Forensic Triage for Mobile Phones with DEC0DE

Robert J. Walls Erik Learned-Miller Brian Neil Levine

Department of Computer Science
University of Massachusetts Amherst

This work was supported in part by NSF award DUE-0830876.

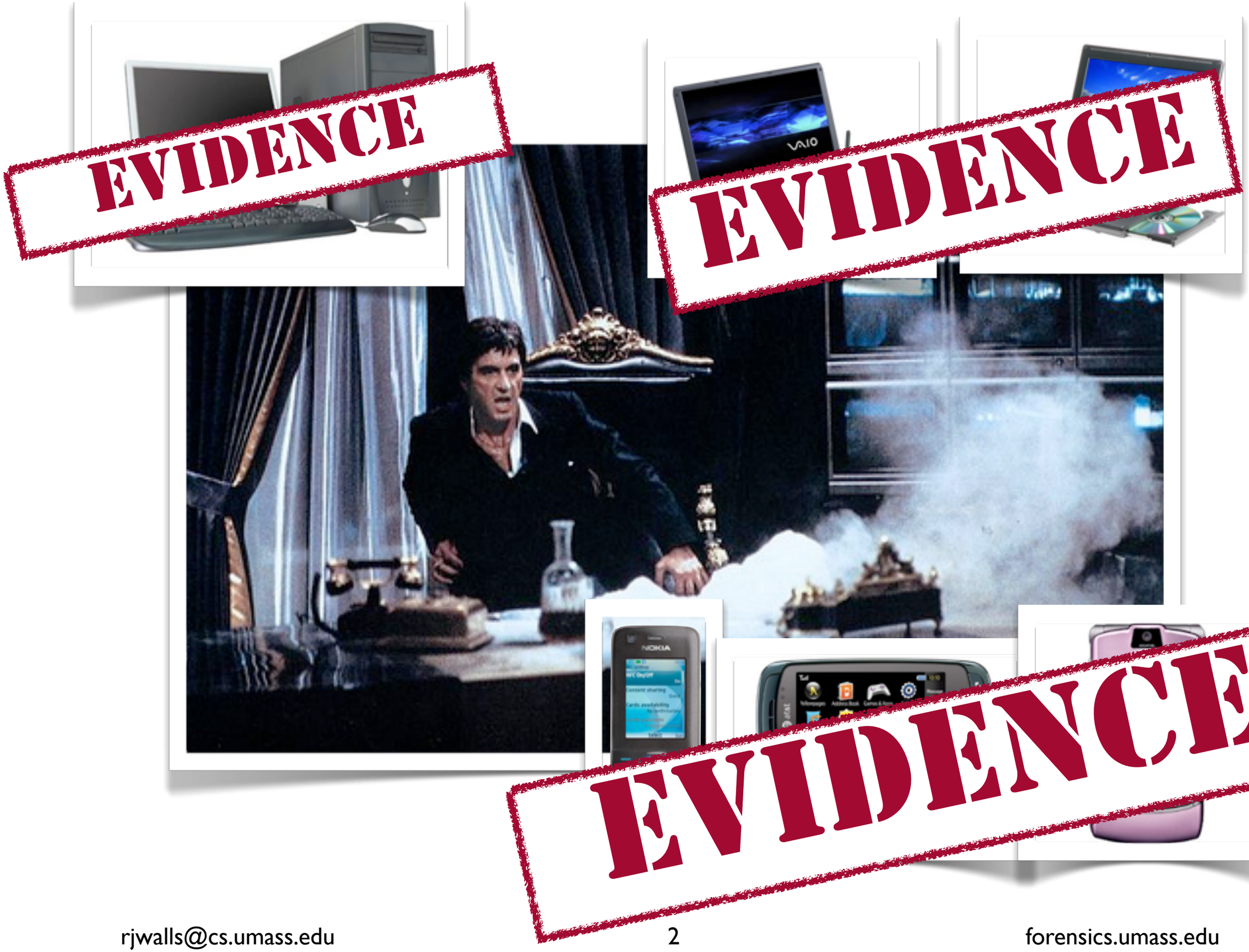
rjwalls@cs.umass.edu forensics.umass.edu

