Maltego3

WHRRROOOOOOOOOOOOM !

Agenda

- What is Maltego ?(bleh)
- What's new in v3?
- Using NER
- Social networks
- Container TAS
- Databases and Maltego
- Community edition

What is Maltego?

- Application that links bits of information
- Information is classified into ‘entity’ types
- Link is created by a piece of code called a transform
- Transform can be:
  - Built by Paterva (100 odd)
  - Yourself (local transforms)
  - Watch this space
- Super flexible (Lego set)

Paterva. This leads to much power...

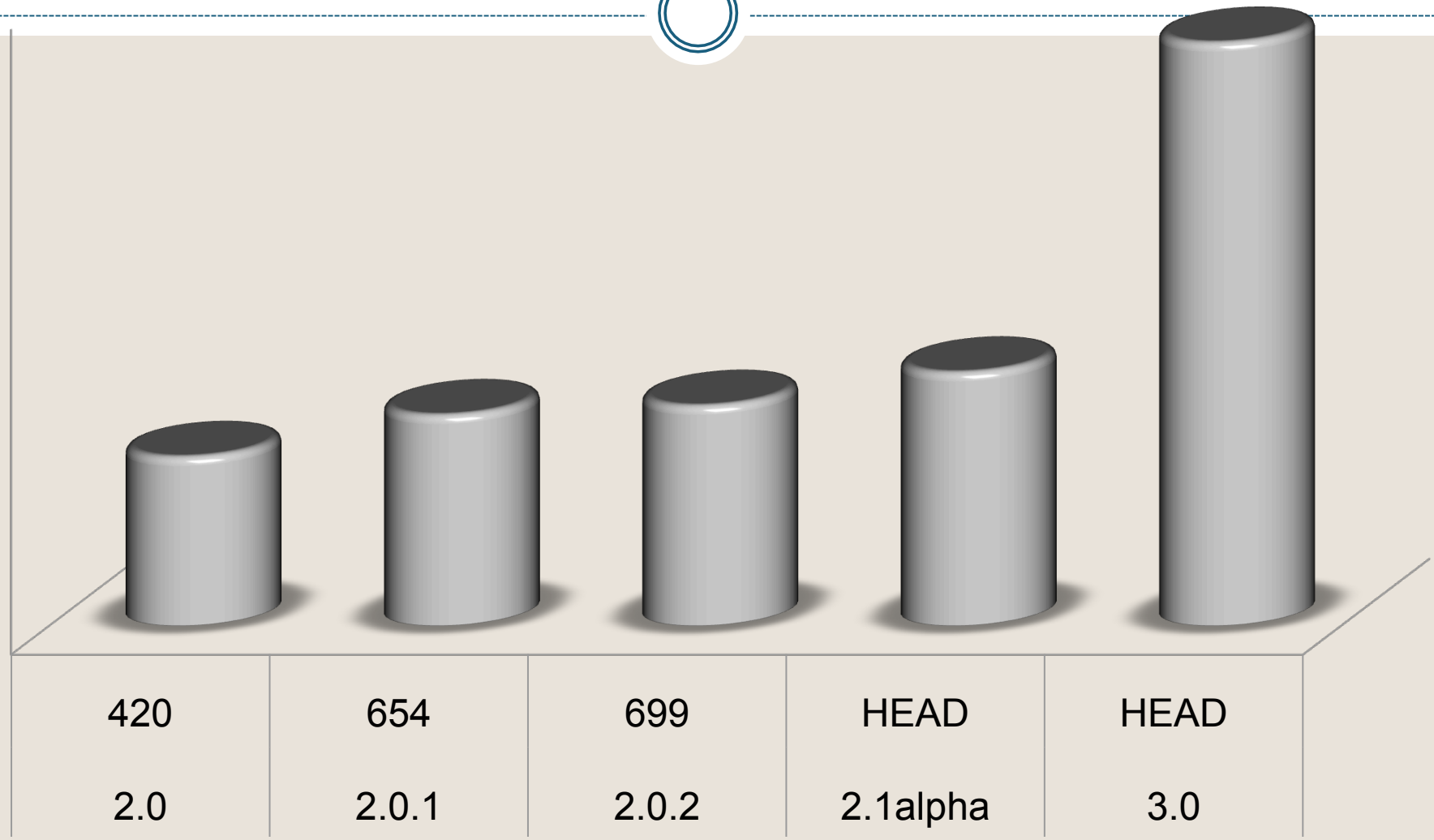
- Maltego can answer important questions like:
  - Which are the most likely weak machines in a network?
  - Which documents hosted on my domain are leaking sensitive information?
  - Who should I friend on Facebook to get invited to the cool parties?
  - Who will win the World Cup?
  - What is the meaning to life?

