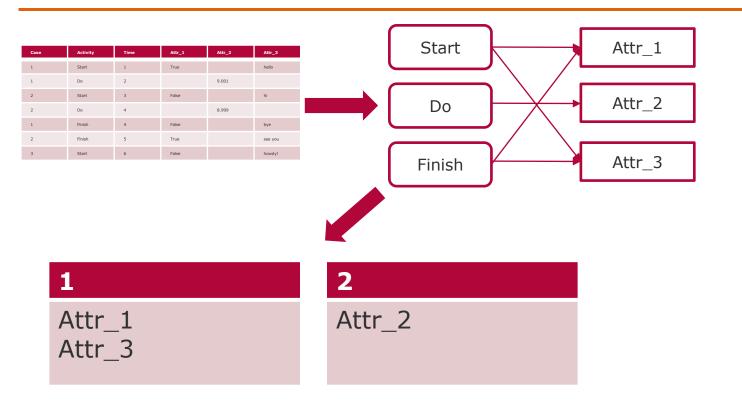


## Recap -Model Discovery on Single Event Log





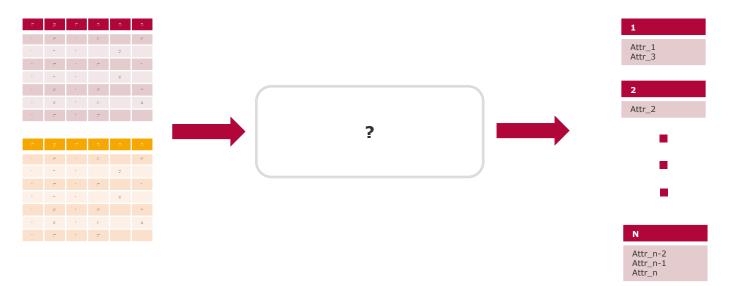
Bano, Dorina & Weske, Mathias. (2020). Discovering Data Models from Event Logs

> Discovering a Global Data Model from Event Logs

# Recap - Problem Definition



Multiple event logs => One data model



### Discovering a Global Data Model from Event Logs

# Recap – "Naive" Approach

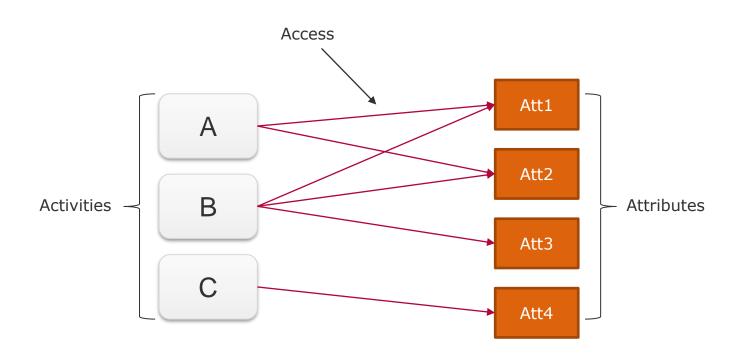




#### Discovering a Global Data Model from Event Logs

## A2A Diagrams - Example

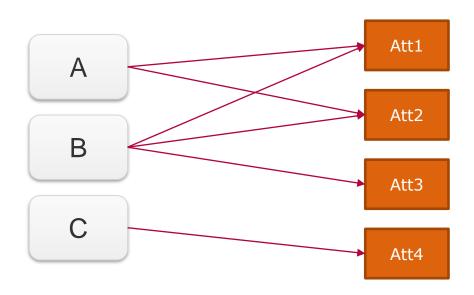




### Discovering a Global Data Model from Event Logs

# A2A Diagrams - Example

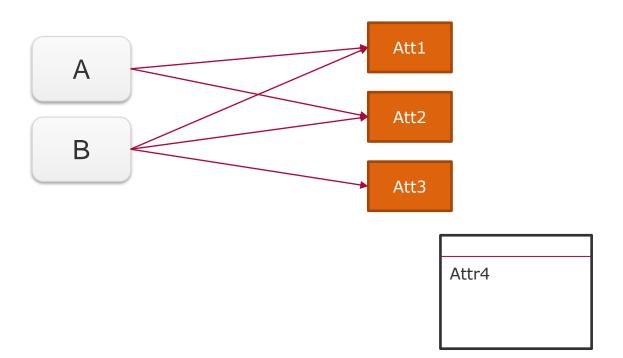




### Discovering a Global Data Model from Event Logs

# A2A Diagrams - Example

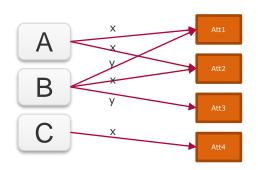




Discovering a Global Data Model from Event Logs

## A2A Diagrams – Data Structure



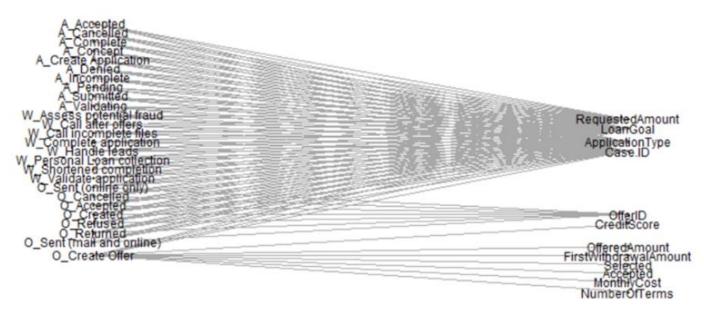


	Att1	Att2	Att3	Att4
Α	Χ	Χ	0	0
В	У	X	У	0
С	0	0	0	X

Discovering a Global Data Model from Event Logs







A2A diagram for one BPI challenge 2017 event log

#### Discovering a Global Data Model from Event Logs

## Intermission: Dataset – BPI Challenge 2017

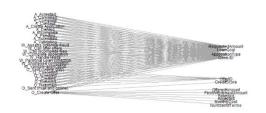


- Two event logs
  - OfferApplications: 1160405 events, 13 attributes
  - Offers: 193849 events, 11 attributes

