On Purpose and by Necessity: Compliance under the GDPR

Søren Debois, IT University of Copenhagen
Joint work with David Basin (ETH) and Thomas Hildebrandt (KU)

FC ’18, Feb 26, 2018

General Data Protection Regulation

May 25, 2018

GDPR

- EU **G**eneral **D**ata **P**rotection **R**egulation
- In force May 25, 2018.
- Teeth! Fines up to 20 million EUR or 4% of **world-wide turnover**, whichever is higher [9, Article 83, §5]
- Long list of mandates: Right to be forgotten, Right of access, Right of rectification, Right to erasure, Right to restrict processing, Right of data portability, Obligation to inform, Right to not be evaluated on the basis of automated processing, …
- In particular …

GDPR

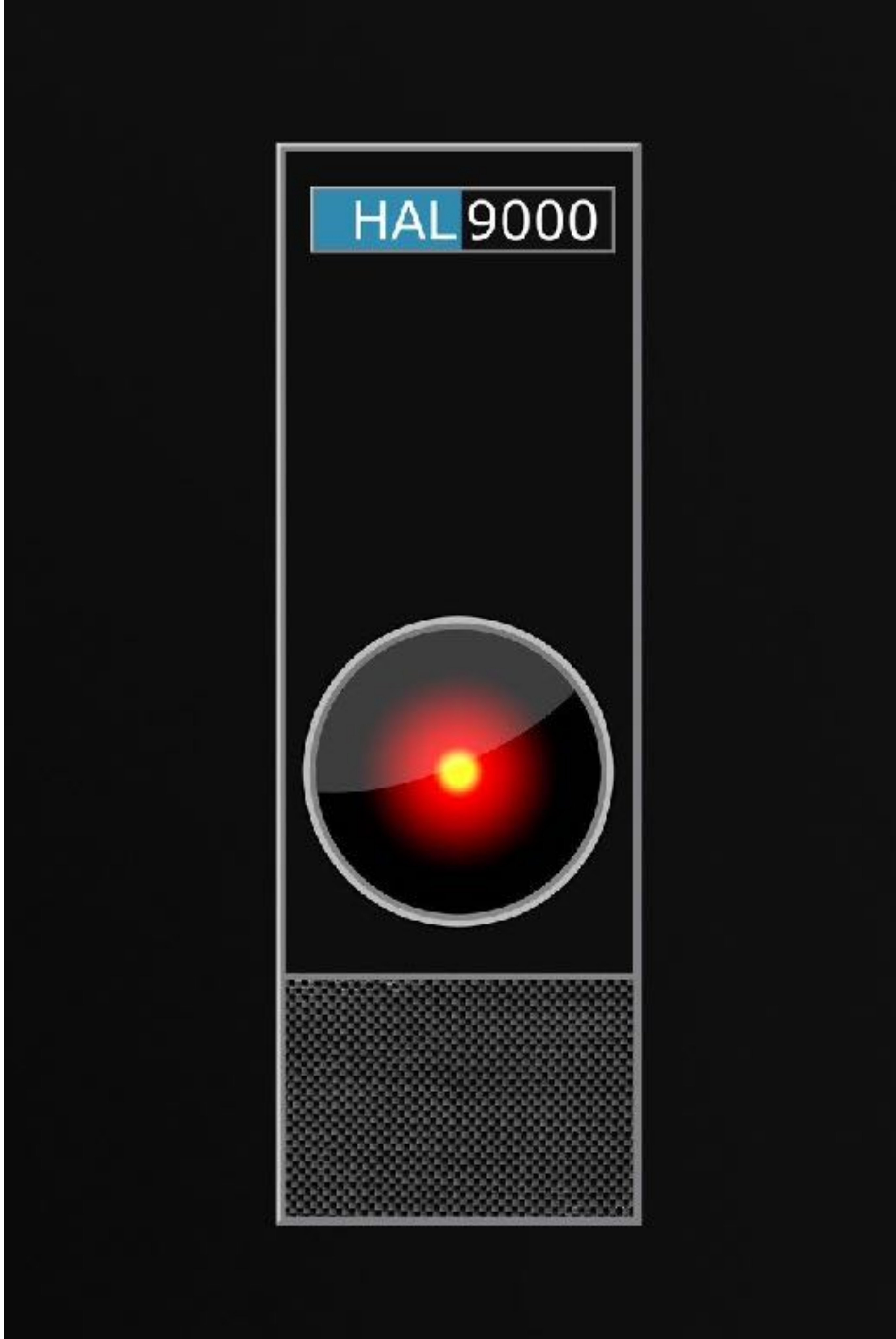
- Purpose limitation [9, Article 5, §1(b)]:
  “[Personal data shall be] collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; [...]”

- Data minimisation [9, Article 5, §1(c)]:
  “[Personal data shall be] adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed [...] ”

- Consent [9, Recital (32)]:
  “Consent should cover all processing activities carried out for the same purpose or purposes.”

Automatic Audits?

- Not at present
- Purpose limitation, connecting purpose to data, is beyond computers—e.g., is a text "advertising" or "political propaganda"?
- Data minimisation is beyond computers—do we really need your images to fulfil grocery orders?
- Proposal: Computer-*supported* audits.



GDPR

That is, the GDPR requires that personal data can only:

1. be collected for a purpose,
2. to which the user has consented, and
3. be necessary to achieve that purpose;
4. moreover the collected data must be deleted when it is no longer necessary for any purpose.