New in V3

- Look and feel
- Custom entities
- Manual linking
- Dynamic layout and Interactive EWV
- Base for 3.1

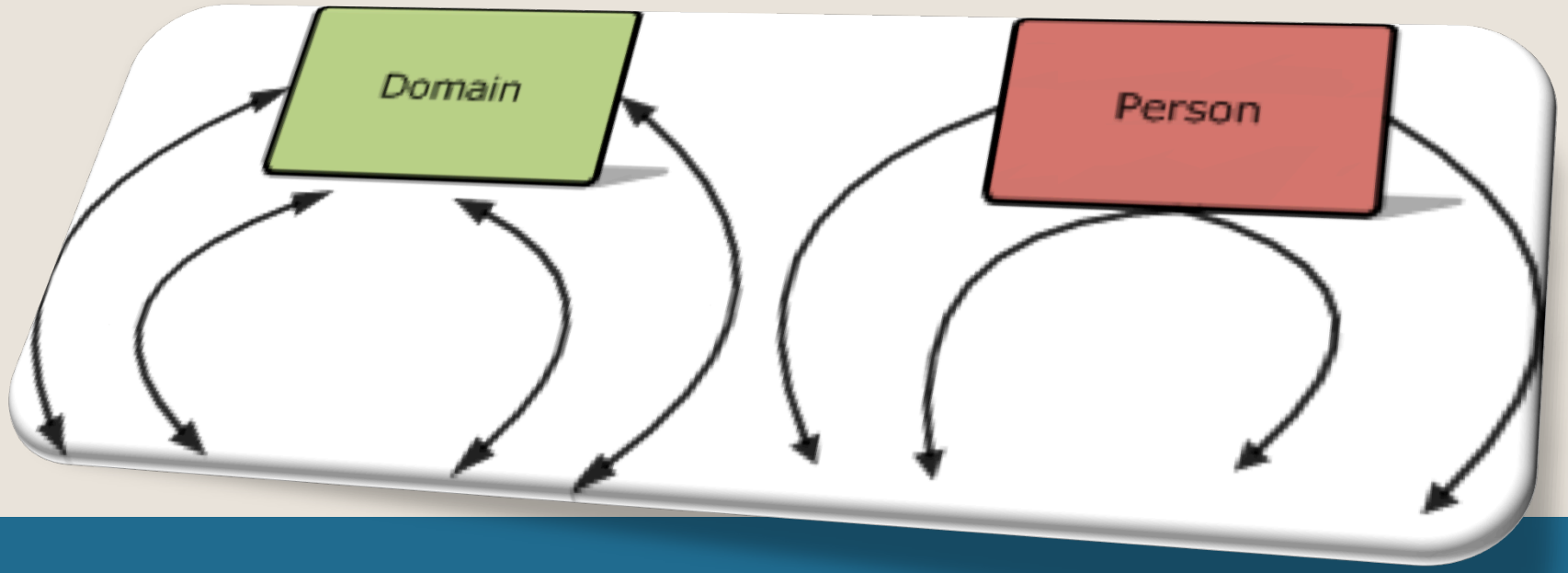
Paterva.

