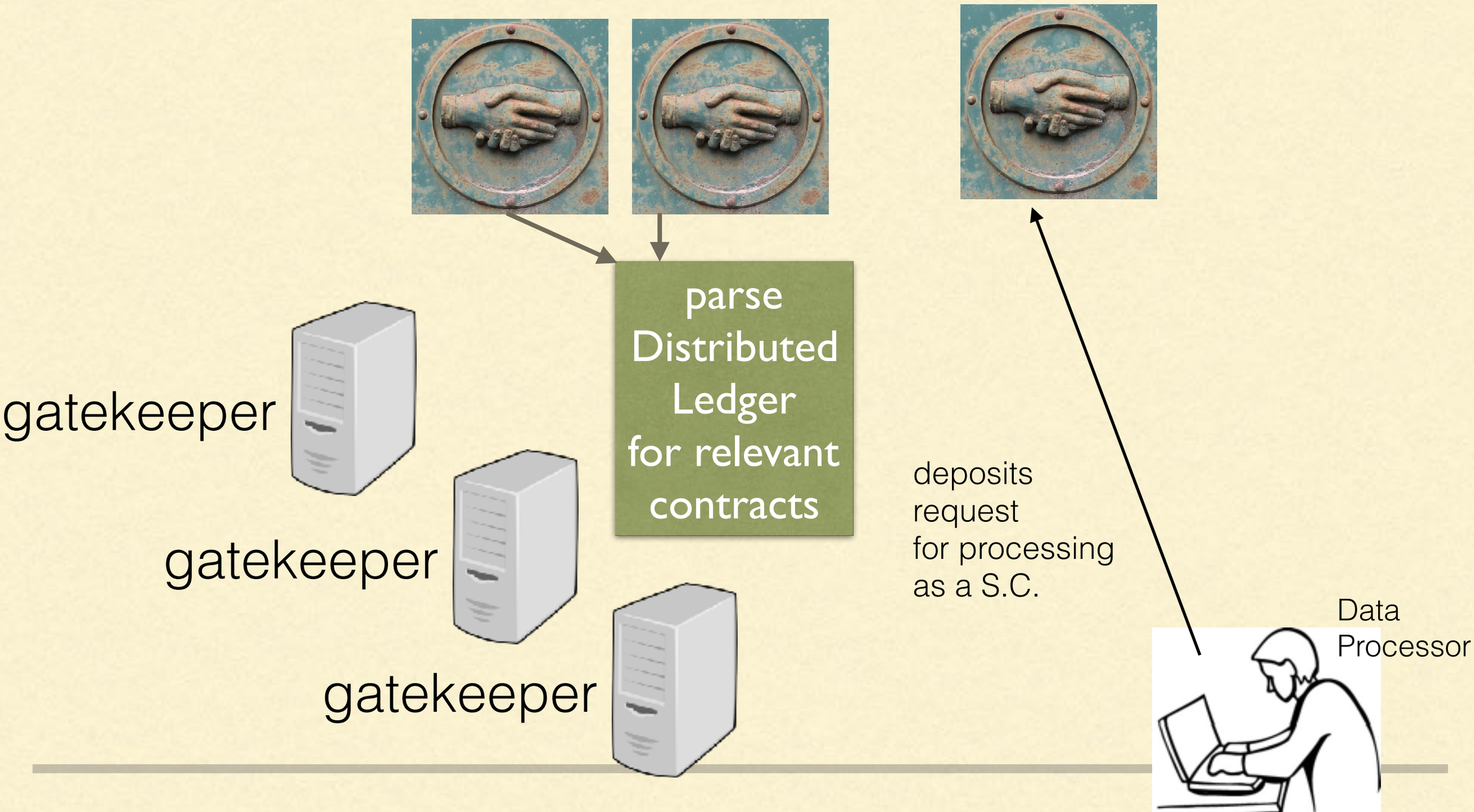
gatekeeper

gatekeeper