GDPR

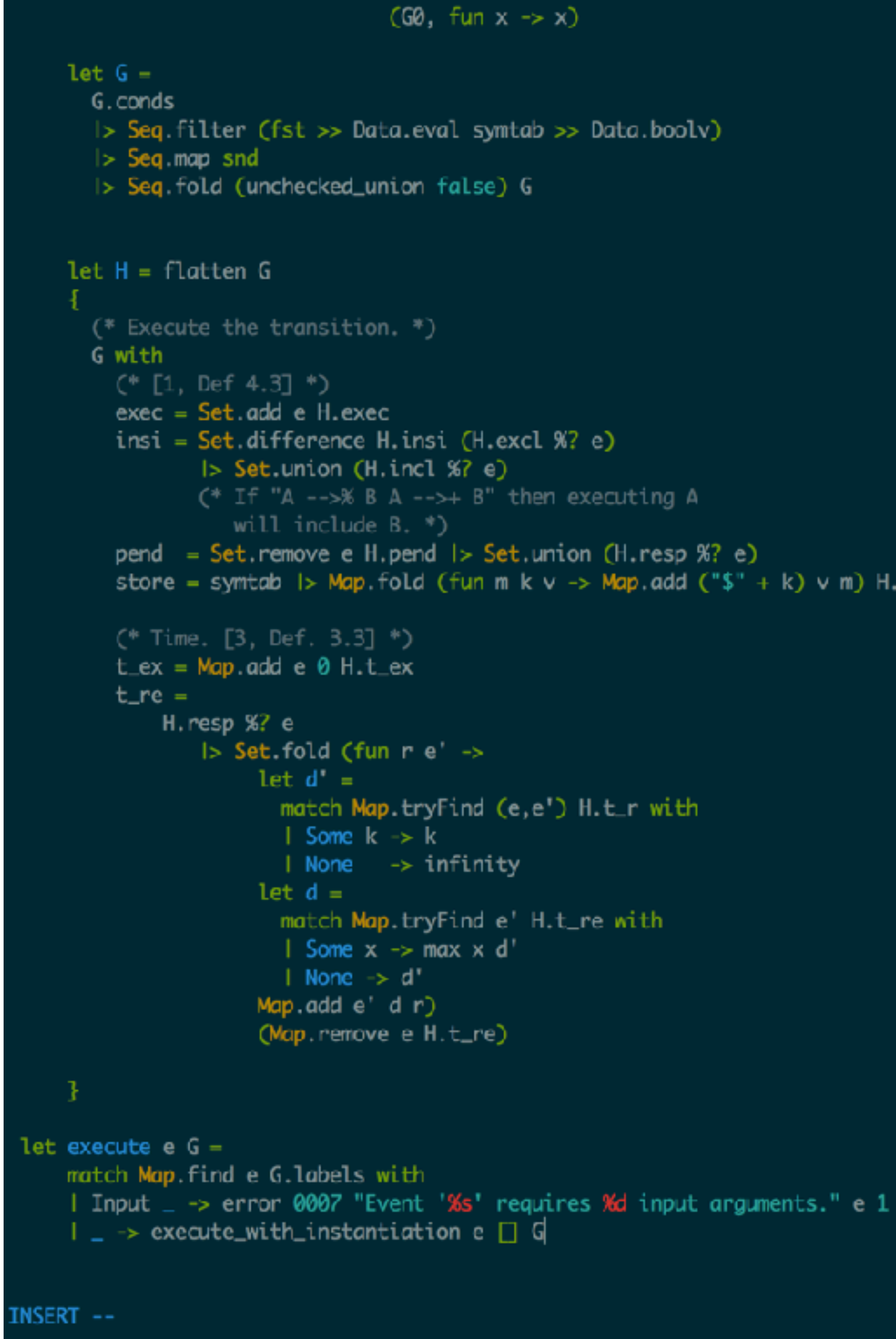
That is, the GDPR requires that personal data can only:

1. be collected for a purpose,
2. to which the user has consented
3. be necessary to achieve that purpose;
4. moreover the collected data must be deleted when it is no longer necessary for any purpose.

We are generally not ready for this!

Where is the purpose?

It’s not a programming concept



```
                               (G0, fun x -> x)

    let G =
      G.conds
      |> Seq.filter (fst >> Data.eval symtab >> Data.boolv)
      |> Seq.map snd
      |> Seq.fold (unchecked_union false) G


    let H = flatten G
    {
      (* Execute the transition. *)
      G with
        (* [1, Def 4.3] *)
        exec = Set.add e H.exec
        insi = Set.difference H.insi (H.excl %? e)
               |> Set.union (H.incl %? e)
               (* If "A -->% B A -->+ B" then executing A
                  will include B. *)
        pend  = Set.remove e H.pend |> Set.union (H.resp %? e)
        store = symtab |> Map.fold (fun m k v -> Map.add ("$" + k) v m) H.

        (* Time. [3, Def. 3.3] *)
        t_ex = Map.add e 0 H.t_ex
        t_re =
            H.resp %? e
               |> Set.fold (fun r e' ->
                    let d' =
                      match Map.tryFind (e,e') H.t_r with
                      | Some k -> k
                      | None    -> infinity
                    let d =
                      match Map.tryFind e' H.t_re with
                      | Some x -> max x d'
                      | None -> d'
                    Map.add e' d r)
                    (Map.remove e H.t_re)

    }

let execute e G =
    match Map.find e G.labels with
    | Input _ -> error 0007 "Event '%s' requires %d input arguments." e 1
    | _ -> execute_with_instantiation e [] G


INSERT --
```