V3 >> V2



Problems in Maltego

- Good at infrastructure
  - Network mapping
  - Port scans blah blah
  - The usual...
- Dead end entities:

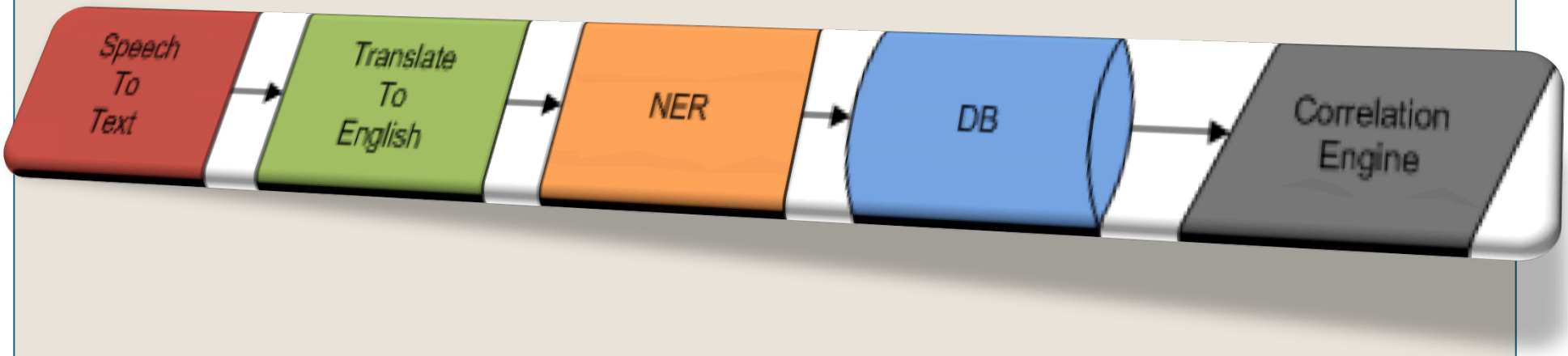


Named Entity Recognition

- What is NER?
  - Takes text and marks entities like person names / companies / phone numbers
- Demo:
  - OpenCalais / AlchemyAPI