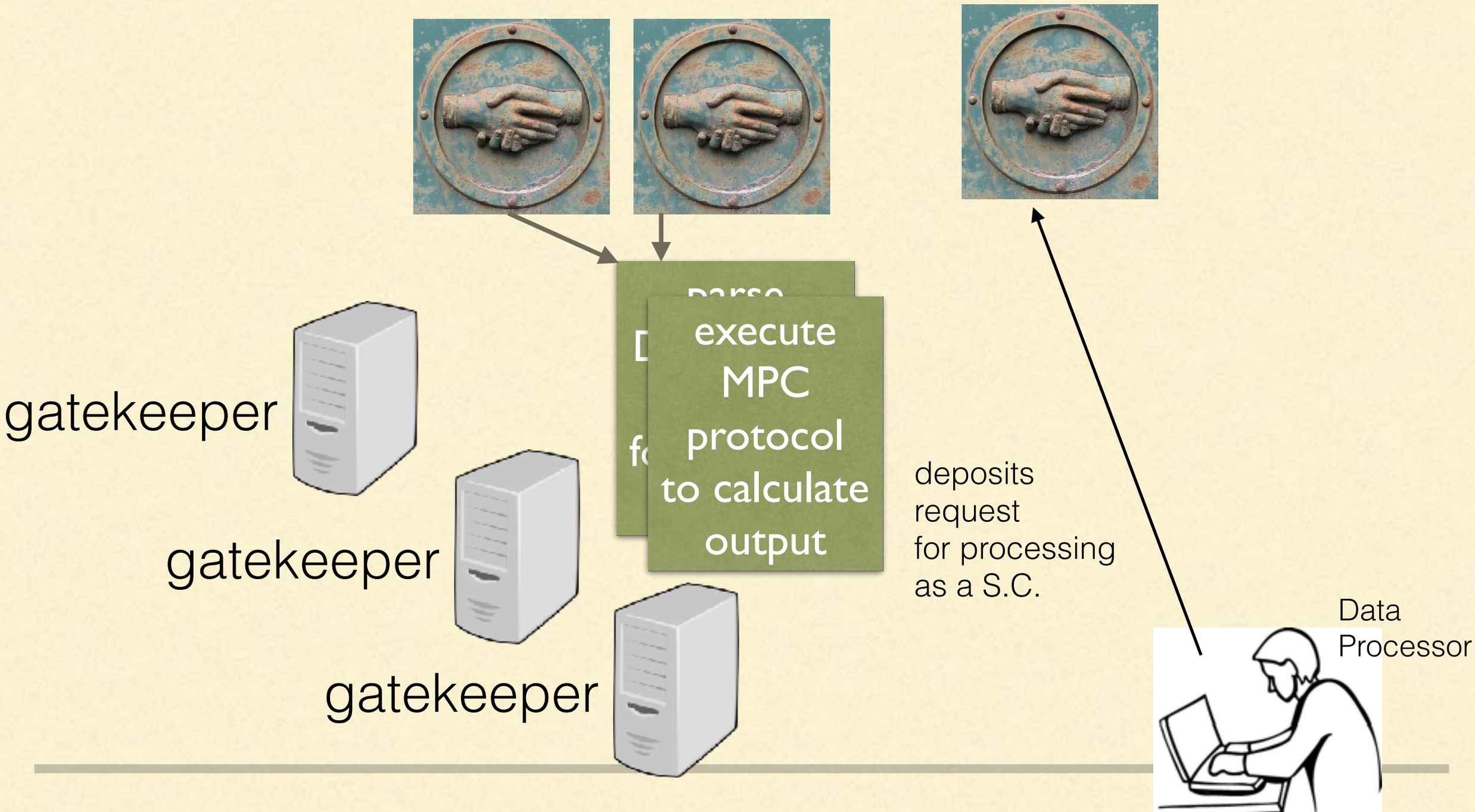
deposits
request
for processing
as a S.C.