Business Processes

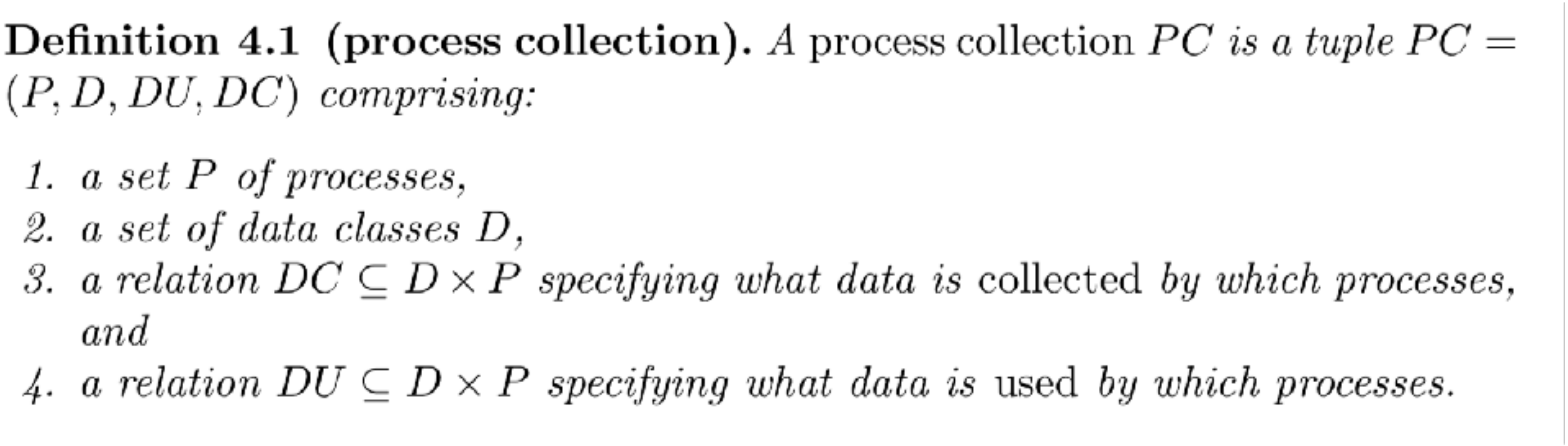
- E.g., sales, marketing, procurement, order fulfillment, loan processing, …

- “a structured, measured set of activities designed to produce a specific output for a particular customer or market. A process is thus a specific ordering of work activities across time and space […]” [Davenport ’94]

- “specific output for […] customer” ~ “purpose”

- But: data collected in one process (for one purpose) may migrate to other processes

Process collections



Definition 4.1 (process collection). *A* process collection *PC is a tuple $PC = (P, D, DU, DC)$ comprising:*

1. *a set P of processes,*
2. *a set of data classes D,*
3. *a relation $DC \subseteq D \times P$ specifying what data is* collected *by which processes, and*
4. *a relation $DU \subseteq D \times P$ specifying what data is* used *by which processes.*

What is a “Data class”?

- You give consent for me using your <data class> for <purpose>.

