

Decentralising Data Collection and Anonymisation

Steven Murdoch University College London

An anonymisation system has architecture, security definition and mechanism

Architecture

How is data collected, stored, processed and distributed?

Security definition

What security and privacy properties must the system preserve, and what level of confidence is required?

Mechanism

How is access to data restricted or data altered in order to meet security definition?

In this presentation I'm going to show an alternative architecture to the traditional centralized data curator allowing the relaxation of the "plausible deniability" security property

A trusted, centralised architecture is common but creates risk and expense



A trusted, centralised architecture is common but creates risk and expense

Trusted In a position to violate security properties



Data submitted to central authority is inherently sensitive

To give data subjects confidence that their data will be protected, strong procedures and technical controls need to be in place, which is expensive in terms of effort, lost opportunity, and monetary cost



Within the centralised architecture, security definition and mechanism may vary

Data released could be noninteractive (anonymised version of original data, aggregate statistics), or interactive (results of a query interface)

		Security definition				
		Differential privacy	<i>k</i> -anonymity			
Architectural variant	Interactive	Adding Laplace noise to queries	Generalisation within query			
	Non-interactive	Aggregate statistics	Record suppression			

To me, differential privacy may be as authoritarian in its conceptual underpinnings as [Identity Based Cryptography]

The Moral Character of Cryptographic Work , Phillip Rogaway

Architecture, security definition and mechanism are linked but often conflated. Authoritarian tendencies come from trusted central authority, not security definition. Criticism above applies mainly to interactive centralized architecture

Confidence in institutions is at an all-time historical low

% who trust the govt in Washington always or most of the time



If data is anonymised by data subject, the central authority is no longer trusted



Google RAPPOR uses randomised response to prevent submissions being sensitive

Architecture

Untrusted central authority

Security definition Differential privacy

Mechanism Randomised response

Are you a member of the Communist Party?

- 1. Toss a fair coin
- 2. If heads, say "Yes"
- 3. If no, tell the truth

Gives **plausible deniability for individuals** but allows accurate aggregate statistics, and the uncertainty, to be calculated

RAPPOR extends this approach to arbitrary values, and shows that it fulfils differential privacy, subject to some caveats

Submissions are linked to individual identity though IP address, but are not sensitive even if linked to identity

RAPPOR offers strong security but is limited in terms of where it can be applied



Homepage URL can be identified only if it is set by at least about 14,000 users which limits the type of organisations that can gain useful insights

Security definition assumes consecutive responses are uncorrelated

Anonymisation can be extended to include not only data but communications too



Anonymisation can be extended to include not only data but communications too



If data is only sensitive if linked to identity and communication is anonymous, the central authority is no longer trusted



Short location traces can be anonymous if unlinked from other traces and identity

- An individual's location trace can be easily linked to an individual through auxiliary data
- Shown by Neustar Research using NYC Taxi dataset



Splitting the location trace into small segments greatly reduces the risk of re-identification provided individual traces cannot be linked either by the trace or by the IP address the trace is submitted from

Intrusion detection logs can be anonymous if unlinked from other logs and identity