Data
Processor

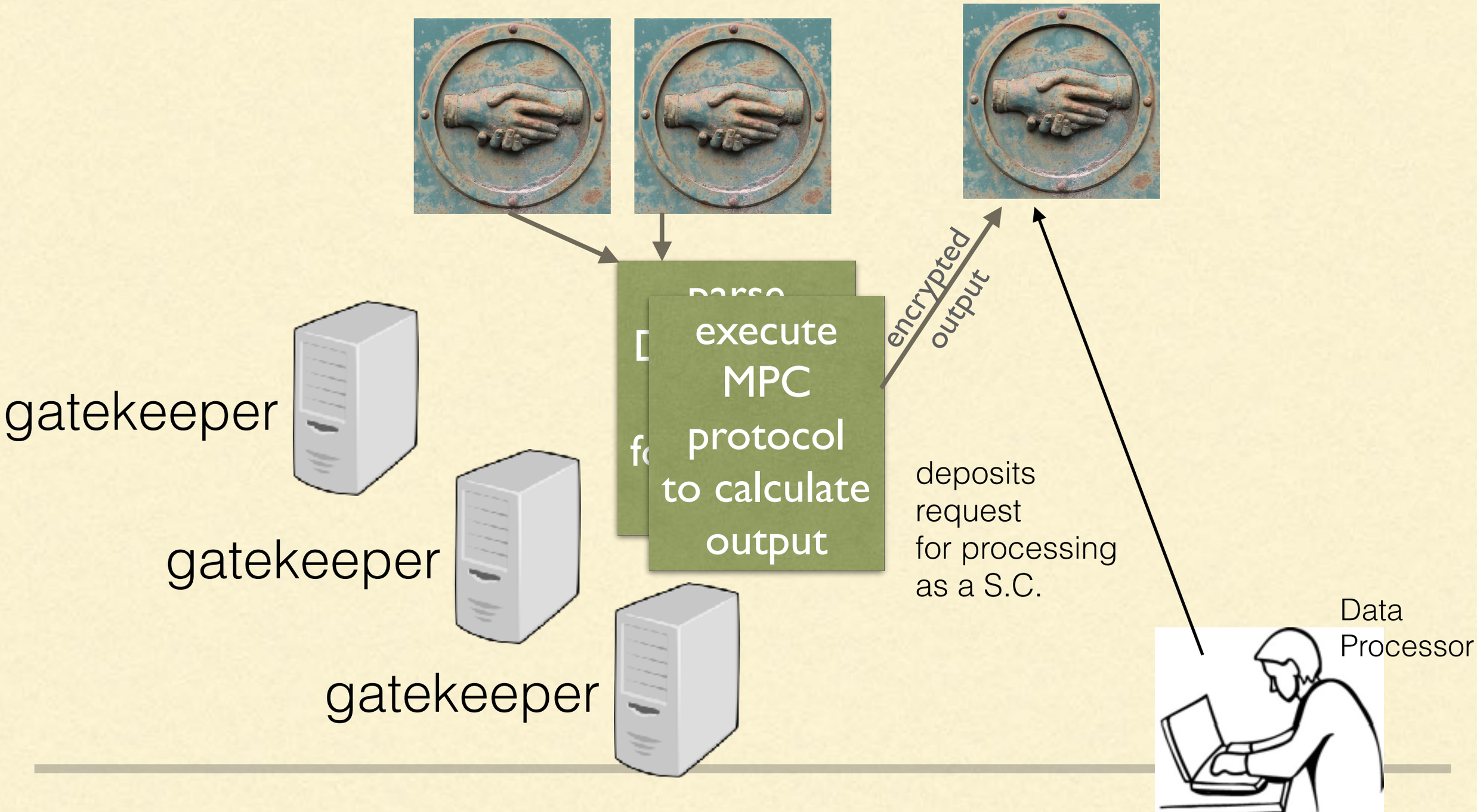
PUTTING IT ALL TOGETHER, (3)