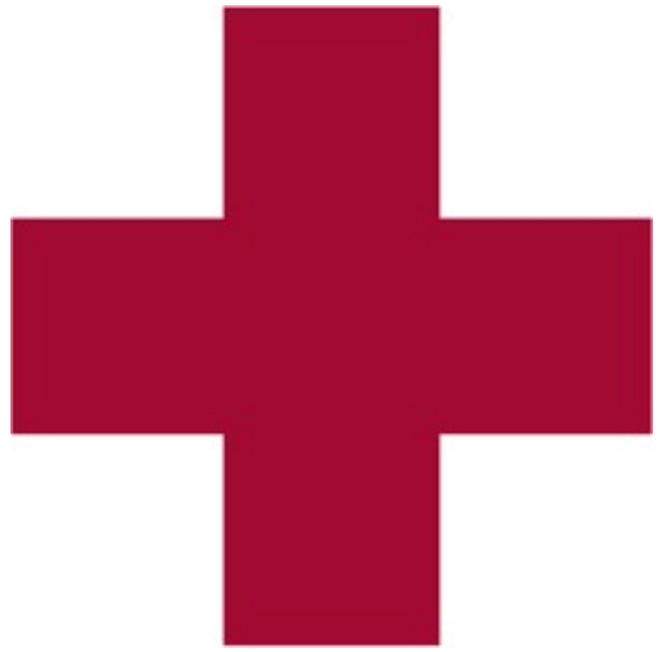

VAIO


VAIO

NOKIA

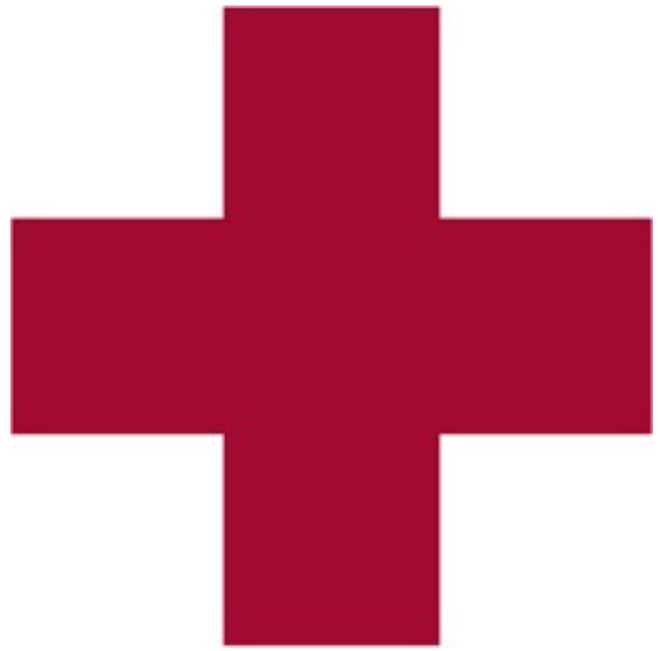
SAMSUNG


EVIDENCE
EVIDENCE
EVIDENCE



Forensic Triage:

Acquire evidence **quickly**, **accurately**, and **on-scene**.

