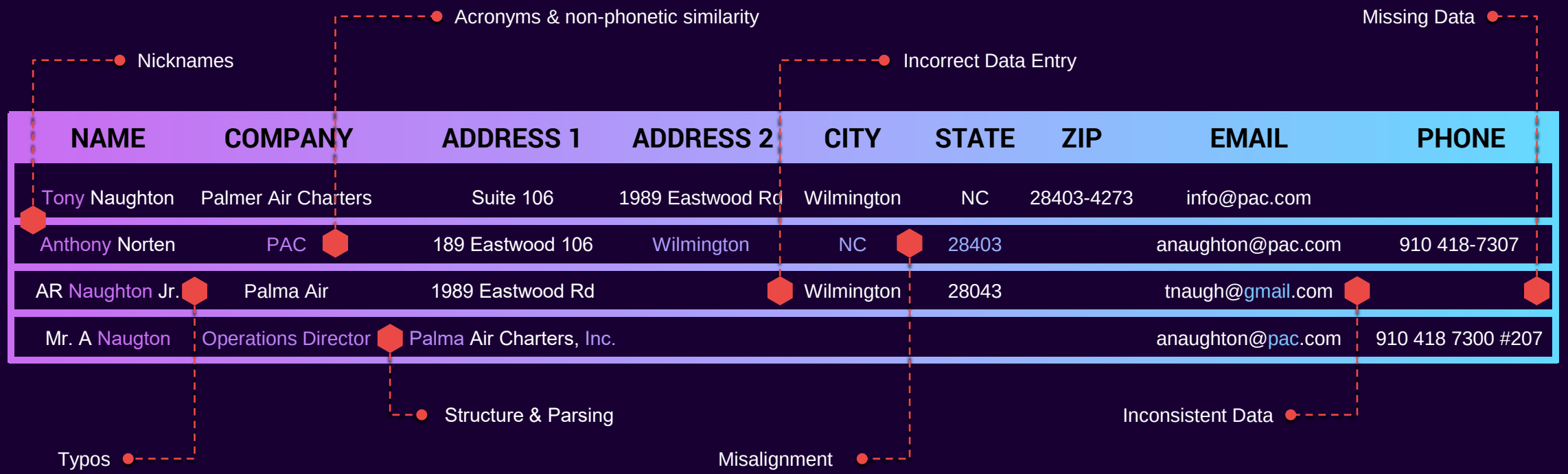

Syniti
MATCHING
+
SOFTPI
.COM

the problem

DATA IS MESSY



Syniti

use cases

WHERE DO WE FIND MATCHING

STRATEGIC

- CUSTOMER 360
- DATA QUALITY & BI
- MDM

DISCRETE

- LOYALTY PROGRAMS
- CRM HYGIENE & DEDUPE
- INCIDENT RESPONSE MANAGEMENT
- INSURANCE CLAIM RECONCILATION
- ENRICHMENT & APPEND

Syniti

matching

CONSIDERATIONS

Preparing Data

Configuration

Processing Speed

Accuracy of Results

Syniti

CONVENTIONAL MATCHING

conventional matching

VULNERABLE TO ERRORS IN DATA

MATCHCODE: Metaphone3(First_Name(3)) + Metaphone3(Last_Name(3)) + Street_Number(4) + ZIP(5)

MATCHKEY	NAME	ADDRESS	CITY	STATE	ZIP
TAMMAR350078746	Tamas Mayer	3500 N Capital of Texas Hwy #230	AUSTIN	TX	78746
TAMMAR350078746	Tom Moore	3500 N Capital of Texas Hwy #502	AUSTIN	TX	78746
TMAR35078746	Mr. T R Moore	350 N Capital of Texas Apt 502	AUSTIN	TX	78746

The matchkey serves as the basis of comparison.