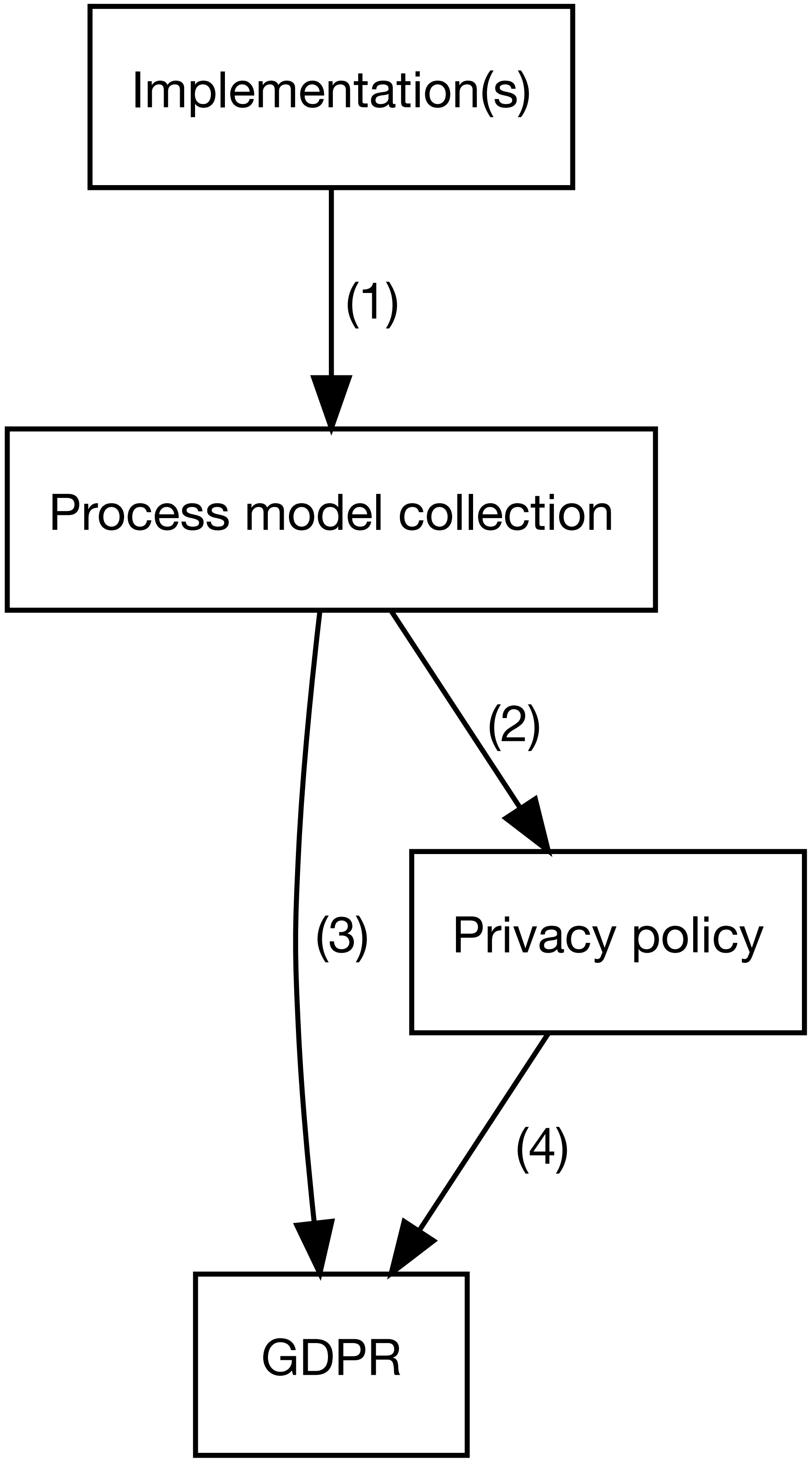
- Examples: Name, address, credit card.

- Also examples: Personal information, payment details, order history.

- Non-examples: Specific data.
  “Søren”, “Thomas”, and “David” are not data classes.

Decomposing compliance

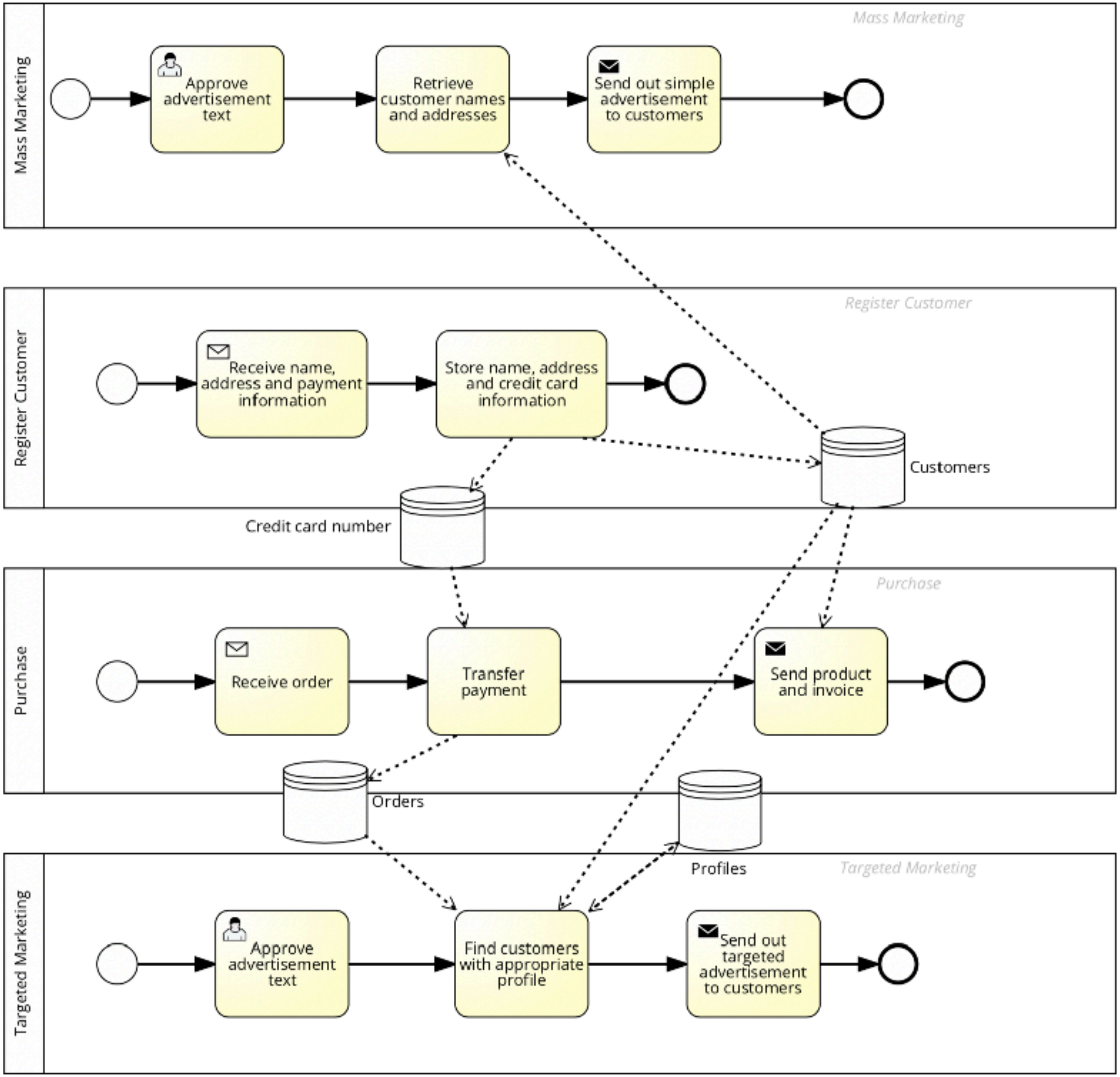
1. Implementation must conform to process model
2. Process model must conform to privacy policy
3. Process model must conform to the GDPR
4. Privacy policy must conform to the GDPR