Bit Dary Bitorial Samt CMI Seventhum: 1921888 Usenthum: Description Setting Seventhum: Setting Stating Seventhum:	>					_			SGUIL-0.7	0-ALP	HA - Conn	ected Te	192.	168.8.8			_	_					00
Statues Events Statues Events 51 GVL 5 Status Ave:10 Over/12/20 (2013) 25 (21) 25 (22) 213) 25 (13) 26 (13) 35 (13) 45 (23) 123 (26) 123 (18) 13) 1 BLEEDINGEDGE Poremal SSH Scan File 1 gareway 2.649 (2073) 23 (21) 28 (23) 123 (28) 123 (Eile Qu	uery B	epor	ts Sound: (off Server	Vame: 192	2.168.8.8 U	serName: bamr	UserID: 2												200	07-03-20	10:49:47 GMT
St. CVI Steer Aver 10 Oxor / me Soc IP SPate Db 11/P OP or P Evert Message I gateway 2.642 2007.33.12 02.51.16 63.26.188.133 1 BLEEDING-EDCE WORM Allade (CMP Sweep Prog Indound I gateway 2.644 2007.33.20 12.5 2.52 2.01.20.58.133 4000 201.20.188.193 1 BLEEDING-EDCE WORM Allade (CMP Sweep Prog Indound I gateway 2.646 2007.33.20 02.21.0 2.41.76.8.17 201.20.188.193 1 BLEEDING-EDCE WORM Allade (CMP Sweep Prog Indound ST CVI Statistic Aver 20 Dev/fines Statistic 6667 201.20.188.193 1 BLEEDING-EDCE WORM Allade (CMP Sweep Prog Indound ST CVI Statistic Aver 20 Dev/fines Statistic 6667 201.20.188.193 1/4 17 MS-SCL writein overflow attempt R1 gateway 2.6307 2007.31.9 0.813.1 2/4 1.41.02 1/7 70.12.10.81.193 1/4 17 MS-SCL writein overflow attempt R1 gateway 4.3 2007.23.10 0.85.73	RealTi	ime Eve	ents	Escalated Ex	vents																		
St CVII Struct Aurulo Double Struct Point P Event Manage 1 geneway 2642 2072-319 20 351 35 135 403 4052 2012.10.108.139 2 6 BLEEDING-EDGE WORM Allade (LMP Sweep Prog Indound 1 geneway 2644 2072-319 20 351 25 21 12 42 2017.2011 2012.10.108.139 1 BLEEDING-EDGE WORM Allade (LMP Sweep Prog Indound 1 geneway 2644 2007.201.90 25 10 24 1376 5116 2012.10.188.193 1 BLEEDING-EDGE WORM Allade (LMP Sweep Prog Indound 1 geneway 26449 2007.201.90 63 25 1 67 44 64 122 1177 2012.10.18.193 1 BLEEDING-EDGE POLICY IRC connection 1 geneway 2639 2007.201.90 63 25 1 27 42 106 45 1 6667 2012.10.18.193 143 41 7 MS-SQL version overflow attempt 1 geneway 2639 2007.201.90 63 25 7 42 10 46 40 2 202.101.62 218 100 40 20 120 188.193 143 41 7 MS-SQL version overflow attempt 1 geneway 2.639 2007.201.80 63 27 42 10 45 82 10 20 04 26 20 20 120 188.193 126 PADS Nex Assex - s	_	_	_		_	_						_	_		_		_		_	_	_	_	
0 1 gateway 2.6442 2.0072-312 213 22 425 12 4 213 124 2.0012.018.193 1 BLEEDING-EDGE WORM halfset (CMP Sweep Prog Innovad 1 gateway 2.6449 2.0072-312 213 22 22 2.13.12.116 20012.018.193 1 BLEEDING-EDGE WORM Alaget (CMP Sweep Prog Innovad 1 gateway 2.6449 2.0072-312 213 22 22 2.13.12.116 20012.018.193 1 BLEEDING-EGGE WORM Alaget (CMP Sweep Prog Innovad 51 CVN Stence 2.6449 2.0072-312 00 2160 241.78.81.77 20012.018.193 1 BLEEDING-EGGE WORM Alaget (CMP Sweep Prog Innovad 51 1.9 gateway 2.6495 20072-312 00 2160 2.6412 12.101.1018.193 5500 00 BLEEDING-EGGE WORM Alaget (CMP Sweep Prog Innovad 51 1.9 gateway 2.6295 20072-319 00 8151 07 244.46122 1107 20012.018.193 144 17 MS-502, version avertime attranse 51 1.9 gateway 2.6295 20072-319 00 8151 07 224.118.120 20012.018.193 144 17 MS-502, version avertime attranse 52 CVN Stence 2.002 72-319 210 212.218 2.001 20.118.193 126 PADS New Asset - sch OpenSSH 13.50 (Pintcol 197) 51 <td>ST</td> <td>CNT</td> <td>S</td> <td>nsor</td> <td>Alert ID</td> <td>Dat</td> <td>te/Time</td> <td>Src IP</td> <td>SPort</td> <td>Dst</td> <td>P</td> <td>DPort</td> <td>Pr</td> <td>Event Messa</td> <td>ge</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td> î</td>	ST	CNT	S	nsor	Alert ID	Dat	te/Time	Src IP	SPort	Dst	P	DPort	Pr	Event Messa	ge								î