Discovering a Global Data Model from Event Logs

## A2A Diagrams – Data Structure





•	Activity	Case.ID <sup>‡</sup>	LoanGoal <sup>‡</sup>	ApplicationType •	RequestedAmount <sup>0</sup>	FirstWithdrawalAmount	NumberOfTerms <sup>‡</sup>	Accepted <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount 0	OfferID 0
1	A_Accepted	31509	31509	31509	31509	0	0	0	0	0	0	0	0
2	A_Cancelled	10431	10431	10431	10431	0	0	0	0	0	0	0	0
3	A_Complete	31362	31362	31362	31362	0	0	0	0	0	0	0	0
4	A_Concept	31509	31509	31509	31509	0	0	0	0	0	0	0	0
5	A_Create Application	31509	31509	31509	31509	0	0	0	0	0	0	0	0
6	A_Denied	3753	3753	3753	3753	0	0	0	0	0	0	0	0
7	A_Incomplete	23055	23055	23055	23055	0	0	0	0	0	0	0	0
8	A_Pending	17228	17228	17228	17228	0	0	0	0	0	0	0	0
9	A_Submitted	20423	20423	20423	20423	0	0	0	0	0	0	0	0
10	A_Validating	38816	38816	38816	38816	0	0	0	0	0	0	0	0
11	O_Accepted	17228	17228	17228	17228	0	0	0	0	0	0	0	17228
12	O_Cancelled	20898	20898	20898	20898	0	0	0	0	0	0	0	20898
13	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created	42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused	4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned	23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail and online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess potential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0
20	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	0
21	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0
22	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	0
23	W_Handle leads	43771	43771	43771	43771	0	0	0	0	0	0	0	0
24	W_Personal Loan collection	22	22	22	22	0	0	0	0	0	0	0	0
25	W_Shortened completion	238	238	238	238	0	0	0	0	0	0	0	0
26	W_Validate application	193648	193648	193648	193648	0	0	0	0	0	0	0	0

### Discovering a Global Data Model from Event Logs

# Underlying Model -



	Which	Att	ribu	tes Be	long To	gether?						Plat Inst	itner titut
^	Activity	Case.ID ‡	LoanGoal <sup>‡</sup>	ApplicationType <sup>‡</sup>	RequestedAmount <sup>‡</sup>	FirstWithdrawalAmount	NumberOfTerms <sup>‡</sup>	Accepted <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>
1	A_Accepted	31509	31509	31509	31509	0	0	0	0	0	0	0	0
2	A_Cancelled	10431	10431	10431	10431	0	0	0	0	0	0	0	0
3	A_Complete	31362	31362	31362	31362	0	0	0	0	0	0	0	0
4	A_Concept	31509	31509	31509	31509	0	0	0	0	0	0	0	0
5	A_Create Application	31509	31509	31509	31509	0	0	0	0	0	0	0	0
6	A_Denied	3753	3753	3753	3753	0	0	0	0	0	0	0	0
7	A_Incomplete	23055	23055	23055	23055	0	0	0	0	0	0	0	0
8	A_Pending	17228	17228	17228	17228	0	0	0	0	0	0	0	0
9	A_Submitted	20423	20423	20423	20423	0	0	0	0	0	0	0	0
10	A_Validating	38816	38816	38816	38816	0	0	0	0	0	0	0	0

11	O_Accepted	17228	17228	17228	17228	0	
12	O_Cancelled	20898	20898	20898	20898	0	
13	O_Create Offer	42995	42995	42995	42995	42995	42
14	O_Created	42995	42995	42995	42995	0	
15	O_Refused		A Accepted				
16	O_Returned		A Cancelled A Complete A Concept	200			
17	O_Sent (mail and online)	A_C	A Denied Incomplete				
18	O_Sent (online only)	W Ass	A Submitted A Validating ess potential fraud			Designation of the last of the	dA mount
19	W_Assess potential fraud	W-Co	Call after offers Ill incomplete files oplete application			Loan Applical Cas	dAmount Goal ionType
20	W_Call after offers	1 Wy Pers	Handle leads onal Loan collection ordened completion			Cas	e.ID
21	W_Call incomplete files	1 '6_8	ent (online only) 2 Cancelled C Accepted				
22	W_Complete application	1	O Refused O Refused			Offe	erID Score
23	W_Handle leads	O_Ser	O_Returned it (mail and online) _Create Offer			Offered FirstWithdra	Amount ryalAmount
24	W_Personal Loan collection					Acce Month Number	oted VCost OfTerms
25	W_Shortened completion						
26	W_Validate application	193048	193048	193048	195048	U	

0	0
0	0
0	0
0	0
42995	42995
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0

# Underlying Model -

24 W\_Personal Loan collection

25 W\_Shortened completion

26 W\_Validate application



		hich A				long	Toge	ethei	?								tner titut
^	Activity		e.ID Ů LoanG	ioal <sup>‡</sup> Applic	:ationType <sup>‡</sup>	RequestedAmou	unt <sup>‡</sup> FirstWi	thdrawalAmoun	t <sup>‡</sup> N	umberOfTerms	Accepted	÷	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore	OfferedAmount	OfferID <sup>‡</sup>
1	Activity	Att1	Att2	Att3	Att4	Att5	Att6	Att7	R1			0	0	0	0	0	0
2												0	0	0	0	0	0
3	Α					0			P	4	Att1	0	0	0	0	0	0
4					-					'		0	0	0	0	0	0
5	В	0	0	0	0	×	0	0	1_		r	0	0	0	0	0	0
6	С					0			E	3	Att2	0	0	0	0	0	0
7										'		0	0	0	0	0	0

-	, · ·	1	I	l	l	ľ	I				i i	U V				
4	_								1			0	0	0	0	
5	В	0	0	0	0	X	0	0	Ш.,		[	0	0	0	0	
6	С					0				В	+ Att2	0	0	0	0	
7												0	0	0	0	
8	D					0			0		0	0	0	0	0	
9						0	D	la 1	0		0	0	0	0	0	
			ı	ı	ı	<b>I</b> 0	IRII									

0	C	- 1			0				Accz	U	U	0	U	0	0
7										0	0	0	0	0	0
8	D				0			0	0	0	0	0	0	0	0
9	_				0	D	1	0	0	0	0	0	0	0	0
10					0	Ru	ile T	0	0	0	0	0	0	0	0
11	O_Accepted	17228	17228	17228		17228		0	0	0	0	0	0	0	17228