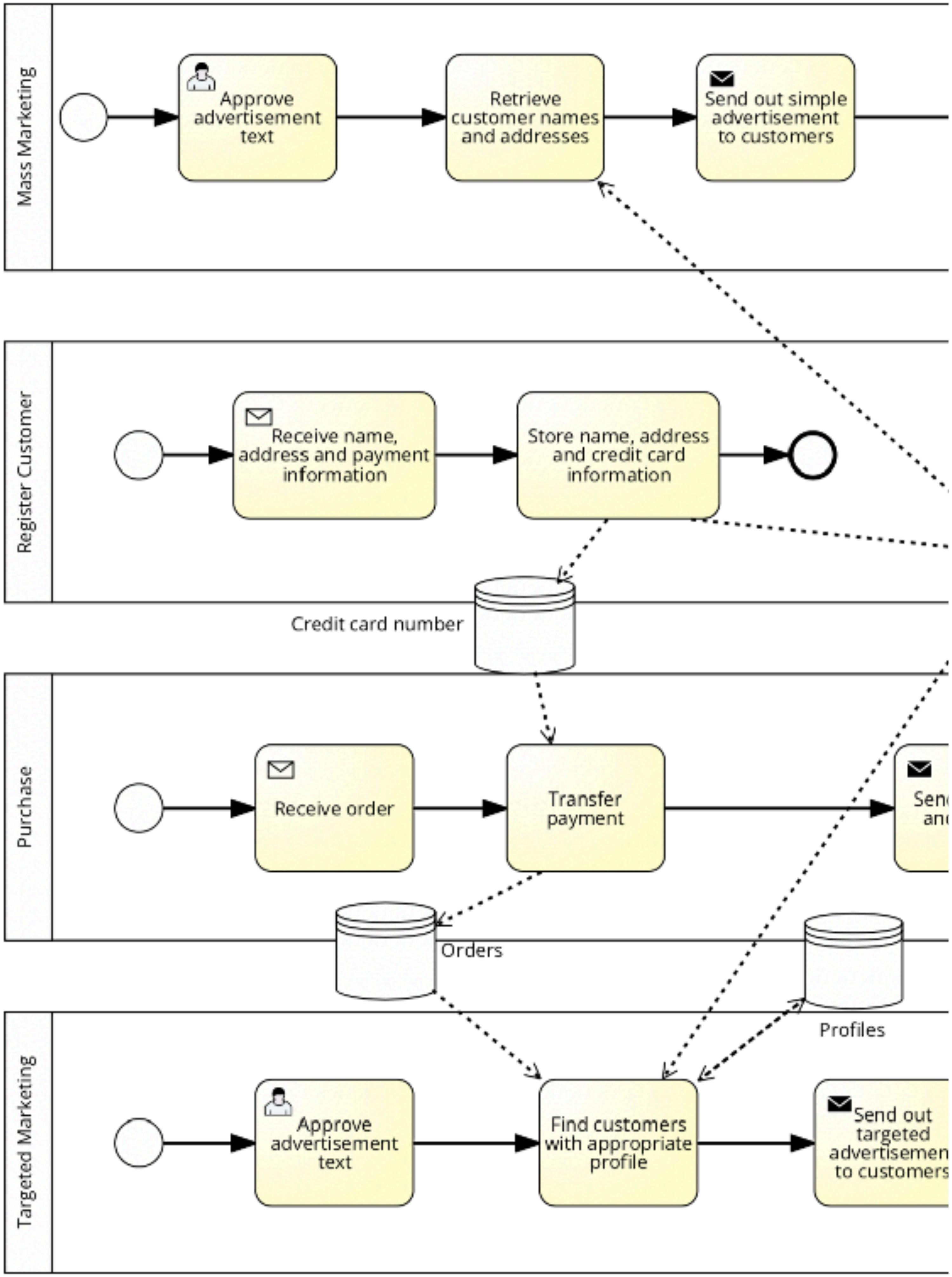
Where do we find process collections?