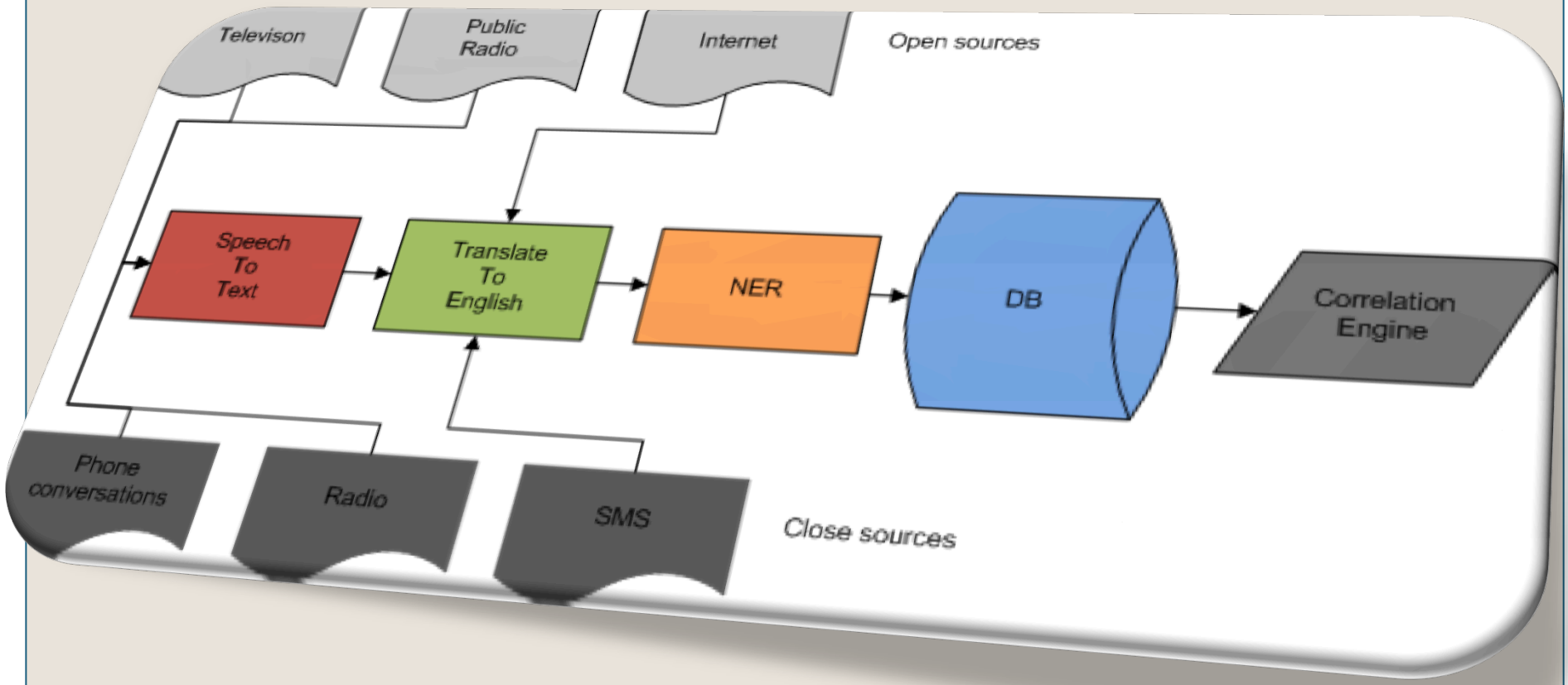
Named Entity Recognition

- Where is NER used?



Named Entity Recognition

- Who finds NER useful?



Using NER in Maltego

- Using it in Maltego:
  - Phrase ->
  - Website ->
  - URL ->
  - Entities
- Phrases can get interesting...we can combine with operators like:
  - Filetype:
  - Site:
  - Etc..
- Can answer the question:

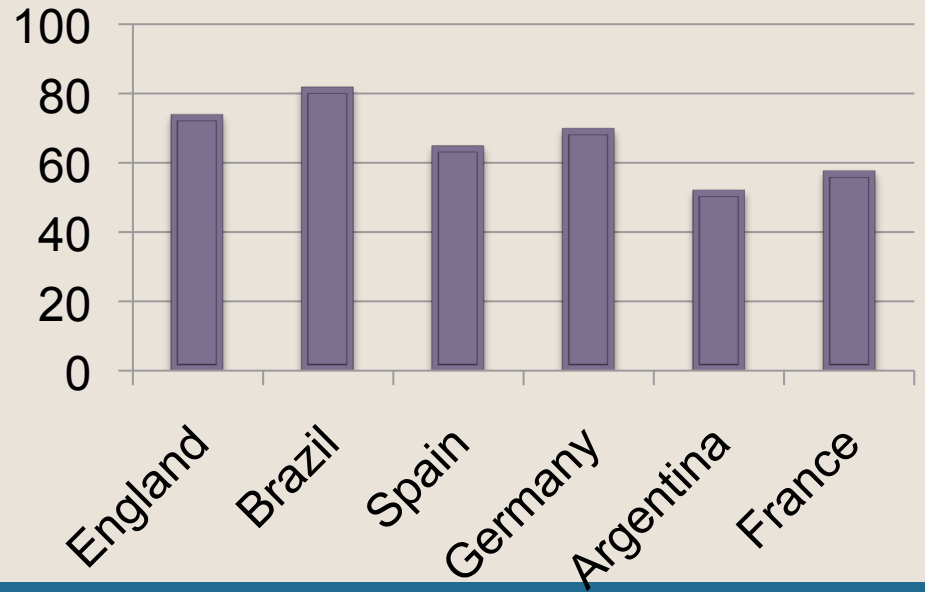
“Who/what/where is connected to phrase X?”

Paterva.