7											0	0	0	0	0	
8	D					0		0	0	0	0	0	0	0	0	
9	<del>                                     </del>						Dula 1	0	0	0	0	0	0	0	0	
10	[ <sup>-</sup>					0	Ruje I	0	0	0	0	0	0	0	0	
11	O_Accepted		17228	17228	17228	17228		0	0	0	0	0	0	0	17228	
12	O_Cancelled		20898	20898	20898	20898		0	0	0	0	0	0	0	20898	
13	O_Create Offer	r	42995	42995	42995	42995	4	2995	42995	42995	42995	42995	42995	42995	0	
14	O_Created		42995	42995	42995	42995	;	0	0	0	0	0	0	0	42995	
15	O_Refused		4695	4695	4695	4695		0	0	0	0	0	0	0	4695	
16	O_Returned		23305	23305	23305	23305	;	0	0	0	0	0	0	0	23305	
17	O_Sent (mail a	nd online)	39707	39707	39707	39707	,	0	0	0	0	0	0	0	39707	
18	O_Sent (online	only)	2026	2026	2026	2026	i l	0	0	0	0	0	0	0	2026	
19	W_Assess pote	ential fraud	3000	3000	3000	3000	)	0	0	0	0	0	0	0	0	
20	W_Call after of	ffers	190750	190750	190750	190750		0	0	0	0	0	0	0	0	
21	W Call incomp	lata filar	165736	165736	165736	165736		٥	0	0	0	0	0	0	0	

13	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created	42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused	4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned	23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail and online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess potential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0
20	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	0
21	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0
22	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	0

13	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created	42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused	4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned	23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail and online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess potential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0
20	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	0
21	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0
22	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	0
23	W_Handle leads	43771	43771	43771	43771	0	0	0	0	0	0	0	0

•	O_/ lecepted			.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		•		_	_			_	
2	O_Cancelled	20898	20898	20898	20898	0	0	0	0	0	0	0	208
3	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	
4	O_Created	42995	42995	42995	42995	0	0	0	0	0	0	0	429
5	O_Refused	4695	4695	4695	4695	0	0	0	0	0	0	0	46
6	O_Returned	23305	23305	23305	23305	0	0	0	0	0	0	0	233
7	O_Sent (mail and online)	39707	39707	39707	39707	0	0	0	0	0	0	0	397
В	O_Sent (online only)	2026	2026	2026	2026	0	0	0	0	0	0	0	20
9	W_Assess potential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	
0	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	
1	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	
2	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	
3	W_Handle leads	43771	43771	43771	43771	0	0	0	0	0	0	0	

# Underlying Model – Which Attributes Belong Together?

24 W Personal Loan collection

25 W\_Shortened completion

26 W\_Validate application



	vity Case.ID CanGoal ApplicationType RequestedAmount NumberOfferms Accepted MonthlyCost Selected CreditScore OfferedAmount OfferlD														
Activity	÷	Case.ID ‡	LoanGoal	ApplicationType	RequestedAmo	ount <sup>‡</sup> First\	VithdrawalAmour	nt ‡	NumberOfTerms	Accepted	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>
Activity	Att1	Att2	Att3	3 Att4	Att5	Att6	Att7	R2		0	0		0	0	0
Δ					10			1	B	1+1					-
								1							0
В	0	0	0	0	x	0	0	Н						0	0
	0			0	<b> </b>	10	0	1			0			0	0
C	U	U		0	x	0	0		E		0	0	0	0	0
D	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0
_						Di	10 2	0	0	0	0	0	0	0	0
_					0	KU	ile Z	0	0	0	0	0	0	0	0
O_Accepted		17228	17228	17228		17228		0	0	0	0	0	0	0	17228
O_Cancelled		20898	20898	20898		20898		0	0	0	0	0	0	0	20898
O_Create Offer		42995	42995	42995		42995		42995	42995	42995	42995	42995	42995	42995	0
O_Created		42995	42995	42995		42995		0	0	0	0	0	0	0	42995
O_Refused		4695	4695	4695		4695		0	0	0	0	0	0	0	4695
O_Returned		23305	23305	23305		23305		0	0	0	0	0	0	0	23305
O_Sent (mail and	d online)	39707	39707	39707		39707		0	0	0	0	0	0	0	39707
O_Sent (online o	only)	2026	2026	2026		2026		0	0	0	0	0	0	0	2026
W_Assess poten	itial fraud	3000	3000	3000		3000		0	0	0	0	0	0	0	0
W_Call after offe	ers	190750	190750	190750		190750		0	0	0	0	0	0	0	0
W_Call incomple	ete files	165736	165736	165736		165736		0	0	0	0	0	0	0	0
W_Complete ap	plication	129796	129796	129796		129796		0	0	0	0	0	0	0	0
W_Handle leads	;	43771	43771	43771		43771		0	0	0	0	0	0	0	0
	A B C D E O_Accepted O_Cancelled O_Create Offer O_Refused O_Returned O_Sent (mail an O_Sent (online of the complete of the com	Activity Att1  A  B  O  C  O  D  O  E  O_Accepted O_Cancelled O_Create Offer O_Created O_Refused	Activity         Att1         Att2           A         0         0           B         0         0           C         0         0           D         0         0           E         0         0           Cancepted         17228           O_Cancelled         20898           O_Create Offer         42995           O_Refused         4695           O_Returned         23305           O_Sent (mail and online)         39707           O_Sent (online only)         2026           W_Assess potential fraud         3000           W_Call after offers         190750           W_Call incomplete files         165736           W_Complete application         129796	Activity         Att1         Att2         Att3           A         Image: Activity of the content o	Activity         Att1         Att2         Att3         ApplicationType           A         Att1         Att2         Att3         Att4           A         D         D         D         D         D           C         D         D         D         D         D         D           D <td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAme           Activity         Att1         Att2         Att3         Att4         Att5           A         0         0         0         0         0         0           B         0         0         0         0         0         x           C         0         0         0         0         x         0           D         0         0         0         0         x         0           E         0         <td< td=""><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         First Variable           Activity         Att1         Att2         Att3         Att4         Att5         Att6           A         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0           C         0         0         0         0         0         0         0           D         0         0         0         0         0         0         0         0           E         0_Accepted         17228         17228         17228         17228         17228         17228         17228         0         0         Refused         0_Cancelled         20898</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           A         0         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0         0         0           C         0</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           B         0         0         0         0         0         x         0         0           C         0         0         0         0         x         0         0         0           D         0         0         0         0         x         0</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7         R2         AXX         AXX         BB         O</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms         Accepted           Activity         Att1         Att2         Att3         Att44         Att5         Att6         Att7         R2         Att11         Att2         Att1         Att44         Att5         Att6         Att7         R2         Att11         0         &lt;</td><td>Activity</td><td>Activity</td><td>Activity</td><td>Activity Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 7 Att 1 Att 1 Att 7 Att 1 Att 1 Att 1 Att 1 Att 7 Att 1 Att</td></td<></td>	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAme           Activity         Att1         Att2         Att3         Att4         Att5           A         0         0         0         0         0         0           B         0         0         0         0         0         x           C         0         0         0         0         x         0           D         0         0         0         0         x         0           E         0 <td< td=""><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         First Variable           Activity         Att1         Att2         Att3         Att4         Att5         Att6           A         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0           C         0         0         0         0         0         0         0           D         0         0         0         0         0         0         0         0           E         0_Accepted         17228         17228         17228         17228         17228         17228         17228         0         0         Refused         0_Cancelled         20898</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           A         0         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0         0         0           C         0</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           B         0         0         0         0         0         x         0         0           C         0         0         0         0         x         0         0         0           D         0         0         0         0         x         0</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7         R2         AXX         AXX         BB         O</td><td>Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms         Accepted           Activity         Att1         Att2         Att3         Att44         Att5         Att6         Att7         R2         Att11         Att2         Att1         Att44         Att5         Att6         Att7         R2         Att11         0         &lt;</td><td>Activity</td><td>Activity</td><td>Activity</td><td>Activity Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 7 Att 1 Att 1 Att 7 Att 1 Att 1 Att 1 Att 1 Att 7 Att 1 Att</td></td<>	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         First Variable           Activity         Att1         Att2         Att3         Att4         Att5         Att6           A         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0           C         0         0         0         0         0         0         0           D         0         0         0         0         0         0         0         0           E         0_Accepted         17228         17228         17228         17228         17228         17228         17228         0         0         Refused         0_Cancelled         20898	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           A         0         0         0         0         0         0         0         0           B         0         0         0         0         0         0         0         0         0           C         0	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7           B         0         0         0         0         0         x         0         0           C         0         0         0         0         x         0         0         0           D         0         0         0         0         x         0	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms           Activity         Att1         Att2         Att3         Att4         Att5         Att6         Att7         R2         AXX         AXX         BB         O	Activity         Case.ID         LoanGoal         ApplicationType         RequestedAmount         FirstWithdrawalAmount         NumberOfTerms         Accepted           Activity         Att1         Att2         Att3         Att44         Att5         Att6         Att7         R2         Att11         Att2         Att1         Att44         Att5         Att6         Att7         R2         Att11         0         <	Activity	Activity	Activity	Activity Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 1 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 1 Att 2 Att 3 Att 4 Att 5 Att 6 Att 7 Att 7 Att 1 Att 7 Att 1 Att 1 Att 7 Att 1 Att 1 Att 1 Att 1 Att 7 Att 1 Att