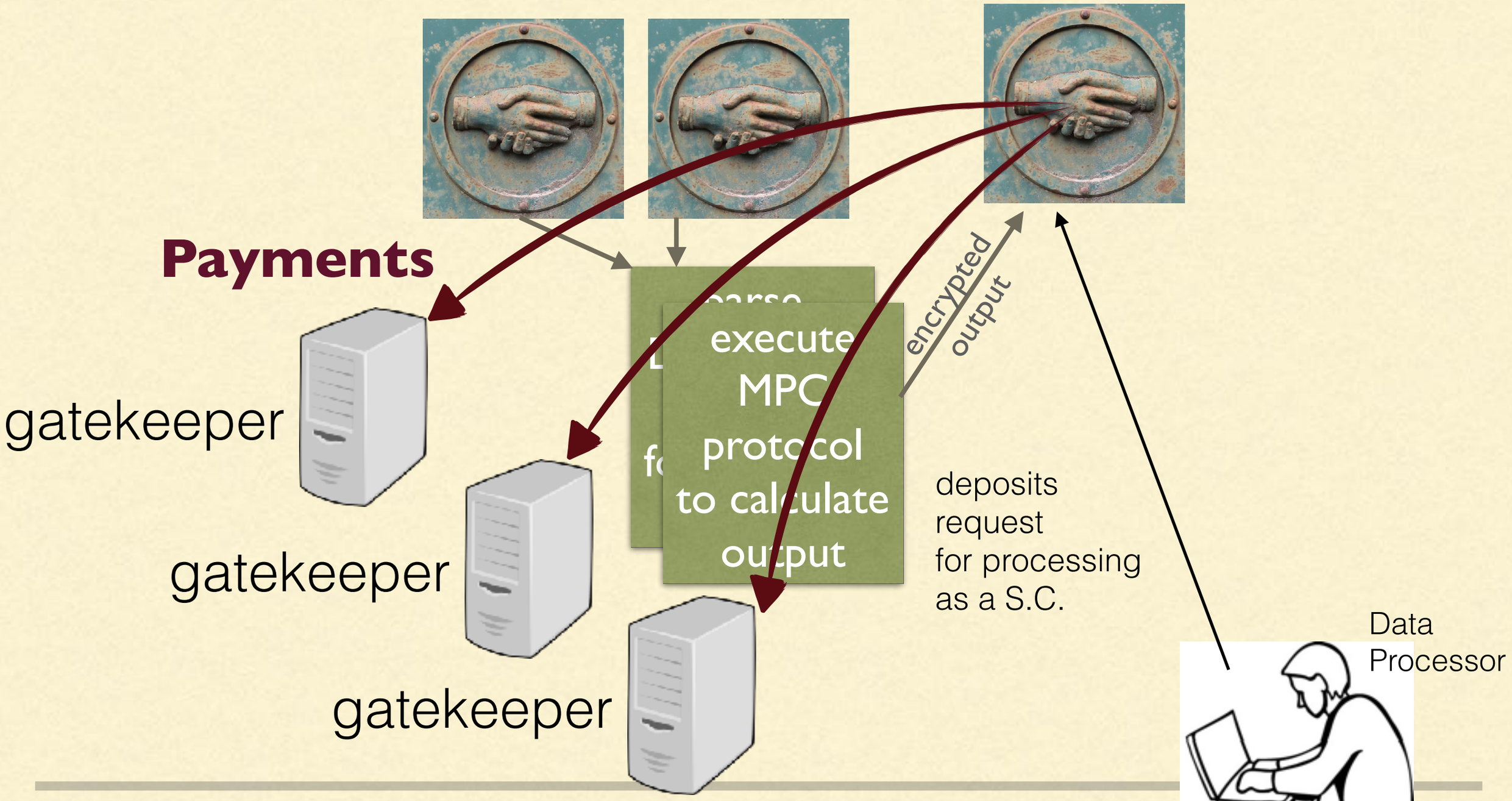
PUTTING IT ALL TOGETHER, (3)