DEMO
Named Entity Recognition
Who/Where – World Cup predictions

Results... ;)

	English	%	French	%	Spanish	%	Portugese	%	Totals
England	0	**0**	6	20	14	33.33333	8	20.51282	73.84615
Brazil	50	23.14814815	9	30	12	28.57143	0	**0**	81.71958
Spain	47	21.75925926	6	20	0	0	9	23.07692	64.83618
Germany	46	21.2962963	5	16.66667	8	19.04762	5	12.82051	69.83109
Argentina	45	20.83333333	4	13.33333	0	0	7	17.94872	52.11538
France	28	12.96296296	0	0	8	19.04762	10	25.64103	57.65161
	216	**100**	**30**	**100**	**42**	**100**	**39**	**100**	400



Paterva.

DEMO
Google Goggles

Paterva

Scraping social networks

- Scraping is against most TOUs.
- They take it seriously!
- Scraping is not cool because:
  - They change their site regularly
  - If you want to hide via TOR the pages looks different
  - FB discourage it by setting cookies for 2038
    - Breaks the Mechanize library
  - Authentication – you need to keep the cookies alive
  - Cannot log in every time – FB checks for frequency of logins

Increasing reliability

- Where possible, use FQL (Facebook query language) or the API
- Use mobile sites – like iPhone Touch interface, m.facebook
  - Less complex results
  - Less likely to change
- Use the AJAX call
  - Data comes in cleaner, easier to parse
- Don’t rely on tags, use regex where possible
  - Eg id=/d{3,15}/&

Paterva.

How to

- Cron – keeping cookie alive
  - Runs every 5 minutes, ‘clicks’ on well known links on Touch FB site
  - If it gets 302 it re-logins
- Email to Facebook profile transform
  - Uses cron cookies, run query at iPhone site
  - Call /s.php?**k=100000020**&q=emailaddress on Touch
  - The historical k parameter means we can search for email addresses on mobile!
  - Returns the Facebook unique ID – pick it up with a regex
  - Get detail on the ID using standard FQL

Paterva.

How to

- Get friends
  - With the ID known, exploits the typeahead_friends AJAX bug.
- Typeahead_friends.php bug:
  1. Can make AJAX call un-authenticated! (typeahead_friends.php?u=ID&__a=1)
     - We don't need to worry about cookies from cron
  2. Get ALL friends of any user
     - Even if they are hidden
- Recently FB close hole 2, but we can still make AJAX call and get friends if profile settings allows it

Paterva.

How to

- Person name to Facebook profile
  - Can use standard FQL
  - Get a list of all matching ID
  - Foreach ID (do FQL lookup)
  - ‘Page’ through results

Paterva.

DEMO

Maltego + Facebook + NER

How to make friends and influence people

Playing with real data

- SQLTAS
- Hooks MySQL, MSSQL, Postgress, DB2 and Oracle into Maltego.
- Talks of making SQLTAS publically available.
  - Bribe us with beer.
- Maltego loves clean, crisp data

Paterva.

DEMO
SQLTAS

What can we learn from the carder's dB?

Paterva.

Join our party

- If you have skill (writing transforms for CCTAS) you can contribute
- If you have data (SQLTAS) you can contribute

Playing with LOTS of data

- Challenge:
  - > 200 databases
  - If only 3 entities = 200 x $3^2 - 3$ = 1200 transforms
  - 9 entities = up to 14 400 transforms
- Clearly we cannot do this by hand
  - Think super transform
  - NER for classification (and then some)
  - User focuses on destination, not path
  - Automatically calculates the best path
  - Context becomes completely lost now