# Underlying Model – Which Attributes Belong Together?

26 W\_Validate application



		ty Case.ID Cas														
^	Activity	÷	Case.ID Date	anGoal ‡	ApplicationType	RequestedAmount	t <sup>‡</sup> FirstW	ithdrawalAmou	nt <sup>‡</sup>	NumberOfTerms	Accepted	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount	OfferID <sup>‡</sup>
1	Activity	Att1	Att2	Att3	Att4	Att5	Att6	Att7	R3		107	0	0	0	0	0
2									-	× Att1	REA!		0	0	0	0
3	Α			0	0	0				x Att2			0	0	0	0
4	В	0	0	×	×	х	0	0	11 0	A Att3			0	0	0	0
5			<u> </u>	<u> </u>	^			Ů	4	z Att	1    "			0	0	0
6	С			0	0	0			Н	Att			-		0	0
7	D			0	0	0			1 .	Name and Advanced			-	0	0	0
8	D					0			0	0				0	0	0
9	Е			0	0	0	Ru	le 3	0	0	0	0	-	0	0	0
10	0.4		17000	17220	17220				0	0	0	0	_	0	0	0
	O_Accepted		17228	17228	17228		7228		0	0	0	0		0	0	.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
	O_Cancelled		20898	20898	20898		0898		0	0	0			0	42005	20898
	O_Create Offer	r	42995	42995	42995		2995		42995	42995	42995	42995			42995	
	O_Created		42995	42995	42995		2995 4695		0	0	0			0	0	42995
	O_Refused		4695	4695 23305	4695 23305		3305		0	0	0			0	0	4695
	O_Returned		23305 39707	39707	39707		9707		0	0	0	0		0	0	23305 39707
	O_Sent (mail a						2026		0					0	0	
	O_Sent (online		2026	2026	2026				0	0	0			•	0	2026
	W_Assess pote		3000 190750	3000 190750	3000 190750		3000 0750		0	0	0	0		0	0	0
	W_Call after of		165736	165736	165736		5736		0	0	0	0		0	0	0
	W_Call incomp		129796	129796	129796		9796		0	0	0	-		0	0	0
	W_Complete a		43771	43771	43771		9796 3771			0				0	0	0
	W_Handle lead					4:			0	0	0	0	0	0	0	0
	W_Personal Lo		22	22	22		22			0	0	-			0	0
25	W_Shortened	completion	238	238	238		238		0	0	0	0	0	0	0	0

# Underlying Model -

20 W\_Call after offers

23 W\_Handle leads

21 W\_Call incomplete files

22 W\_Complete application

24 W\_Personal Loan collection

25 W\_Shortened completion

26 W\_Validate application



	WI	hich i	Attri	bute	s Be	long	Toge	ethe	r?							ttner titut
^	Activity	Cas	e.ID Ů LoanG	ioal <sup>‡</sup> Applic	ationType <sup>‡</sup>	RequestedAmou	nt <sup>‡</sup> FirstWi	thdrawalAmoun	t ‡	NumberOfTerms   Acc	epted	MonthlyCost	\$ Selected	CreditScore	OfferedAmount	OfferID <sup>‡</sup>
1	Activity	Att1	Att2	Att3	Att4	Att5	Att6	Att7	R4			0	0 (	) (	0	0
2										A X Att1		0	0 (	) (	0	0
3	Α	Z	0	0	m	n	Z	0		X X		0	0 (	) (	0	0
4	_									B V Att2		0	0	) (	0	0
5	В	У	У	X	X	X	0	Z				0	0 (	) (	0	0
6	С	0	х	0	z	0	У	х	1	C 7 2		0	0 (	) (	0	0
7					l	I	·					0	0			0