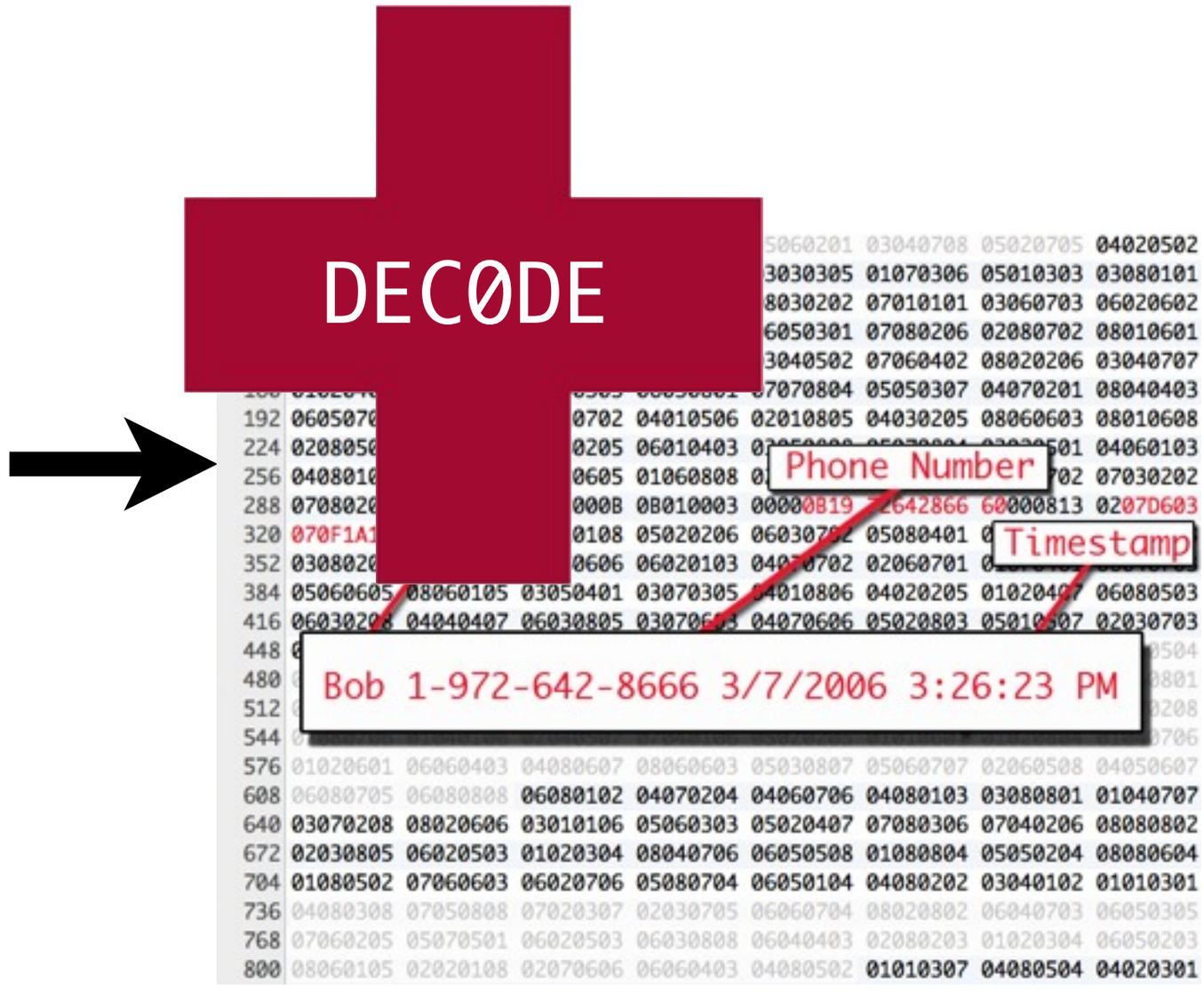


Forensic Triage:

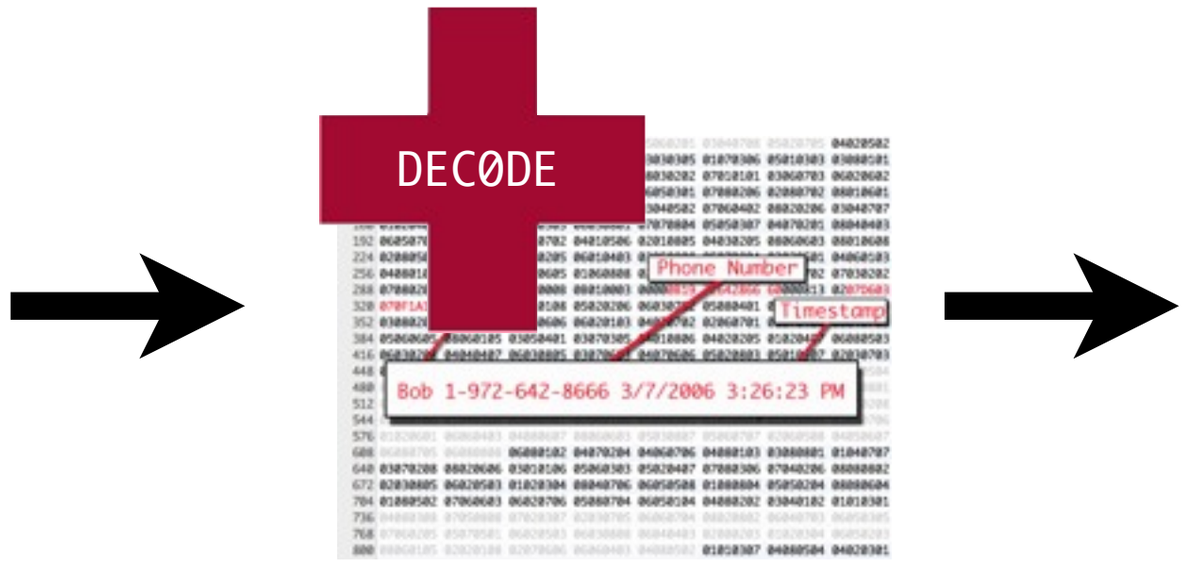
Acquire evidence **quickly**, **accurately**, and **on-scene**.