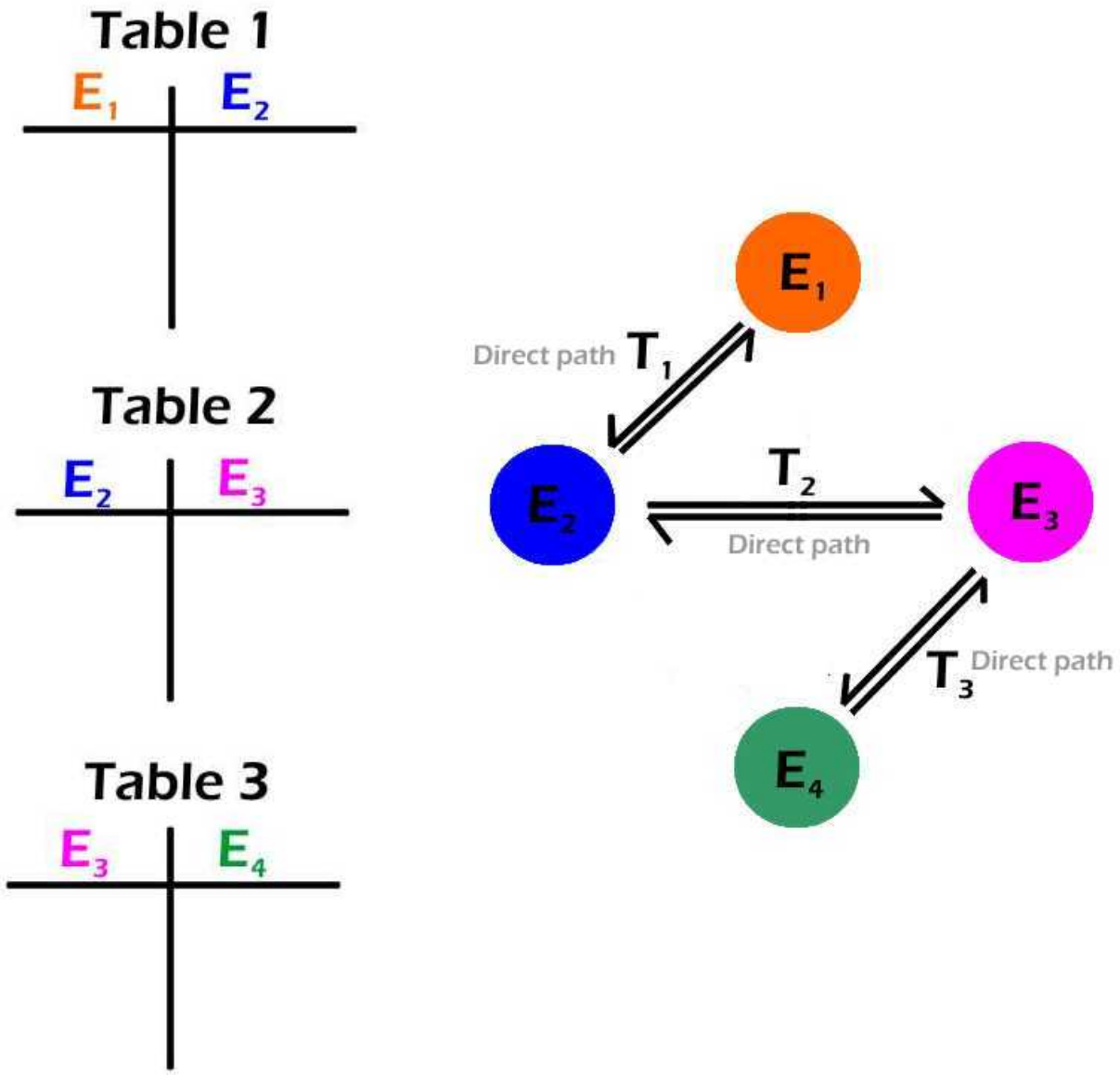


Table 1
E1
E2
Table 2
E2
E3
Table 3
E3
E4
E1
Direct path T1
E2
T2
E3
Direct path
T3 Direct path
E4