											_		_	_
4	Б			,,	, , , , , , , , , , , , , , , , , , ,	,,			B V Att2	0	0	0	0	0
5	В	У	У	X	X	X	0	Z		0	0	0	0	0
6	С	0	х	0	z	0	У	х		0	0	0	0	0
7							,			0	0	0	0	0
8	D	х	0	0	0	У	Z	0	0 0	0	0	0	0	0
o l									Hoo	0	0	0	0	0
	_	l 0	.,	l 0	l .,	I 0		1 1				•	•	•

6	С	0	х	0	z	0	У	х		-	0	0	0	0	0	0
7									V	U	0	0	0	0	0	0
8	D	X	0	0	0	У	z	0	0	0	0	0	0	0	0	0
9	_	0	1,,	0		0	o Du	10 1	0	0	0	0	0	0	0	0
10	L	U	y	U U	^	"	<sup>°</sup> Ru	16 4	0	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228		17228		0	0	0	0	0	0	0	17228

- 1																
8	D	Х	0	0	0	y z		0	0	0	0	0	0	0	0	0
9	_	0	.,				Dii	10 1	0	0	0	0	0	0	0	0
10	_	U	У		x		Ru	18 4	0	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228	1722	28		0	0	0	0	0	0	0	17228
12	O_Cancelled		20898	20898	20898	2089	98		0	0	0	0	0	0	0	20898
13	O_Create Offer		42995	42995	42995	4299	95	4	2995	42995	42995	42995	42995	42995	42995	0
14	O_Created		42995	42995	42995	4299	95		0	0	0	0	0	0	0	42995
15	O Refused		4695	4695	4695	469	95		0	0	0	0	0	0	0	4695

0		_ ^	ľ	ľ	ľ	1 '   -	l v L v	0	U	U	U	U	U	U
9	_	0					Dula 4	0	0	0	0	0	0	0
10	_	U	У	0	×		Rule 4	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228	17228	0	0	0	0	0	0	0	17228
12	O_Cancelled		20898	20898	20898	20898	0	0	0	0	0	0	0	20898
13	O_Create Offer	r	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created		42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused		4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned		23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail a	nd online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online	only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess pote	ential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0

8	l D	X	0	0	0	y z		0	0	0	0	0	0	0	0
9	E		<del> </del>				Dula 1	0	0	0	0	0	0	0	0
10		U	У	0	X		Ruje 4	0	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228	17228	8	0	0	0	0	0	0	0	17228
12	O_Cancelled		20898	20898	20898	20898		0 42995		0	0	0	0	0	20898
13	O_Create Offer		42995	42995	42995	42995	42	2995	42995	42995	42995	42995	42995	42995	0
14	O_Created		42995	42995	42995	42995		0	0	0	0	0	0	0	42995
15	O_Refused		4695	4695	4695	4695		0	0	0	0	0	0	0	4695
16	O_Returned		23305	23305	23305	23305		0	0	0	0	0	0	0	23305
17	O_Sent (mail ar	nd online)	39707	39707	39707	39707	,	0	0	0	0	0	0	0	39707
18	O_Sent (online	only)	2026	2026	2026	2026	i e	0	0	0	0	0	0	0	2026
19	W Assess note	ntial fraud	3000	3000	3000	3000		0	0	0	0	0	0	0	0

# Underlying Model – Which Attributes Belong Together?

26 W\_Validate application



	Wl	nich	Attr	ibut	tes Be	long	y Tog	ethei	?							titut
^	Activity	\$	Case.ID <sup>‡</sup> Lo	oanGoal <sup>‡</sup>	ApplicationType <sup>‡</sup>	RequestedAn	nount <sup>‡</sup> FirstW	/ithdrawalAmoun	t <sup>‡</sup>	NumberOfTerms <sup>‡</sup>	Accepted <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>
1	Activity	Att1	Att2	Att3	Att4	Att5	Att6	Att7		First Fragment	0	0	0	0	0	0
2						_			f1		0	0	0	0	0	0
3	Α	z	0	0	m	n	z	0		A X Att1	0	0	0	0	0	0
4	D.			-						A Att1	0	0	0	0	0	0
5	В	У	У	×	×	X	0	Z		Att2	0	0	0	0	0	0
6	С	0	X	0	Z	0	V	×		Attz	0	0	0	0	0	0
7							,			5	0	0	0	0	0	0
8	D	х	0	0	0	У	Z	0	0	0	0	0	0	0	0	0
9	_	0	.,	0		0	0 D.	10 1	0	0	0	0	0	0	0	0
10		U	У	1 0	×	0	KU	ile 4	0	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228		17228		0	0	0	0	0	0	0	17228
12	O_Cancelled		20898	20898	20898		20898		0	0	0	0	0	0	0	20898
13	O_Create Offer		42995	42995	42995		42995	4	2995	42995	42995	42995	42995	42995	42995	0
14	O_Created		42995	42995	42995		42995		0	0	0	0	0	0	0	42995
15	O_Refused		4695	4695	4695		4695		0	0	0	0	0	0	0	4695
16	O_Returned		23305	23305	23305		23305		0	0	0	0	0	0	0	23305

					- 1									
9	F	0	V	0	×	0 0	Rule 4	0	0	0	0	0	0	0
10	_	Ü	y	Ŭ	^		Rule 4	0	0	0	0	0	0	0
11	O_Accepted		17228	17228	17228	17228	0	0	0	0	0	0	0	17228
12	O_Cancelled		20898	20898	20898	20898	0	0	0	0	0	0	0	20898
13	O_Create Offer	r	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created		42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused		4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned		23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail a	nd online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online	only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess pote	ential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0
20	W_Call after of	fers	190750	190750	190750	190750	0	0	0	0	0	0	0	0
21	W_Call incomp	lete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0
22	W_Complete a	pplication	129796	129796	129796	129796	0	0	0	0	0	0	0	0
23	W_Handle lead	İs	43771	43771	43771	43771	0	0	0	0	0	0	0	0
24	W_Personal Lo	an collection	22	22	22	22	0	0	0	0	0	0	0	0
25	W_Shortened	completion	238	238	238	238	0	0	0	0	0	0	0	0