> Done before a full examination

DEC0DE: Forensic Triage for Phones

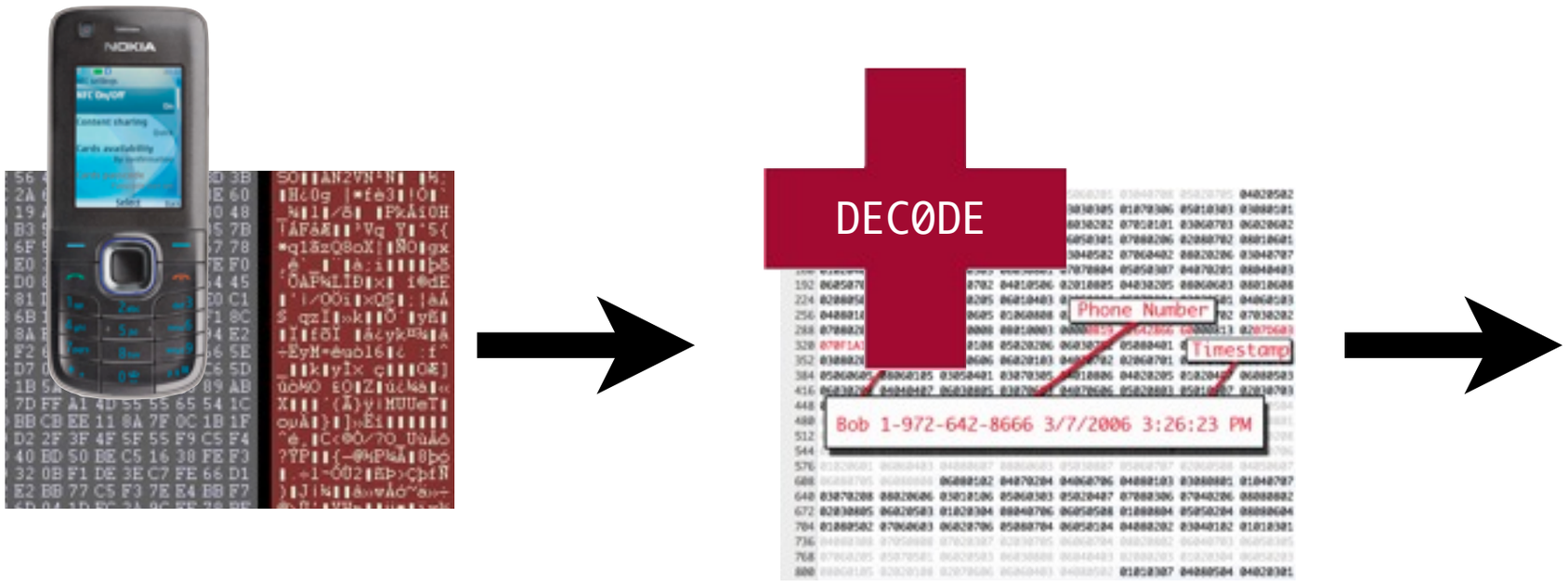


DEC0DE: Forensic Triage for Phones



DEC0DE: Forensic Triage for Phones

DEC0DE: Forensic Triage for Phones



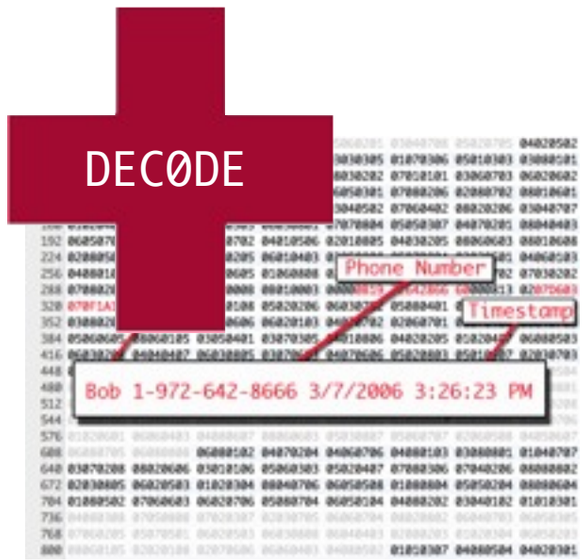
DEC0DE: Forensic Triage for Phones

SMS Text Messages

Call Logs

Address Book		
Robert	Walls	443-380-2823
Erik	Learned-Miller	505-309-3769
Brian	Levine	718-774-8034
William	Enck	714-206-4569
David	Wagner	276-251-8823
Ian	Goldberg	909-994-6768

DEC0DE: Forensic Triage for Phones



SMS Text Messages

Call Logs

Address Book

Robert	Walls	443-380-2823
Erik	Learned-Miller	505-309-3769
Brian	Levine	718-774-8034
William	Enck	714-206-4569
David	Wagner	276-251-8823
Ian	Goldberg	909-994-6768

Why phones?

Phones record our lives.

Phones contain evidence.


ACE MOBILE
acer
ALCATEL
alltel
AMOI
Anycool
ASUS
AUDIOVOX
AURO
AXIA
Bell
BenQ-SIEMENS
BLU
Camangi
CASIO
CECT 中电手机
Changjiang
China Clones
MEDIATEK
based phones
cingular
CONCORD
cricket
dalab
DELL
doopoo