Table 1

E_1	E_2

	E_1	E_2
E_1		a
E_2	b	

Table 2

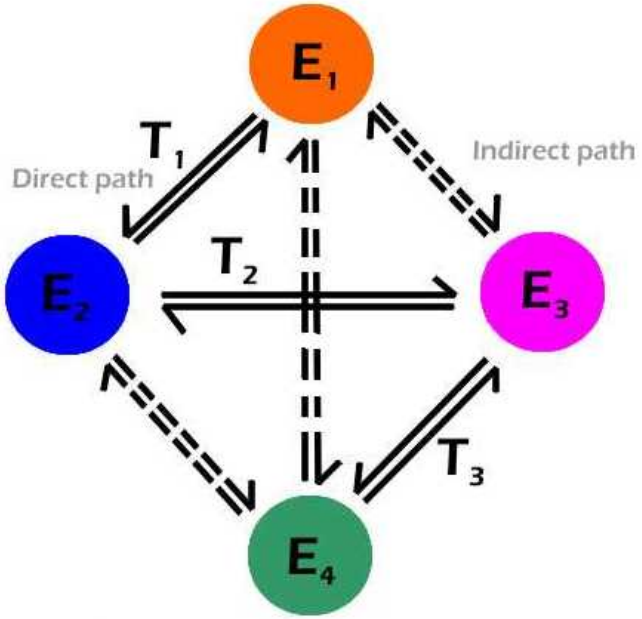
E_2	E_3

	E_2	E_3
E_2		c
E_3	d	

Table 3

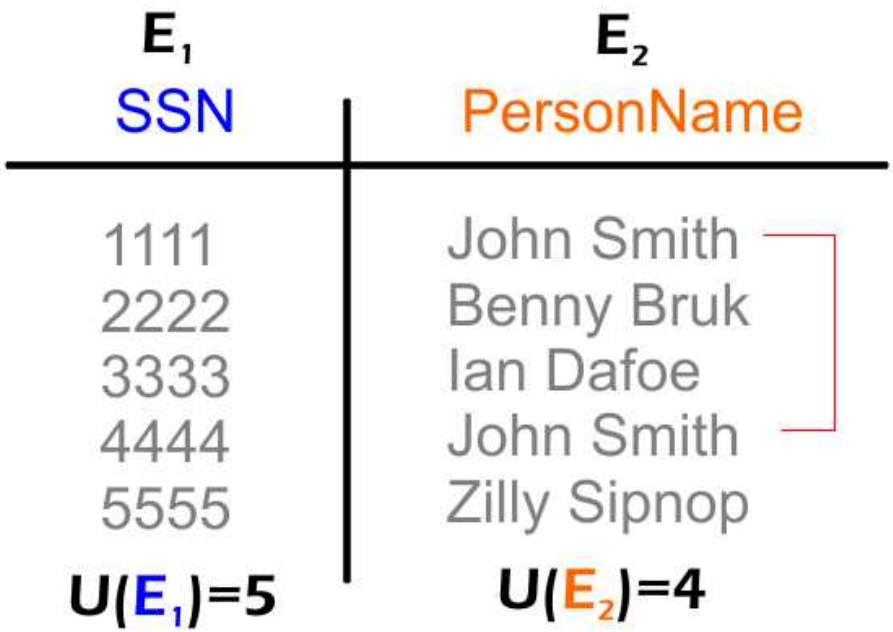
E_3	E_4

	E_3	E_4
E_3		e
E_4	f	



Reliability:

Path (E1 -> E2) : a
Path (E1 -> E3) : a.c
Path (E1 -> E4) : a.c.e



Number of rows: RC=5

E_1 SSN	E_2 PersonName	E_3 CarReg
1111	John Smith	ABC123GP
2222	Benny Bruk	BNB448GP
3333	Ian Dafoe	EMN667GP
4444	John Smith	ZGN321GP
5555	Zilly Sipnop	PEE101GP
1111	John Smith	DEF456GP
$U(E_1)=5$	$U(E_2)=4$	$U(E_3)=6$

Number of rows: RC=6

$$R(E_1 \rightarrow E_2) = \frac{U(E_1)}{RC} = \frac{5}{6} = 0.833$$

$$R(E_2 \rightarrow E_1) = \frac{U(E_2)}{RC} = \frac{4}{6} = 0.666$$

FAIL!