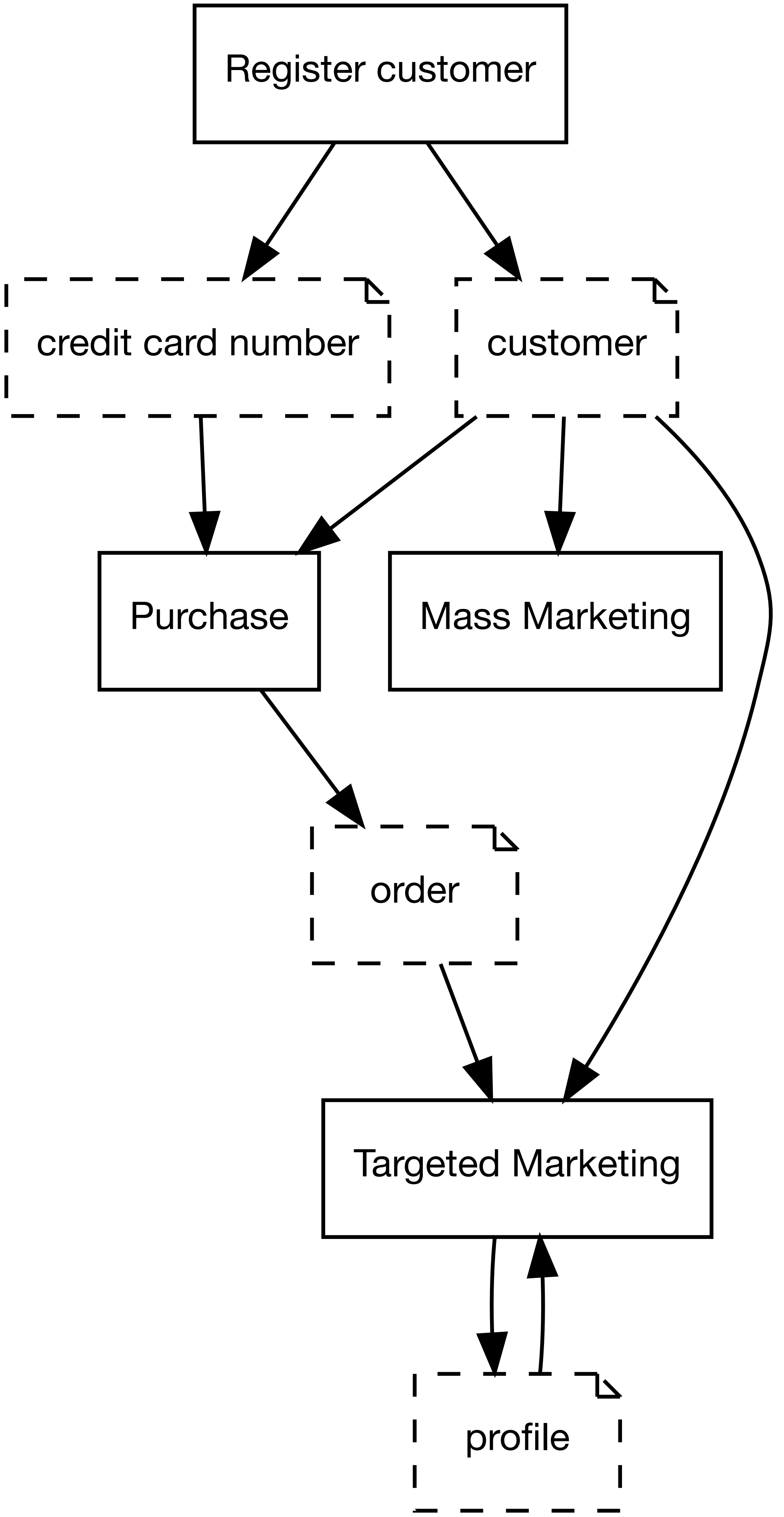
Requirements specifications, process models (BPMN), implementation artifacts