No. Display CMI 4 Stand Zord 32 bits 13 bits Zord 32 bits 12 bits <thzord 12="" 32="" bits="" bits<="" th=""></thzord>	R	1	ga	teway	2.6422	2007-03	-19 23:05:14	63.26.198.32		209.	120.188.193		1	BLEEDING-E	DGE WO	HM AI	llaple ICM	IN 2466	p Ping Inb	ound			_
1 garewy 2.0499 2007/33/0 / 25 / 27 27 / 17.1116 2001/12.188.193 1 BLEEDING-EDGE WORM Alapte (CMP Sweep Prog Inhound SI CVI Server W 2.6496 2007/33.00 28.10 24.178.8.170 2001/20.188.193 1 BLEEDING-EDGE WORM Alapte (CMP Sweep Prog Inhound SI CVI Server W 2.6496 2007/33.18 0.216 65.188.126 6667 2001/20.188.193 149 DELEDING-EDGE WORM Alapte (CMP Sweep Prog Inhound SI 1 gareway 2.6289 2007/33.18 0.251 67.444.61.22 11/7 2001/20.188.193 149 17 MS-50L version verifice attempt N 1 gareway 2.6387 2007/20.19 0.812.02 2001/20.188.193 149 17 MS-50L version verifice attempt N 1 gareway 2.6387 2007/20.19 0.812.03 2001/20.188.193 128 6 PADS New Asset - sch OpenS2H 3.90.1 Portcor.1.90 N 1 gareway 4.3 2007/20.10 0.428 2001/20.188.193 128 6 PADS New Asset - sch OpenS2H 3.90.1 Portcor.1.90 N 1 gareway 4.6 2007/23.20 0.446.23 2001/20.188.1	RI	64	ga	teway	2.6424	2007-03	-19 23:13:26	203.193.56.1	33 45026	209.	120.166.193	22	6	BLEEDING-E	DGE Pot	tential 3	SSH Scar	n					
I gareway 2.849 201/21-20 (0.2 k) [2] X1/2 k L/2 201/21.9 k L/3 1 BLELDING-ELGE: WORM Allage ICMP Sweep Prig Informal ST CN1 Sensor Auer ID DearTime SCIP SPart Out IP Op on TP Event Message RT I gareway 2.589 207/23-38 00 130 65 (1.8 L/2.56 6677 201/20.18.193 14 kl 17 MS-SOL vertice vertifies attempt RT I gareway 2.589 2007/23-38 00 130 CALL (kl L/2.10 kl	R	1	ga	reway	2.6489	2007-03	-19 23:52:22	213.7.12.116		209.	120.188.193		1	BLEEDING-E	DGE WO	RM AI	llaple ICM	P Swee	p Ping Inb	ound			_
ST CNT Streture Autritio Description Streture Streure Stretur	R	1	ga	teway	2.6496	2007-03	-20 00:28:10	24.178.8.170		209.	120.188.193		1	BLEEDING-E	DGE WO	RM AI	llaple ICM	IP Sweet	p Ping Inb	ound			
St. C VIT Status Color File Distance File Distance File Distance File File<		L CNT			Alers ID	0.4		[Sec 10	[cnu	1.000	0	Lonu	l n.	Constants	_	_							
No. 1 general 2.509 1207/25120.15.0 021121.15.0 021121.15.0123 344 11 MS-SCL version overflow attempt R1 1 general 2.509 2207/23120.15.15 07.214.41.21 DOI: DOI: MS-SCL version overflow attempt R1 1 general 2.509 2207/23120.15.15 07.214.41.22 DOI:	DT		100	town	2 5090	2007.02	18 20-21-06	STO 192	6667	200	120 169 102	5997	6	DI CEDINC E	DCE PO		PC como	etion					
Bit In general CMI P (2017) 10 (15 / 17 / 12	DT	1	ya	towny	2.3500	2007-03	10 00 03 51	87 348 4E 12	1173	209.	120.100.193	1434	17	MS-SOL ware	ion avoid	Los atta	areat	scroff					_
Description Listen Control bases / Listen <thcontrol bases="" listen<="" th=""> Control ba</thcontrol>	DT	1	ga	temay	2.0296	2007-03	10 18 27 24	220 179 42 9	2064	209.	120.108.193	1424	17	MS-SQL Vers	ion over	low atte	ernpt						_
Image: Normal States State Normal States Normal Normal Normal Normal Normal Normal Normal Norm	DT	1	90	towny	2.0300	2007-03	10.10-16-20	200 101 62 2	1020	200	120 100 102	1424	17	MS SOL war	ion quod	low atte	onect						