# Underlying Model -



	W	hich i	Attri	bute	s Be	long	Tog	ether	?						ttner titut
*	Activity		e.ID <sup>‡</sup> Loan(	Goal <sup>‡</sup> Applic	ationType <sup>‡</sup>	RequestedAmou	ınt <sup>‡</sup> FirstW	/ithdrawalAmount	NumberOfTerms	Accepted	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>
1	Activity	Att1	Att2	Att3	Att4	Att5	Att6	Att7	Second Fragment	0 0	0	0	0	0	0
2									f2	о о	0	0	0	0	0
3	А	Z	0	0	m	n	Z	0	Att1	0 0	0	0	0	0	0
4						-		_	/	0 0	0	0	0	0	0
5	В	У	У	X	X	X	0	Z	B /y Att2	0 0	0	0	0	0	0
6	C	0	X	0	7	0	V	×		0 0	0	0	0	0	0

А	Z	0	0	m	n	Z	0	Н
В	У	У	х	х	х	0	Z	
С	0	Х	0	Z	0	У	Х	
D	Х	0	0	0	У	Z	0	0
Е	0	У	0	X	0	Ru	le 4	0

A	Z	U	U	m	n	Z	U
В	у	У	×	×	×	0	z
С	0	х	0	Z	0	У	х
D	х	0	0	0	У	Z	0
							I - A

11 O\_Accepted

12 O Cancelled

14 O\_Created

15 O\_Refused

16 O\_Returned

17 O\_Sent (mail and online)

19 W\_Assess potential fraud

21 W\_Call incomplete files

22 W\_Complete application

24 W\_Personal Loan collection

25 W\_Shortened completion

26 W\_Validate application

18 O\_Sent (online only)

20 W\_Call after offers

23 W\_Handle leads

13 O\_Create Offer

asso attner istitut

HPI	Hasso Plattner Institut
 ÷ 04414-	÷ 044ID

## Offers - A2A Matrix



^	Activity	Case.ID <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	ApplicationID +	FirstWithdrawalAmount +	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	NumberOfTerms <sup>‡</sup>	Accepted <sup>‡</sup>	OfferID <sup>‡</sup>
1	O_Accepted	17228	17228	17228	17228	17228	17228	17228	17228	17228	17228
2	O_Cancelled	20898	20898	20898	20898	20898	20898	20898	20898	20898	20898
3	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
4	O_Created	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995
5	O_Refused	4695	4695	4695	4695	4695	4695	4695	4695	4695	4695
6	O_Returned	23305	23305	23305	23305	23305	23305	23305	23305	23305	23305
7	O_Sent (mail and online)	39707	39707	39707	39707	39707	39707	39707	39707	39707	39707
8	O_Sent (online only)	2026	2026	2026	2026	2026	2026	2026	2026	2026	2026

### Discovering a Global Data Model from Event Logs



# OfferApplications – A2A Matrix

•	Activity	Case.ID <sup>‡</sup>	LoanGoal <sup>‡</sup>	ApplicationType	RequestedAmount	FirstWithdrawalAmount <sup>‡</sup>	NumberOfTerms	Accepted <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected	CreditScore <sup>‡</sup>	OfferedAmount +	OfferID <sup>‡</sup>
1	A_Accepted	31509	31509	31509	31509	0	0	0	0	0	0	0	0
2	A_Cancelled	10431	10431	10431	10431	0	0	0	0	0	0	0	0
3	A_Complete	31362	31362	31362	31362	0	0	0	0	0	0	0	0
4	A_Concept	31509	31509	31509	31509	0	0	0	0	0	0	0	0
5	A_Create Application	31509	31509	31509	31509	0	0	0	0	0	0	0	0
6	A_Denied	3753	3753	3753	3753	0	0	0	0	0	0	0	0
7	A_Incomplete	23055	23055	23055	23055	0	0	0	0	0	0	0	0
8	A_Pending	17228	17228	17228	17228	0	0	0	0	0	0	0	0
9	A_Submitted	20423	20423	20423	20423	0	0	0	0	0	0	0	0
10	A_Validating	38816	38816	38816	38816	0	0	0	0	0	0	0	0
11	O_Accepted	17228	17228	17228	17228	0	0	0	0	0	0	0	17228
12	O_Cancelled	20898	20898	20898	20898	0	0	0	0	0	0	0	20898
13	O_Create Offer	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	0
14	O_Created	42995	42995	42995	42995	0	0	0	0	0	0	0	42995
15	O_Refused	4695	4695	4695	4695	0	0	0	0	0	0	0	4695
16	O_Returned	23305	23305	23305	23305	0	0	0	0	0	0	0	23305
17	O_Sent (mail and online)	39707	39707	39707	39707	0	0	0	0	0	0	0	39707
18	O_Sent (online only)	2026	2026	2026	2026	0	0	0	0	0	0	0	2026
19	W_Assess potential fraud	3000	3000	3000	3000	0	0	0	0	0	0	0	0
20	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	0
21	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0
22	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	0
23	W_Handle leads	43771	43771	43771	43771	0	0	0	0	0	0	0	0
24	W_Personal Loan collection	22	22	22	22	0	0	0	0	0	0	0	0
25	W_Shortened completion	238	238	238	238	0	0	0	0	0	0	0	0
26	W_Validate application	193648	193648	193648	193648	0	0	0	0	0	0	0	0