doro
Duet
EM
EMOBILE
ERICSSON
ETEN
E-TEN
EVERTEK
TECHNOLOGY FOREVER
fly
garmin asus
GENERAL
MOBILE
G'FIVE
GIGABYTE
GiONEE金立
GLOBAL HIGH TECH
Google
G-TIDE
MOBILE PHONE
HITACHI
Inspire the Next
HKC
hp
htc
HUAWEI
i-mate
iNQ
iPro
i-teq
izenMobile
JCB
K-Touch
Keep in Touch
KYOCERA
lenovo
LG
MAXWEST
micromax
mio
MOTOROLA
MWg

Proprietary OS + Little Documentation

= Unknown Formats


Nokia 1006


Nokia 1100


Nokia 1100b


Nokia 1101


Nokia 1108


Nokia 1110


Nokia 1110i


Nokia 1112


Nokia 1112b


Nokia 1116


Nokia 1


5-2

Proprietary OS + Little Documentation

= Unknown Formats


Nokia 1650


Nokia 1650b
(Untested)


Nokia 1661-2


Nokia 2115i


Nokia 2116i


Nokia 2118
(Untested)


Nokia 2125i


Nokia 2126i


Nokia 2128i


Nokia 2220 slide


Nokia 2270


Nokia 2272
(Untested)


Nokia 2275
(Untested)


Nokia 2280
(Untested)


Nokia 2285


Nokia 2300


Nokia 2310


Nokia 2320 classic


Nokia 2323 classic


Nokia 2330 classic

Nokia 2366i

Nokia 2600

Nokia 2600 classic

Nokia 2600b

Triage options now?