ST CNT Sterior Aver ID Over/Time Size IP Spect Dit IP Open P/L Everet Message RT 1 gareway 4.3 2207/343 (D 24) 52 2201.28.10.03 45398 201.20.18.139 22 6 PADS Nex Asset - sith Open/SSH 3.9p1 (Protocol 1.9) RT 1 gareway 4.4 2207/343 (D 23) 21.147/253.20 000 200.12.01.88.139 23 6 PADS Nex Asset - sith Open/SSH 3.9p1 (Protocol 1.9) RT 1 gareway 4.5 2207/34 (D 23) 21.147/253.20 4044 200.12.01.88.193 24 6 PADS Nex Asset - sith Open/SSH 3.9p1 (Protocol 1.9) RT 1 gareway 4.6 2207/34 (D 12) 162.18.202.86 4044 200.12.01.88.193 24 6 PADS Nex Asset - sith Open/SSH 3.9p1 (Protocol 2.0) IP Resolution Agent Status S	IN .	1	ya	teway	2.038/	2007-03	-15 15.10.20	202.101.02.2	10 1030	205.	120.100.193	14.54	17	Marage Pers	ion oven	IOW ALL	empt						-
RT 1 generacy 4.3 2027/33/31 D44 52 202.188.103 4529 202.10.188.133 22 6 PAOS New Asset - son OpenSSH 3.91 (Protocol 1.99) RT 1 generacy 4.4 2027/33/31 D144 62 (20.66.3) 301/21 D181/33 11/8 6 PAOS New Asset - son OpenSSH 3.91 (Protocol 1.99) RT 1 generacy 4.4 2027/33/31 D144 62 (20.66.3) 301/31 D201/20.188.133 218 6 PAOS New Asset - son OpenSSH 3.91 (Protocol 1.99) RT 1 generacy 4.6 2027/33/3 D144 64 (20.66.3) 301/32 D112/10.188.133 216 PAOS New Asset - son OpenSSH 4.3 (Protocol 2.9) RT 1 generacy 4.6 2007/33/3 D144 64 (20.66.5) 301/32 D112/10.188.193 24 6 PAOS New Asset - son OpenSSH 4.3 (Protocol 2.9) RT 1 generacy 4.6 2007/33/2 D144 64 Source IP Paol New Asset - son OpenSSH 4.3 (Protocol 2.9) Store Paol Source IP Source IP Source IP Paol New Asset - son OpenSSH 4.3 (Protocol 2.9) Paol New Asset - son OpenSSH 4.3 (Protocol 2.9) Store Paol Source IP Description IPIOT (equere 1.15 (1.1004 (1.100 (1.100 (1.100 (1.100 (1.100 (1.100 (1.100 (1.100 (1.100 (ST	CNT	Is	7507	Alert ID	Dat	to/Time	Sec 10	SPort	Det	P	DRog	D,	Event Marra	a e	_							
RT 1 gareway 4.4 2207/33/21 D3 93 31 21.147/250 20 6000 2201/20.188.139 3128 6 PAOS New Asset - www stand 25.55TABLE6 RT 1 gareway 4.5 2207/33/21 D3 40.44 62.96.6.3 39113 201/20.188.139 23 6 PAOS New Asset - wsh OpenSSH 3.90 (Pontcol 1.90) RT 1 gareway 4.5 2207/33/21 D3 40.44 62.96.6.3 39113 201/20.188.139 24 6 PAOS New Asset - sth OpenSSH 3.90 (Pontcol 2.0) IP Residence Apent Status Stat	RT	1	03	teway	4.3	2007-03	-01 02:49:52	202.188.160	53 45398	209.	120.188.193	22	б	PADS New A	55.01 - 55	h Onen	SSH 3.9r	n1 (Proto	col 1.99)				_
RT 1 gareway 4.5 2207/23/33 (0.4246) 682.96.96.3 39113 231.20.188.193 23 6 PADS New Asset - sch OpenSSH 3.91 (Printed 1.97) RT 1 gareway 4.6 2007/23/36 (0.5112) 162.182.022.88 49412 201.20.188.193 24 6 PADS New Asset - sch OpenSSH 3.91 (Printed 1.97) RT 1 gareway 4.6 2007/23/36 (0.5112) 162.182.022.88 49442 201.20.188.193 44 6 PADS New Asset - sch OpenSSH 3.91 (Printed 1.97) RT 1. Store Store Store Market Data Store Store Store Store Store Market Data Store Mark	RT	1	03	teway	4.4	2007-03	01 05:35:33	211.147.250	0 6000	209.	120.188.193	3128	6	PADS New A	sset - we	w saui	d/2.5.5T/	ABLES					- B