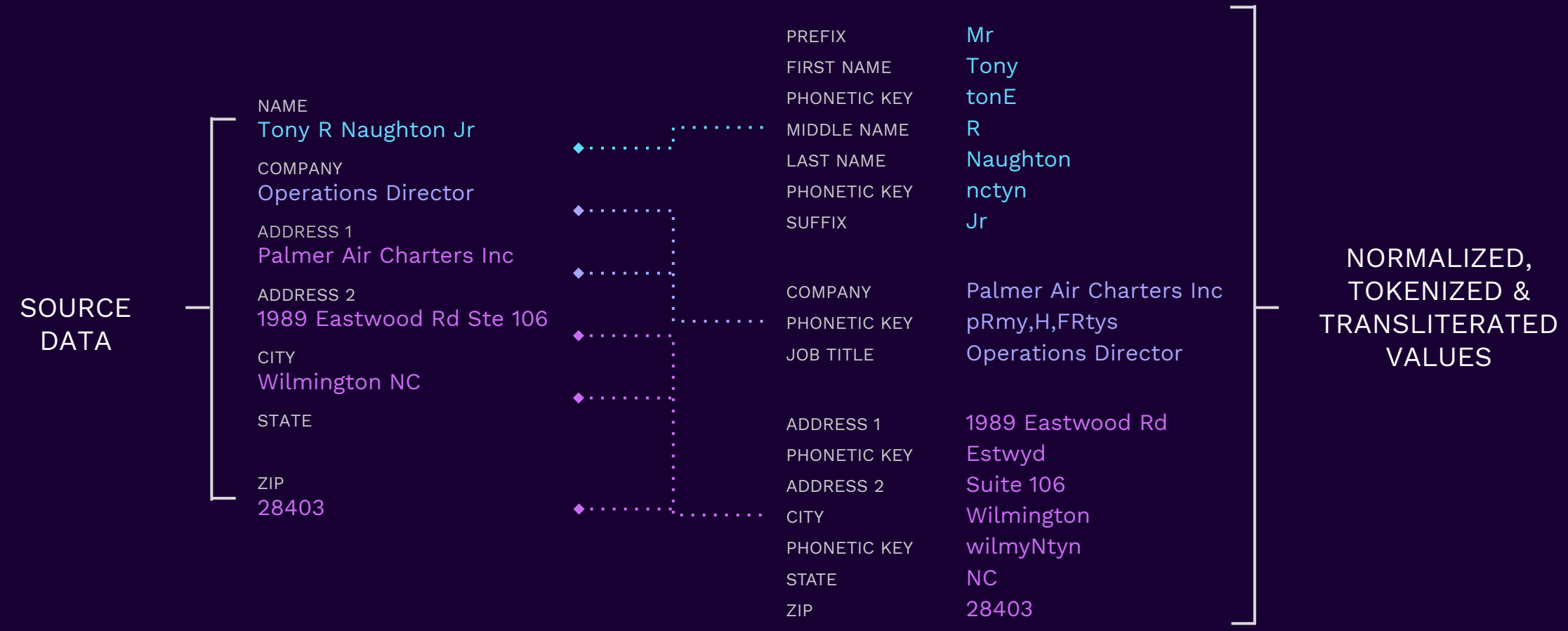
Syniti

Syniti

DATA MATCHING

intelligent parsing

BRING YOUR DATA AS IS



Syniti

enhanced algorithm

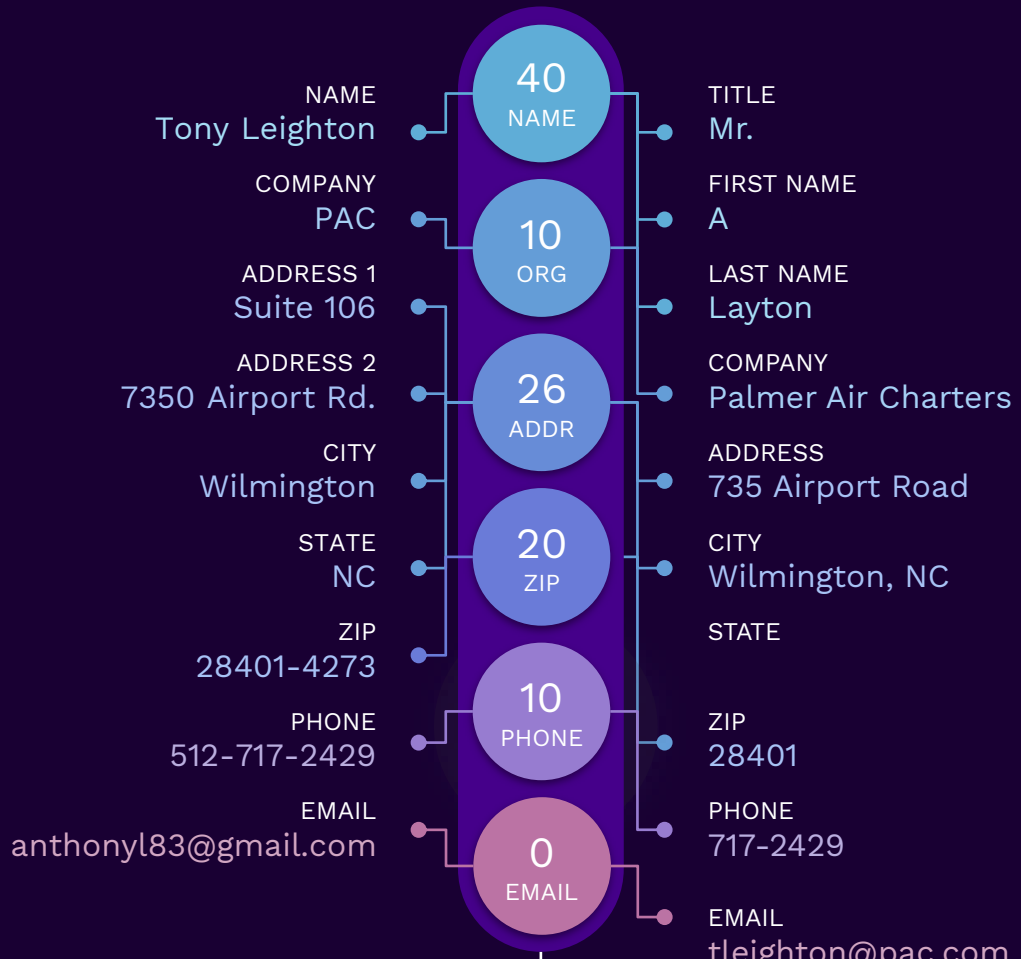
UNDERSTANDING PRONOUNCIATION

	Conventional Algorithms			
Surname	**Soundex**	**Metaphone**	**Metaphone3**	**soundit**
Thompson	T512	0MPSN	TAMPSAN	tomsyn
Thomson	T525	0MSN	TAMSAN	tomsyn
Tompson	T512	TMPSN	TAMPSAN	tomsyn
Moore	M600	MR	MAR	my
Mayer	M600	MYR	MAR	myy

Syniti

contextual scoring

HUMAN-LIKE PERCEPTION



MATCH SCORE

106

Syniti

PERFORMANCE

BUILT FOR SPEED

Whether you're dealing with a few thousand records or more than a billion, Syniti has you covered.

In-memory Processing is significantly faster than competing solutions that rely on sluggish disk storage.

Scanning for similarity to build Candidate Groups limits the number of comparisons that need to be made.

BATCH

DEDUPE

49m 10s

49 million records

MATCH TWO FILES

05m 03s

1 million : 30 million records

16-Core hyper-threaded Windows PC with 64GB RAM

REAL-TIME

LOOKUP MODE

00s 14ms

1 : 98 million records

MATCH TWO FILES

00m 15s

100,000 : 50 million records

16-Core hyper-threaded Windows PC with 64GB RAM

BIG DATA

DEDUPE

20m 15s