Retailer example

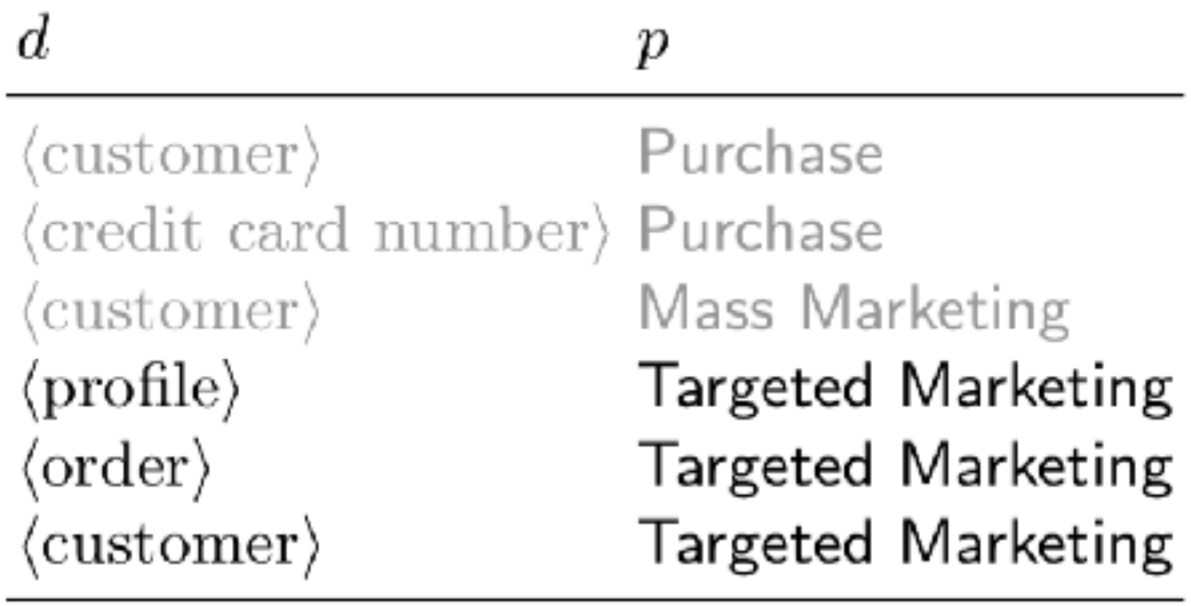

Mass Marketing
Mass Marketing
Approve advertisement text
Retrieve customer names and addresses
Send out simple advertisement to customers
Register Customer
Register Customer
Receive name, address and payment information
Store name, address and credit card information
Customers
Credit card number
Purchase
Purchase
Receive order
Transfer payment
Send product and invoice
Orders
Profiles
Targeted Marketing
Targeted Marketing
Approve advertisement text
Find customers with appropriate profile
Send out targeted advertisement to customers

BPMN


Mass Marketing
Approve advertisement text
Retrieve customer names and addresses
Send out simple advertisement to customers
Register Customer
Receive name, address and payment information
Store name, address and credit card information
Credit card number
Purchase
Receive order
Transfer payment
Send and
Orders
Profiles
Targeted Marketing
Approve advertisement text
Find customers with appropriate profile
Send out targeted advertisement to customers

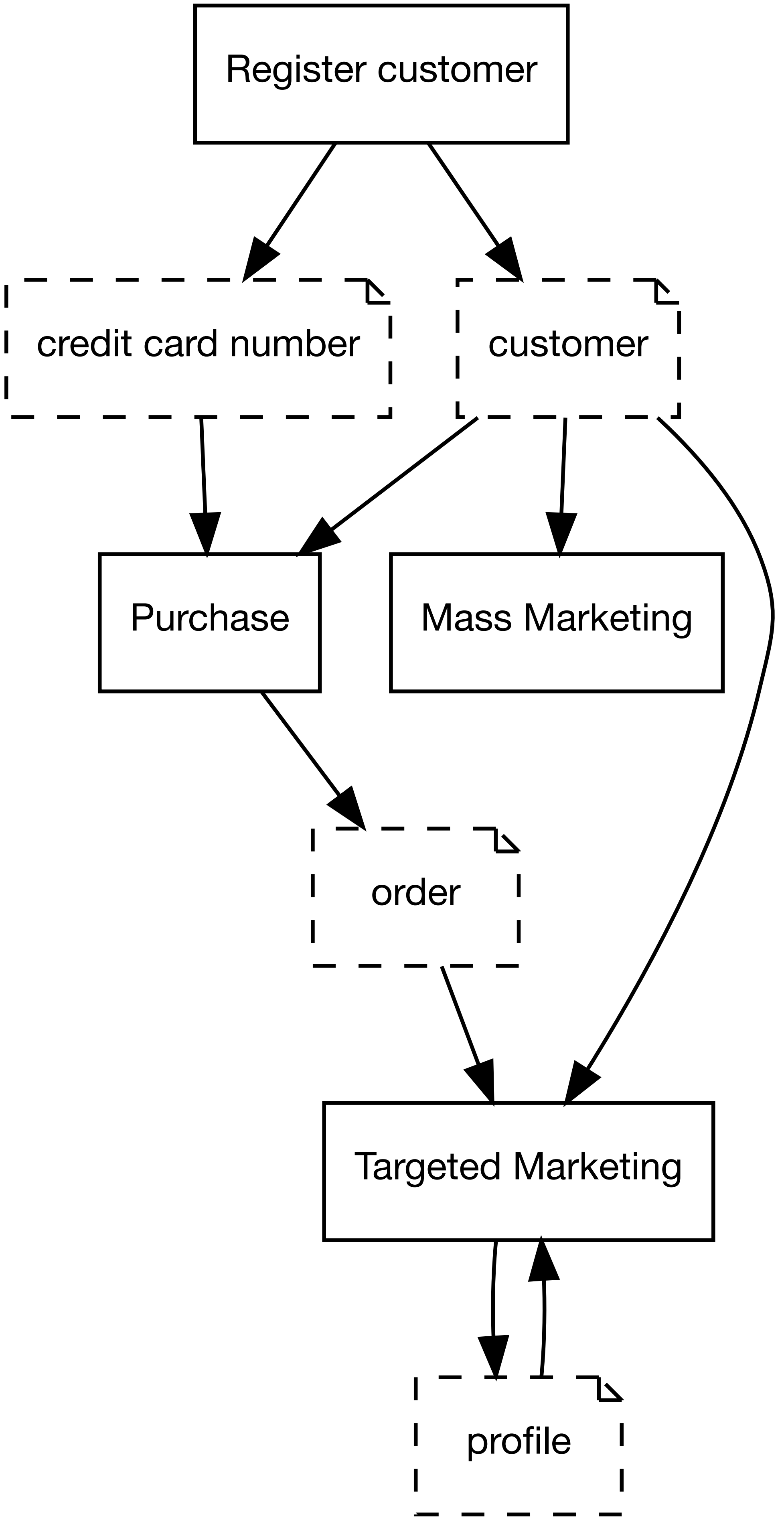

Register customer
credit card number
customer
Purchase
Mass Marketing
order
Targeted Marketing
profile