RT 1 gateway 4.6 2027/33/28 15 12 20 162.18 202.88 40942 203.120.188.193 44 6 PADS New Asset -sch OpenSSH 4.3 (Proncul 2.0) IP Resolution Agent Status <td< td=""><td>BT</td><td>1</td><td>03</td><td>teway</td><td>4.5</td><td>2007-03</td><td>03 03:40:46</td><td>82.96.96.3</td><td>38419</td><td>209.</td><td>120.188.193</td><td>23</td><td>6</td><td>PADS New A</td><td>sset - ss</td><td>h Open</td><td>SSH 3.9</td><td>ol (Proto</td><td>col 1.99)</td><td></td><td></td><td></td><td></td></td<>	BT	1	03	teway	4.5	2007-03	03 03:40:46	82.96.96.3	38419	209.	120.188.193	23	6	PADS New A	sset - ss	h Open	SSH 3.9	ol (Proto	col 1.99)				
IP Residure Strict Status </td <td>BT</td> <td>1</td> <td>03</td> <td>teway</td> <td>4.6</td> <td>2007-03</td> <td>-08 15:13:20</td> <td>162.18.202.8</td> <td>8 40942</td> <td>209.</td> <td>120.188.193</td> <td>44</td> <td>6</td> <td>PADS New A</td> <td>sset - ss</td> <td>h Open</td> <td>SSH 4.3</td> <td>Protoco</td> <td>1 2.0)</td> <td></td> <td></td> <td></td> <td></td>	BT	1	03	teway	4.6	2007-03	-08 15:13:20	162.18.202.8	8 40942	209.	120.188.193	44	6	PADS New A	sset - ss	h Open	SSH 4.3	Protoco	1 2.0)				
Weiner Source IP Description Mark Statestick Systems Marger List Source IP Description Mark Statestick Description Description Mark Statestick Description Description <thdescriptin< th=""> Descriptin Descri</thdescriptin<>	-																		,				Ļ
IP Residuars System Korg Laser Mage System Korg Laser Mage Sol Magert Status System Korg Laser Mage Status Show Rade Status System Korg Laser Mage Sol Magert Status System Korg Laser Mage Status Show Rade Status System Korg Laser Mage Status Show Rade Status Maint.gov 1 Ext. Ner gateway sand 2007/03/00 02 810 PB PM Laser Magert Status Maint.gov PM Laser Magert Status Show Rade Status Show Rade Status Maint.gov PM Laser Magert Status Show Rade Status Show Rade Status Show Rade Status Show Rade Status Maint.gov PM Laser Magert Status Show Rade			_																				
Stor Hots Hotsenane Type Last Strature 10641044<	IP R	lesolutio	n Ì	Agent Status	Snort St	atistics	System Msg	User Msg		📕 Sho	w Packet Da	ta 📕 Sho	w Rule	www.snort.or	g nvd	.nist.go	w						
Control 100	Sid	Mar		Hostop			1.		Statue	alert ud	p \$EXTERN	AL_NET a	ny -> \$i	IOME_NET 143	4 (msg:"	MS-SQ	L version	ı overflov	v attempt"	; flowbits:	isnotset, m	is_sql_se	en_dns; 🗄
2 Ext.Nerr gateway snort 2007/03/20 00 28 10 LP 3 Sac.Nerr gateway sac.p 2007/03/20 10 15 93 4 GP 4 Ext.Nerr gateway sac.p 2007/03/20 10 15 93 4 GP 7 GP Port Dest Port Chestion Chestion 10 0 0.01 5 13 20 0.01 0.01 10 10 10 10 10 10 10 10 10 10 10 10 1	1	Ext_Ne	et	gateway	/ DC	ap	2007-03-20	00:49:23	UP	dsize:>	100; content	1041 ; dej	itn:1; re	rerence: bugirad	5310; ret	a ozo .	Cve, 2007	2-0649; 1	eterence	105505,10	074; :0: 0-3		17
3 Ext.Ner gateway sance 2007/03100159.34 120 4 Ext.Ner gateway pads 2007/031061511220 144 5 Ext.Ner gateway pads 2007/031061511220 144 Update Interval (sect): 120 ▼ NOW Update Interval (sect): 120 ▼ NOW	2	Ext_Ne	et .	gateway	/ 50	ort	2007-03-21	00:28:10	UP	IP	Sou	rce IP		Dest IP	Ver	HL	TOS	len	ID	Flags	Offset	TTL	ChkSum
4 Ext_Ner gateway pads 2007/03 06 15 13 20 LP 4 Ext_Ner gateway pads 2007/03 06 15 13 20 LP 9 064 F044 F044 <t< td=""><td>3</td><td>Ext. Ne</td><td>er</td><td>gateway</td><td>/ 5a</td><td>ncp</td><td>2007-03-1</td><td>01:59:34</td><td>UP</td><td></td><td>220.178.4</td><td>13.82</td><td>209.</td><td>120.188.193</td><td>4</td><td>5</td><td>0</td><td>404</td><td>31898</td><td>0</td><td>0</td><td>115</td><td>13184</td></t<>	3	Ext. Ne	er	gateway	/ 5a	ncp	2007-03-1	01:59:34	UP		220.178.4	13.82	209.	120.188.193	4	5	0	404	31898	0	0	115	13184