1 billion records

MATCH TWO FILES

07m 30s

100 million : 500 million records

20-machine cluster in AWS, each with 48 Cores & 192GB RAM

Syniti

TRANSLITERATION

GLOBAL DATA

(UTF8 or UTF16) Unicode input data is transliterated into English Latin characters (ISO 8859-1)

So, whether your data is Cyrillic Россия, Hellenic Ελλάδα, Hebrew ישראל, KANJI 日本, Simplified Chinese 中国, Arabic العربية, Thai ประเทศไทย, Hangul 대한민국, or yes, even Klingon - 'tlhIngan, the 360Science Matching Engine has you covered.

Full Name	Phonetic Name Key
昌李	lyC
厂里	lyC
Chang Li	lyC

Syniti

scorecard

ONE SLIDE TO REMEMBER

CONVENTIONAL MATCHING

Requires “match ready” data which needs significant preprocessing.

Users must understand the nuances of various off-the-shelf algorithms.

Match keys are the basis for comparison, which are vulnerable to errors in data.

Processing with traditional blocking and storage is slow.

Preparing Data

Configuration

Results

Speed

Syniti

No need for preprocessing, bring your data as it is.

Utilizes a proprietary phonetic algorithm built for contact and business data.

Contextual scoring which mirrors human-like perception is much more accurate.

Processing in minutes instead of hours or days

DEMO