Option 1:

Browsing

Option 2:

Commercial tools

Option 1:
Browsing

Drawbacks

Option 1:
Browsing

Drawbacks

- May not be possible

Option 1:
Browsing

Drawbacks

- May not be possible
- Modifies the phone

Option 1:
Browsing

Drawbacks

- May not be possible
- Modifies the phone
- Misses important information

Option 2:
Commercial Tools

Drawbacks

Option 2:
Commercial Tools

Drawbacks

- Cost Prohibitive

Option 2:
Commercial Tools

Drawbacks

- Cost Prohibitive
- Does not support all phones

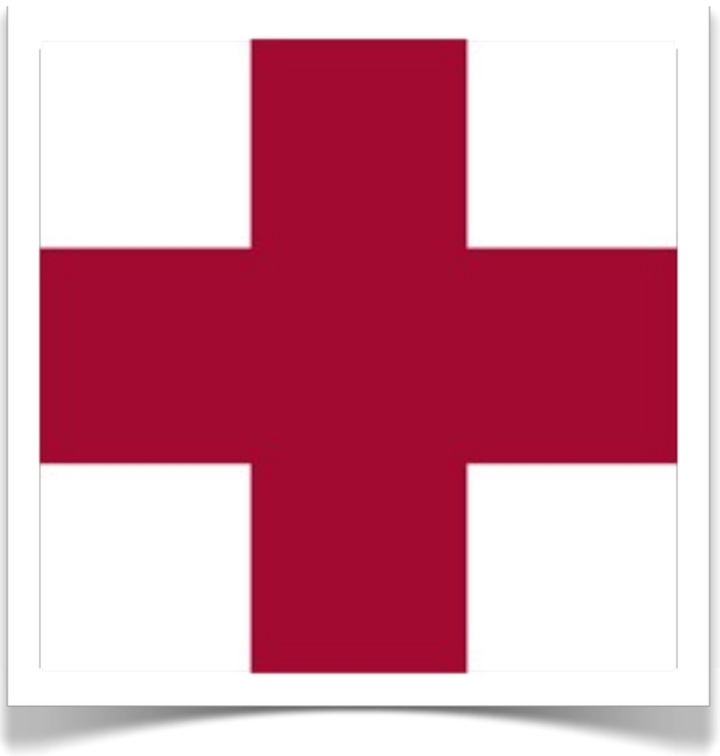
Option 2:
Commercial Tools

Drawbacks

> Cost Prohibitive

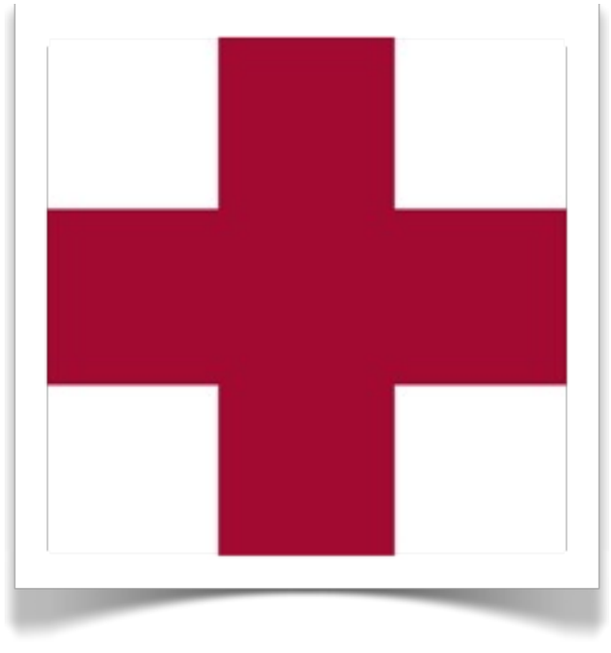
> Does not support all phones

> Still misses important information!