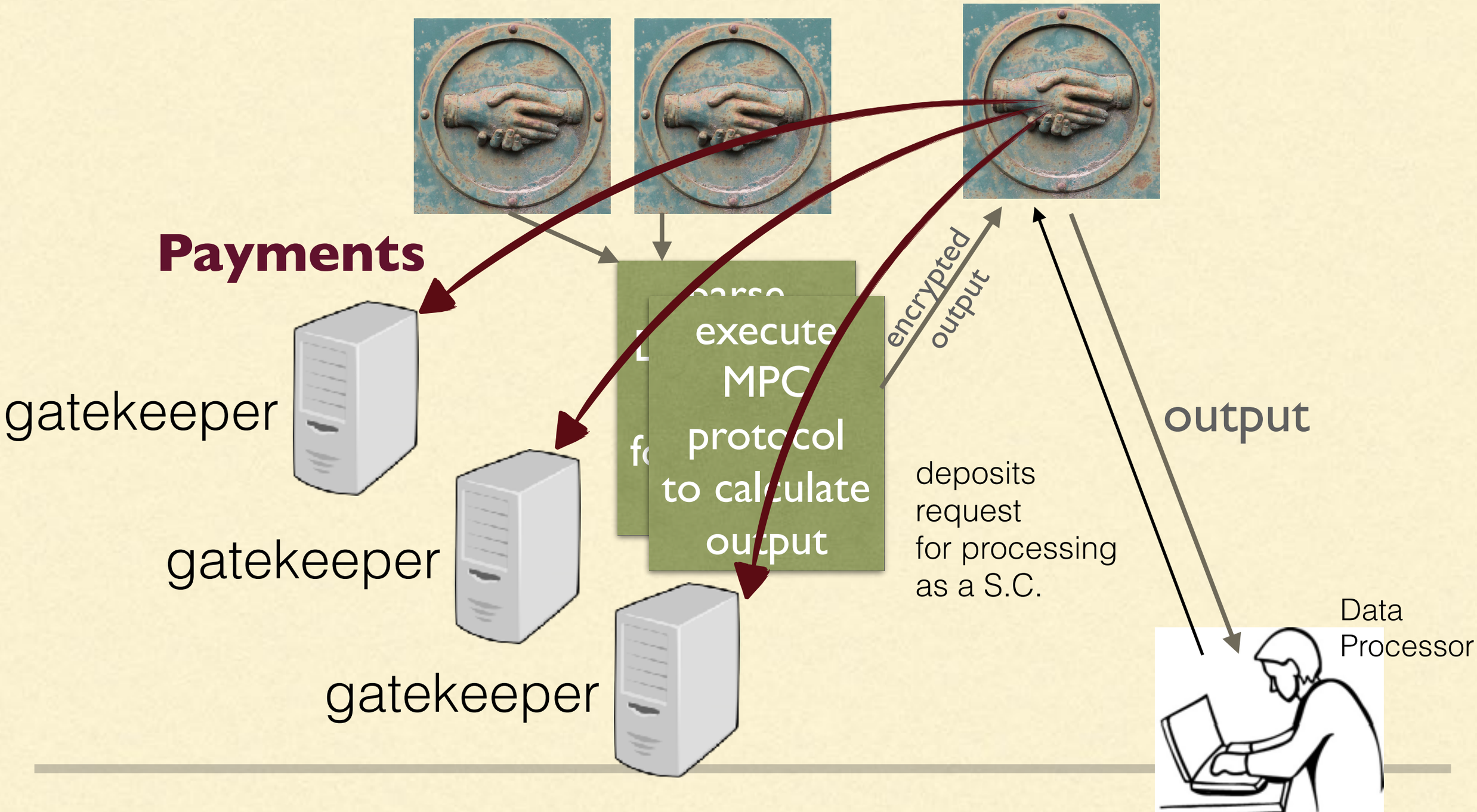
PUTTING IT ALL TOGETHER, (3)



PUTTING IT ALL TOGETHER, (3)



PUTTING IT ALL TOGETHER, (3)



CONCLUSIONS

- Positive use of DLT for improving GDPR inspired compliance issues.
- Many open questions remain:
  - improve performance of secure multiparty computation protocols.
  - **integration** of MPC / blockchain, DLT.
  - security & game theoretic **analysis**.

BLOCKCHAIN SYSTEMS & PRIVACY

Aggelos Kiayias

University of Edinburgh & IOHK

project PANORAMIX