Formalising privacy policies



d	p
⟨customer⟩	Purchase
⟨credit card number⟩	Purchase
⟨customer⟩	Mass Marketing
⟨profile⟩	Targeted Marketing
⟨order⟩	Targeted Marketing
⟨customer⟩	Targeted Marketing

“We collect your customer information, order history, and profile, and use them to send you targeted advertising”

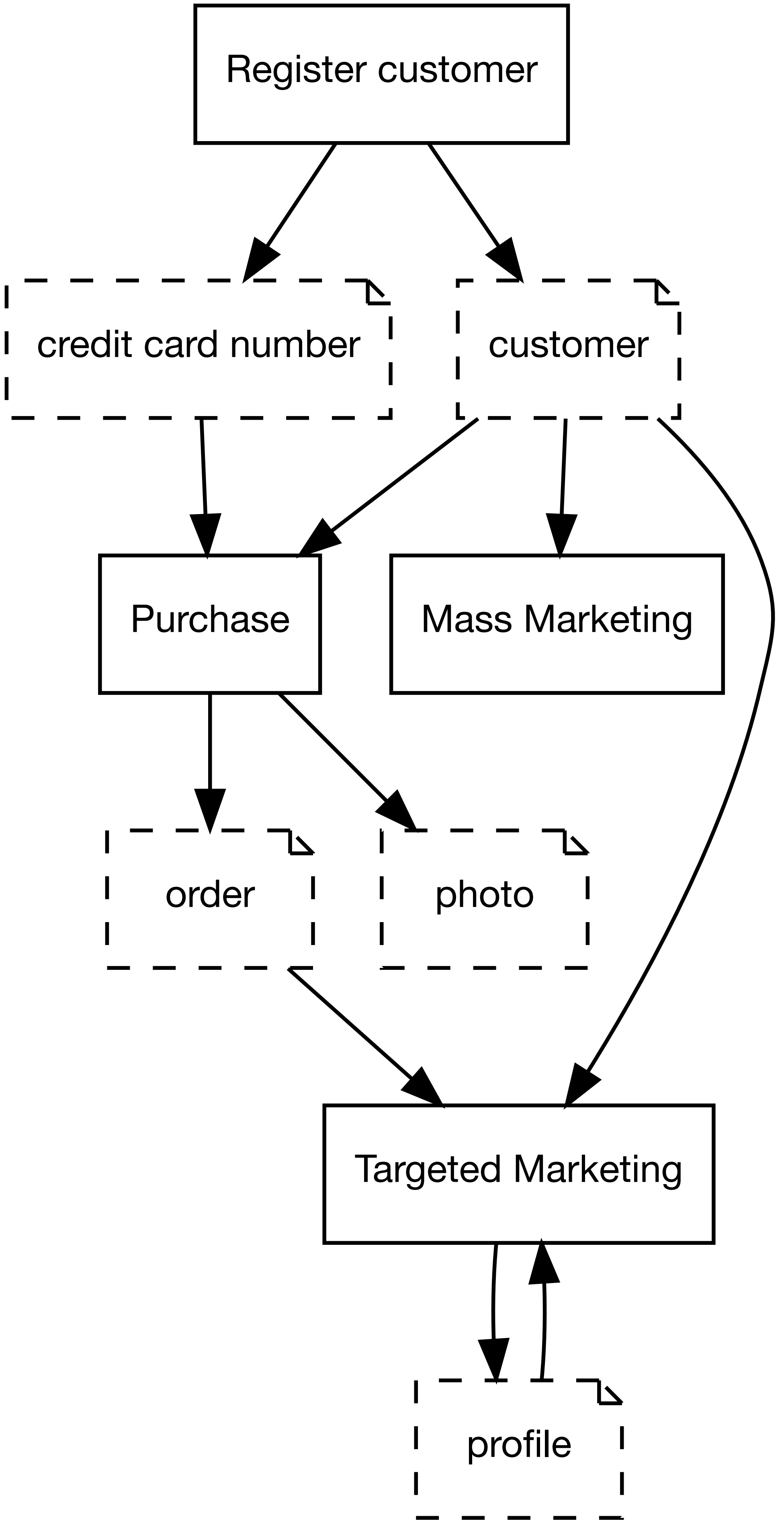


What’s wrong with this picture?

Purpose limitation:

Can’t collect “photo” during “Purchase” process/purpose without use.

May detect automatically



Related work

- Purpose-based access control, e.g.:

  Milan Petkovic, Davide Prandi, and Nicola Zannone. Purpose Control: Did You Process the Data for the Intended Purpose? In Secure Data Management, LNCS 6933, pp. 145–168. Springer, 2011.

- [Privacy-aware / Purpose-aware] role based access control, e.g.:
  Ni, Bertino, Lobo, Brodie, Karat, Karat, and Trombeta. Privacy-aware Role-based Access Control. ACM Transactions on Information System Security 13(3):24:1–24:31, July 2010.

Conclusion

- Regulations (GDPR) require consent to purposes
- How do we audit a computer system’s adherence to a purpose?
- Business Process Models, process model collections!
- Connecting the formal (process model, implementation) to the informal (purpose, privacy policy, unnecessary data)

Thank you!

Søren Debois, IT University of Copenhagen
Joint work with David Basin and Thomas Hildebrandt

FC ’18, Feb 26, 2018