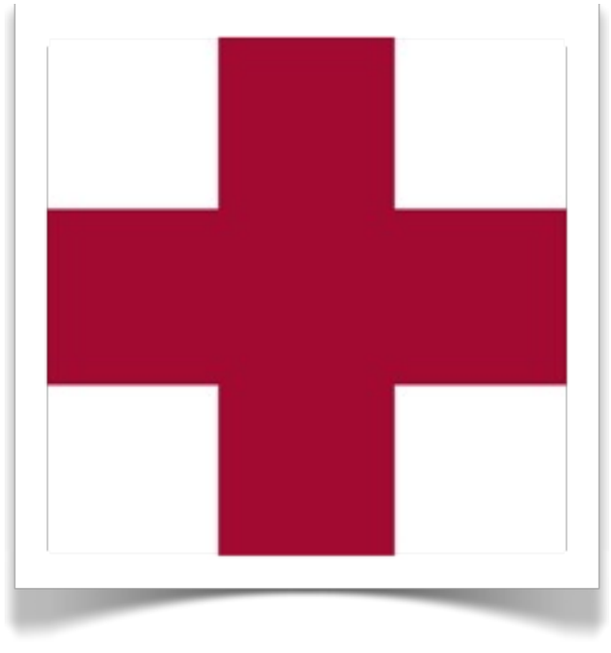
Option 3:

DEC0DE



Option 3:
DEC0DE

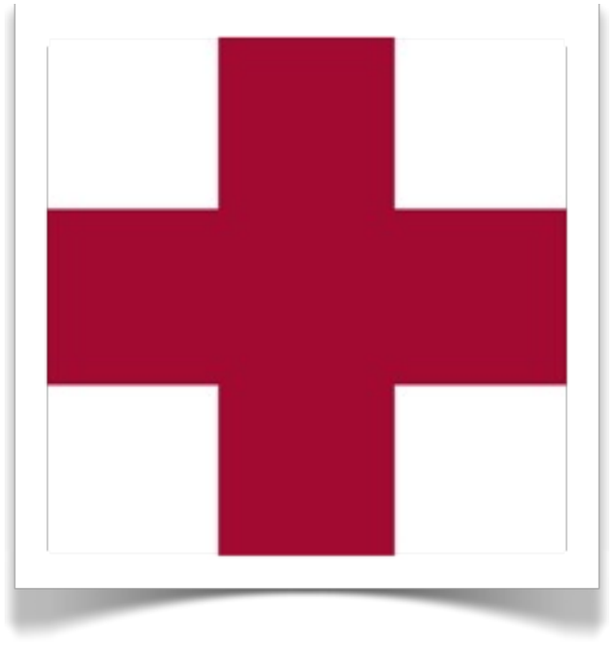
Advantages



Option 3:
DEC0DE

Advantages

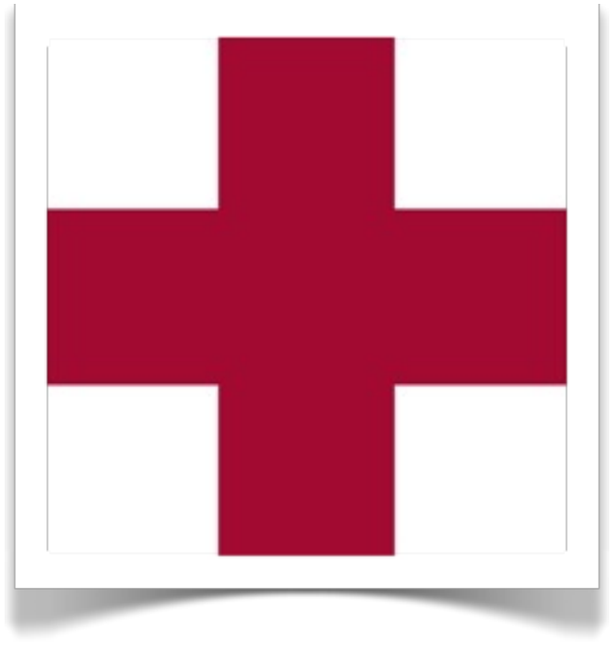
> Extracts information directly from storage