## Union(OfferApplications, Offers) – A2A Matrix

^	Activity	Case.ID <sup>‡</sup>	LoanGoal <sup>‡</sup>	ApplicationType <sup>‡</sup>	RequestedAmount +	FirstWithdrawalAmount <sup>‡</sup>	NumberOfTerms	Accepted	MonthlyCost	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>	ApplicationID <sup>‡</sup>	
1	A_Accepted	31509	31509	31509	31509	0	0	(	0	0	0	0	0	0	
2	A_Cancelled	10431	10431	10431	10431	0	0	(	0	0	0	0	0	0	
3	A_Complete	31362	31362	31362	31362	0	0	(	0	0	0	0	0	0	
4	A_Concept	31509	31509	31509	31509	0	0	(	0	0	0	0	0	0	
5	A_Create Application	31509	31509	31509	31509	0	0	(	0	0	0	0	0	0	
6	A_Denied	3753	3753	3753	3753	0	0	(	0	0	0	0	0	0	
7	A_Incomplete	23055	23055	23055	23055	0	0	(	0	0	0	0	0	0	
8	A_Pending	17228	17228	17228	17228	0	0	0	0	0	0	0	0	0	
9	A_Submitted	20423	20423	20423	20423	0	0	(	0	0	0	0	0	0	
10	A_Validating	38816	38816	38816	38816	0	0	0	0	0	0	0	0	0	
11	O_Accepted	34456	17228	17228	17228	17228	17228	17228	17228	17228	17228	17228	34456	17228	
12	O_Cancelled	41796	20898	20898	20898	20898	20898	20898	20898	20898	20898	20898	41796	20898	
13	O_Create Offer	85990	42995	42995	42995	85990	85990	85990	85990	85990	85990	85990	0	42995	
14	O_Created	85990	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	85990	42995	
15	O_Refused	9390	4695	4695	4695	4695	4695	4695	4695	4695	4695	4695	9390	4695	
16	O_Returned	46610	23305	23305	23305	23305	23305	23305	23305	23305	23305	23305	46610	23305	
17	O_Sent (mail and online)	79414	39707	39707	39707	39707	39707	39707	39707	39707	39707	39707	79414	39707	
18	O_Sent (online only)	4052	2026	2026	2026	2026	2026	2026	2026	2026	2026	2026	4052	2026	
19	W_Assess potential fraud	3000	3000	3000	3000	0	0	(	0	0	0	0	0	0	
20	W_Call after offers	190750	190750	190750	190750	0	0	0	0	0	0	0	0	0	
21	W_Call incomplete files	165736	165736	165736	165736	0	0	0	0	0	0	0	0	0	
22	W_Complete application	129796	129796	129796	129796	0	0	0	0	0	0	0	0	0	
23	W_Handle leads	43771	43771	43771	43771	0	0	0	0	0	0	0	0	0	
24	W_Personal Loan collection	22	22	22	22	0	0	(	0	0	0	0	0	0	
25	W_Shortened completion	238	238	238	238	0	0	0	0	0	0	0	0	0	
26	W_Validate application	193648	193648	193648	193648	0	0	0	0	0	0	0	0	0	





"Each set of cells accessed by multiple event logs is treated as representing the same data model class."

^	Activity	Ca	se.ID <sup>‡</sup>	LoanGoal <sup>‡</sup>	ApplicationType <sup>‡</sup>	RequestedAmount <sup>‡</sup>	FirstWithdrawalAmount $^{\ddagger}$	NumberOfTerms <sup>‡</sup>	Accepted <sup>‡</sup>	MonthlyCost <sup>‡</sup>	Selected <sup>‡</sup>	CreditScore <sup>‡</sup>	OfferedAmount <sup>‡</sup>	OfferID <sup>‡</sup>	ApplicationID <sup>‡</sup>
11	O_Accepted	$\overline{}$	34456	17228	17228	17228	17228	17228	17228	17228	17228	17228	17228	34456	17228
12	O_Cancelled		41796	20898	20898	20898	20898	20898	20898	20898	20898	20898	20898	41796	20898
13	O_Create Offer		85990	42995	42995	42995	85990	85990	85990	85990	85990	85990	85990	0	42995
14	O_Created		85990	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	85990	42995
15	O_Refused		9390	4695	4695	4695	4695	4695	4695	4695	4695	4695	4695	9390	4695
16	O_Returned		46610	23305	23305	23305	23305	23305	23305	23305	23305	23305	23305	46610	23305
17	O_Sent (mail and online)		79414	39707	39707	39707	39707	39707	39707	39707	39707	39707	39707	79414	39707
18	O_Sent (online only)		4052	2026	2026	2026	2026	2026	2026	2026	2026	2026	2026	4052	2026

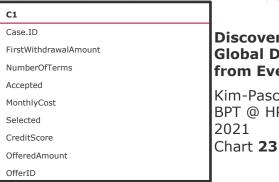
### Discovering a Global Data Model from Event Logs





"Each set of cells accessed by multiple event logs is treated as accessed by the same data model class."

^	Activity	Case.ID	LoanGoal	ApplicationType	RequestedAmount	FirstWithdrawalAmount	NumberOfTerms	Accepted	MonthlyCost	Selected	CreditScore	OfferedAmount **	OfferID	ApplicationID
11	O_Accepted	34456	17228	17228	17228	17228	17228	17228	17228	17228	17228	17228	34456	17228
12	O_Cancelled	41796	20898	20898	20898	20898	20898	20898	20898	20898	20898	20898	41796	20898
13	O_Create Offer	85990	42995	42995	42995	85990	85990	85990	85990	85990	85990	85990	0	42995
14	O_Created	85990	42995	42995	42995	42995	42995	42995	42995	42995	42995	42995	85990	42995
15	O_Refused	9390	4695	4695	4695	4695	4695	4695	4695	4695	4695	4695	9390	4695
16	O_Returned	46610	23305	23305	23305	23305	23305	23305	23305	23305	23305	23305	46610	23305
17	O_Sent (mail and online)	79414	39707	39707	39707	39707	39707	39707	39707	39707	39707	39707	79414	39707
18	O_Sent (online only)	4052	2026	2026	2026	2026	2026	2026	2026	2026	2026	2026	4052	2026



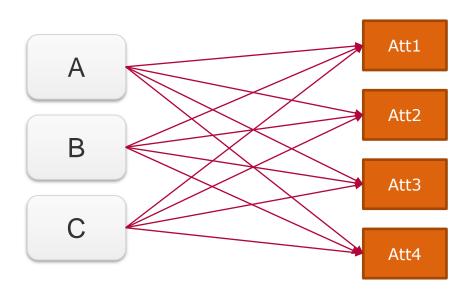
Discovering a **Global Data Model** from Event Logs Kim-Pascal Borchart

BPT @ HPI Summer





"A2A diagram/matrix in which all activities access all attributes"



	Att1	Att2	Att3
Α	X	У	Z
В	X	У	X
С	X	X	У

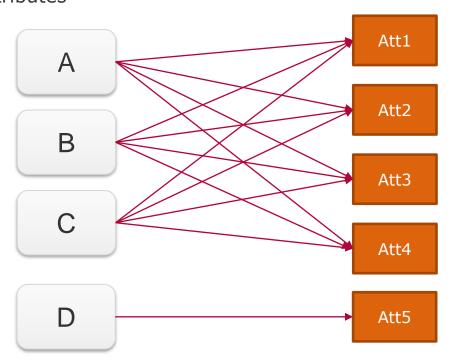
### Discovering a Global Data Model from Event Logs





## New Definition: Partially Fully Connected A2A

"A2A diagrams / matrices in which a subset of activities always accesses a subset of attributes"