UD Odd (144) (14	4	Ext_Ne	et .	gateway	/ pa	ds	2007-03-0	8 15:13:20	UP			Source			Dest				onath			ChkSur	
Update Interval (sect): 120 ▼ NW Update Interval (sect): 120 ▼ NW										UUP	3064 1434 384								37652				
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0											04 01 01	01 01 0	1 01 0	1 01 01 01 01	01 01 0	01 01	1001	_				_	12
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0											01 01 01 01 01	01 01 0	1 01 0	1 01 01 01 01 01 1 01 01 01 01 01 01 01	01 01 0	1 01							
01 01 01 01 01 01 01 01 01 01 01 01 01 0											01 01 01 01	01 01 0	1 01 0	1 01 01 01 01 01 1 01 01 01	01 01 0	1 01							
Update Interval (sect): 120 NOW Sector 10 feed 19 so 10 s											01 01 01	01 01 0 B0 42 E	1 01 0 B 0E 0	1 01 01 01 01 01	01 01 0	1 01							- 11
0 r r r r r r r r r r r r r r r r r r r										DATA	42 01 70	RE 42 9	0 90 9	0 90 90 90 90	90 68 1	C C9			B.p	.в			-
											01 01 05	5 50 89 E	5 51 6	8 2E 64 6C 6C	68 65 6	C 33				P. Qh.dl	lhe13		
06 77 75 26 56 59 66 59 67 45 66 76 32 76 35 86 66 19 19 74 55 16 75 67 36 76 35 86 66 19 19 74 55 16 75 67 56 75 75											43 68 47	65 74 5	4 66 B	9 6C 6C 51 68	33 32 2	ы 68 Е 64			ChGe	ernuncunt etTf,110	ласк h32.d		
Update Interval (secs): 120 👻 NOW Search Packet Payload 🗸 Hex ^ Text 🗆 NoCase											68 77 73 B9 74 6F	5 32 5F 6 51 68 7	ь в9 6 3 65 6	5 74 51 68 73 E 64 BE 18 10	67 63 6 RE 42 8	ag 66 3D 45			hus.	2_f.etQh: Qhsend	DOCKF		
						Updat	te Interval (se	ecs): 120	NOW			Se	arch Pa	cket Payload	🔆 Hex	$^{-}$ Te	ext 🔟 No	Case					

Organisations may wish to share information on attacks they are subject to but still hide their identity

Real example of a consortium of large companies, competing with each other but subject to same threats

Techniques like prefix-preserving anonymization can hide details from logs but still need to submit data without disclosing identity



Network nodes are run by volunteers, and changes to structure are made by consensus decision of 8 directory authority operators



Each node knows the prior step in the circuit, and the next step, but unless there is collusion nobody can link source to destination



Encryption prevents data flowing into a node from being linked to data flowing out, preserving unlinkability property

Tor and other common anonymity systems do not hide traffic patterns

There are useful alternatives between the extremes of fully centralised and full anonymisation by data subject

- Fully centralised approaches are versatile but have disadvantages
 - Risk of data compromise, expensive to set up and maintain security
 - Public lacks confidence in institutions
 - Centralisation of power lends itself to authoritarian tendencies
- Local anonymisation such as RAPPOR has limitations
 - Requires large amounts of data to get meaningful results
 - Care needed to preserve security if there is correlated data
- Data can be anonymised such that it is not sensitive in itself but does not offer plausible deniability either
 - In which case, submitting data over an anonymous communication system such as Tor can protect privacy without having to trust a central authority to anonymise data