E_1	E_2	E_3
SSN	PersonName	CarReg
1111	John Smith	ABC123GP
2222	Benny Bruk	BNB448GP
3333	Ian Dafoe	EMN667GP
4444	John Smith	ZGN321GP
5555	Zilly Sipnop	PEE101GP
1111	John Smith	DEF456GP

$U(E_1)=5$ $U(E_2)=4$ $U(E_3)=6$

$U(E_1 + E_2) = 5$ $U(E_2 + E_3) = 6$

$U(E_1 + E_3) = 6$

E_1 SSN	E_2 PersonName	E_3 CarReg
1111	John Smith	ABC123GP
2222	Benny Bruk	BNB448GP
3333	Ian Dafoe	EMN667GP
4444	John Smith	ZGN321GP
5555	Zilly Sipnop	PEE101GP
1111	John Smith	DEF456GP

$U(E_1)=5$ $U(E_2)=4$ $U(E_3)=6$

$U(E_1 + E_2) = 5$ $U(E_2 + E_3) = 6$

$U(E_1 + E_3) = 6$

$$R(E_1 \to E_2) = \frac{U(E_1)}{U(E_1+E_2)} = 5/5 = 1$$

$$R(E_2 \to E_3) = \frac{U(E_2)}{U(E_2+E_3)} = 4/6 = 0.666$$

$$R(E_1 \to E_3) = \frac{U(E_1)}{U(E_1+E_3)} = 5/6 = 0.833$$

$$R(E_2 \to E_1) = \frac{U(E_2)}{U(E_1+E_2)} = 4/5 = 0.8$$

$$R(E_n \to E_m) = \frac{U(E_n)}{U(E_n+E_m)}$$

TABLE 1

FROM \ TO	E_1	E_2	E_3
E_1		1	0.833
E_2	0.8		0.666
E_3	1	1	

TABLE 1

FROM \ TO	E_1	E_2	E_3
E_1		1	0.833
E_2	0.8		0.666
E_3	1	1	

TABLE 2

FROM \ TO	E_7	E_2	E_5
E_7		0.75	0.3
E_2	0.25		1
E_5	0	0.473	

TABLE 1

FROM \ TO	E_1	E_2	E_3
E_1		1	0.833
E_2	0.8		0.666
E_3	1	1	

TABLE 2

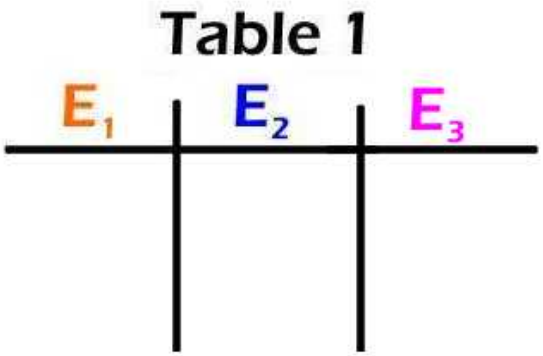
FROM \ TO	E_1					E_7
E_1						
E_7						

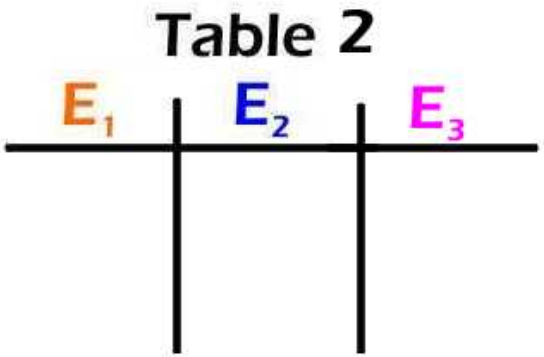
	SSN	Fullname	ZIP	TAX
SSN		0.929	0.819	0.999
Fullname	0.867		0.811	0.867
ZIP	0	0.01		0
TAX	0.947	0.851	0.808	

	TAX	Phone
TAX		0.397
Phone	0.470	

Reliability:
Fullname -> Phone

Path:
Fullname -> Tax -> Phone
0.867 x 0.397 = 0.344





T1:R(E_1->E_3) = 0.5

T1:A(E_1->E_3) = true

T2:R(E_1->E_3)= 0.99

T2:A(E_1->E_3)= false

Reliability vs. Availability

Reliability: 0 to 1
pre computed

Availability: 0 or 1
determined in real time