	Att1	Att2	Att3	Att4
Α	X	У	Z	0
В	X	X	X	0
С	X	X	X	0
D	0	0	0	У

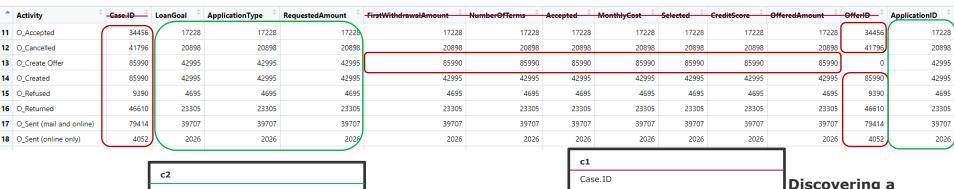
### Discovering a Global Data Model from Event Logs



### HPI Hasso Plattner Institut

## New Rule: (Partially) Fully Connected A2A

 Attributes belonging to partially fully connected a2a subsets are treated as belonging to one data class



c2

LoanGoal

ApplicationType

RequestedAmount

ApplicationId

c1

Case.ID

FirstWithdrawalAmount

NumberOfTerms

Accepted

MonthlyCost

Selected

CreditScore

OfferedAmount

Discovering a Global Data Model from Event Logs

Kim-Pascal Borchart BPT @ HPI Summer 2021

Chart 26

## BPI 2017 Challenge Data Definition vs. Result

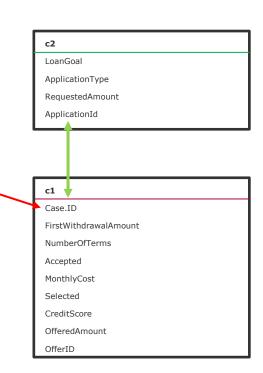


For all applications, the following data is available:

- · Requested load amount (in Euro),
- The application type,
- . The reason the loan was applied for (LoanGoal), and
- An application ID.

For all offers, the following data is available:

- · An offer ID,
- · The offered amount,
- · The initial withdrawal amount,
- · The number of payback terms agreed to,
- · The monthly costs,
- The creditscore of the customer,
- · The employee who created the offer,
- · Whether the offer was selected, and
- · Whether the offer was accepted by the customer.



https://www.win.tue .nl/bpi/doku.php?id= 2017:challenge

Discovering a Global Data Model from Event Logs





c1	]	c2
Case.ID		LoanGoal
FirstWithdrawalAmount		ApplicationType
NumberOfTerms		RequestedAmount
Accepted		ApplicationId
MonthlyCost	'	
Selected		
CreditScore		
OfferedAmount		
OfferID		

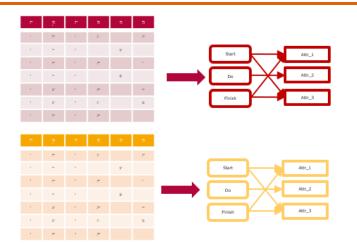
In performance experiments on five diverse datasets, our algorithm reaches an average recall of 88% and 91% in primary keys and foreign keys discovery, respectively. We show with an experiment that without knowledge of primary keys (which is assumed in related work), the performance of foreign key discovery is much worse, indicating the necessity to discover primary keys in advance or simultaneously. We compared precision and recall with the state-of-art algorithms.

Jiang, L., Naumann, F. Holistic primary key and foreign key detection. *J Intell Inf Syst* **54**, 439–461 (2020)

### Discovering a Global Data Model from Event Logs



- A2A for all EL
- Multilog-Rule + Combine (Union)
- Fully Connected Rule
- Rules 1-4 as normal



### Discovering a Global Data Model from Event Logs



- A2A for all EL
- Multilog-Rule + Combine (Union)
- Fully Connected Rule
- Rules 1-4 as normal



Case.ID

FirstWithdrawalAmount

NumberOfferms

Accepted

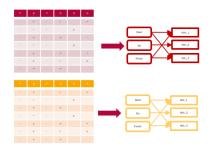
MonthlyCost

Selected

CreditScore

OfferdAmount

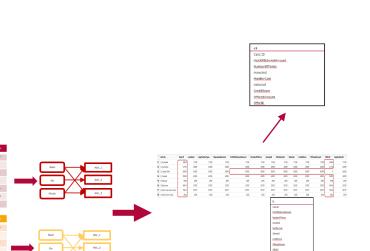
OfferID

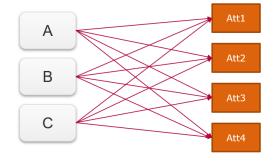


#### Discovering a Global Data Model from Event Logs



- A2A for all EL
- Multilog-Rule + Combine (Union)
- Fully Connected Rule
- Rules 1-4 as normal





#### Discovering a Global Data Model from Event Logs



- A2A for all EL
- Multilog-Rule + Combine (Union)
- Fully Connected Rule
- Rules 1-4 as normal

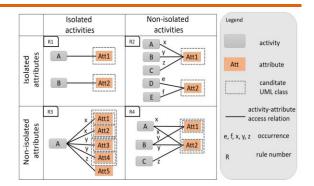
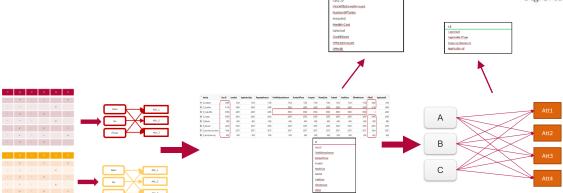


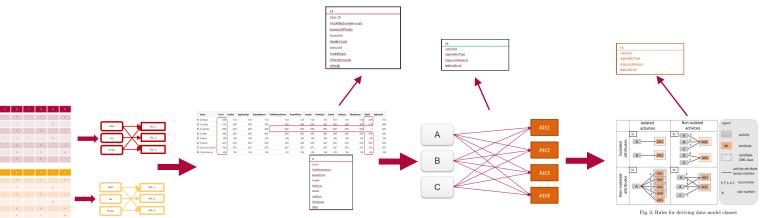
Fig. 3: Rules for deriving data model classes



#### Discovering a Global Data Model from Event Logs



- A2A for all EL
- Multilog-Rule + Combine (Union)
- Fully Connected Rule
- Rules 1-4 as normal



### Discovering a Global Data Model from Event Logs

## Dataset Overview – BPI Challenge 2020



- Dataset of five event logs
  - Permits: 86581 events, 170 attributes
  - DomesticDeclarations: 56437 events, 7 attributes
  - International Declarations: 72151 events, 20 attributes
  - PrepaidTravelCost: 18246 events, 19 attributes
  - RequestForPayment: 36796 events, 11 attributes

Discovering a Global Data Model from Event Logs