Option 3:
DEC0DE

Advantages

- Extracts information directly from storage
- File system and OS agnostic



Option 3:
DEC0DE

Advantages

- Extracts information directly from storage
- File system and OS agnostic
- Quick (< 20 minutes)

```
  0 02030108 04010404 02040301 06020106 05060201 03040708 05020705 04020502
 32 08050303 03010401 01020206 05060304 03030305 01070306 05010303 03080101
 64 07020101 07060404 02020802 03020407 08030202 07010101 03060703 06020602
 96 01020803 06020506 08070308 08020403 06050301 07080206 02080702 08010601
128 04070102 02010206 06050308 07010201 03040502 07060402 08020206 03040707
160 01020401 02020205 04050303 06030801 07070804 05050307 04070201 08040403
192 06050705 06010801 03010702 04010506 02010805 04030205 08060603 08010608
224 02080502 04040406 07080205 06010403 03050808 05070804 03020501 04060103
256 04080104 05040207 06020605 01060808 02030804 01060403 01010702 07030202
288 07080207 0042006F 0062000B 0B010003 00000B19 72642866 60000813 0207D603
320 070F1A17 05010703 05010108 05020206 06030702 05080401 07050304 04020403
352 03080205 05020103 07050606 06020103 04070702 02060701 01070401 06040703
384 05060605 08060105 03050401 03070305 04010806 04020205 01020407 06080503
416 06030208 04040407 06030805 03070603 04070606 05020803 05010507 02030703
448 08010301 06030108 06020802 03050303 04020308 03080704 03070404 08060504
480 04060502 04060602 06060803 02070108 07020301 08080801 02070608 08060801
512 05050606 08010807 04020205 01040305 06040507 03040404 06050105 08070208
544 07080706 01040106 02040507 07040106 05020205 01010601 01020804 01040706
576 01020601 06060403 04080607 08060603 05030807 05060707 02060508 04050607
608 06080705 06080808 06080102 04070204 04060706 04080103 03080801 01040707
640 03070208 08020606 03010106 05060303 05020407 07080306 07040206 08080802
672 02030805 06020503 01020304 08040706 06050508 01080804 05050204 08080604
704 01080502 07060603 06020706 05080704 06050104 04080202 03040102 01010301
736 04080308 07050808 07020307 02030705 06060704 08020802 06040703 06050305
768 07060205 05070501 06020503 06030808 06040403 02080203 01020304 06050203
800 08060105 02020108 02070606 06060403 04080502 01010307 04080504 04020301
```

```
  0 02030108 04010404 02040301 06020106 05060201 03040708 05020705 04020502
 32 08050303 03010401 01020206 05060304 03030305 01070306 05010303 03080101
 64 07020101 07060404 02020802 03020407 08030202 07010101 03060703 06020602
 96 01020803 06020506 08070308 08020403 06050301 07080206 02080702 08010601
128 04070102 02010206 06050308 07010201 03040502 07060402 08020206 03040707
160 01020401 02020205 04050303 06030801 07070804 05050307 04070201 08040403
192 06050705 06010801 03010702 04010506 02010805 04030205 08060603 08010608
224 02080502 04040406 07080205 06010403 03050808 05070804 03020501 04060103
256 04080104 05040207 06020605 01060808 02030804 01060403 01010702 07030202
288 07080207 0042006F 0062000B 0B010003 00000B19 72642866 60000813 0207D603
320 070F1A17 05010703 05010108 05020206 06030702 05080401 07050304 04020403
352 03080205 05020103 07050606 06020103 04070702 02060701 01070401 06040703
384 05060605 08060105 03050401 03070305 04010806 04020205 01020407 06080503
416 06030208 04040407 06030805 03070603 04070606 05020803 05010507 02030703
448 08010301 06030108 06020802 03050303 04020308 03080704 03070404 08060504
480 04060502 04060602 06060803 02070108 07020301 08080801 02070608 08060801
512 05050606 08010807 04020205 01040305 06040507 03040404 06050105 08070208
544 07080706 01040106 02040507 07040106 05020205 01010601 01020804 01040706
576 01020601 06060403 04080607 08060603 05030807 05060707 02060508 04050607
608 06080705 06080808 06080102 04070204 04060706 04080103 03080801 01040707
640 03070208 08020606 03010106 05060303 05020407 07080306 07040206 08080802
672 02030805 06020503 01020304 08040706 06050508 01080804 05050204 08080604
704 01080502 07060603 06020706 05080704 06050104 04080202 03040102 01010301
736 04080308 07050808 07020307 02030705 06060704 08020802 06040703 06050305
768 07060205 05070501 06020503 06030808 06040403 02080203 01020304 06050203
800 08060105 02020108 02070606 06060403 04080502 01010307 04080504 04020301
```

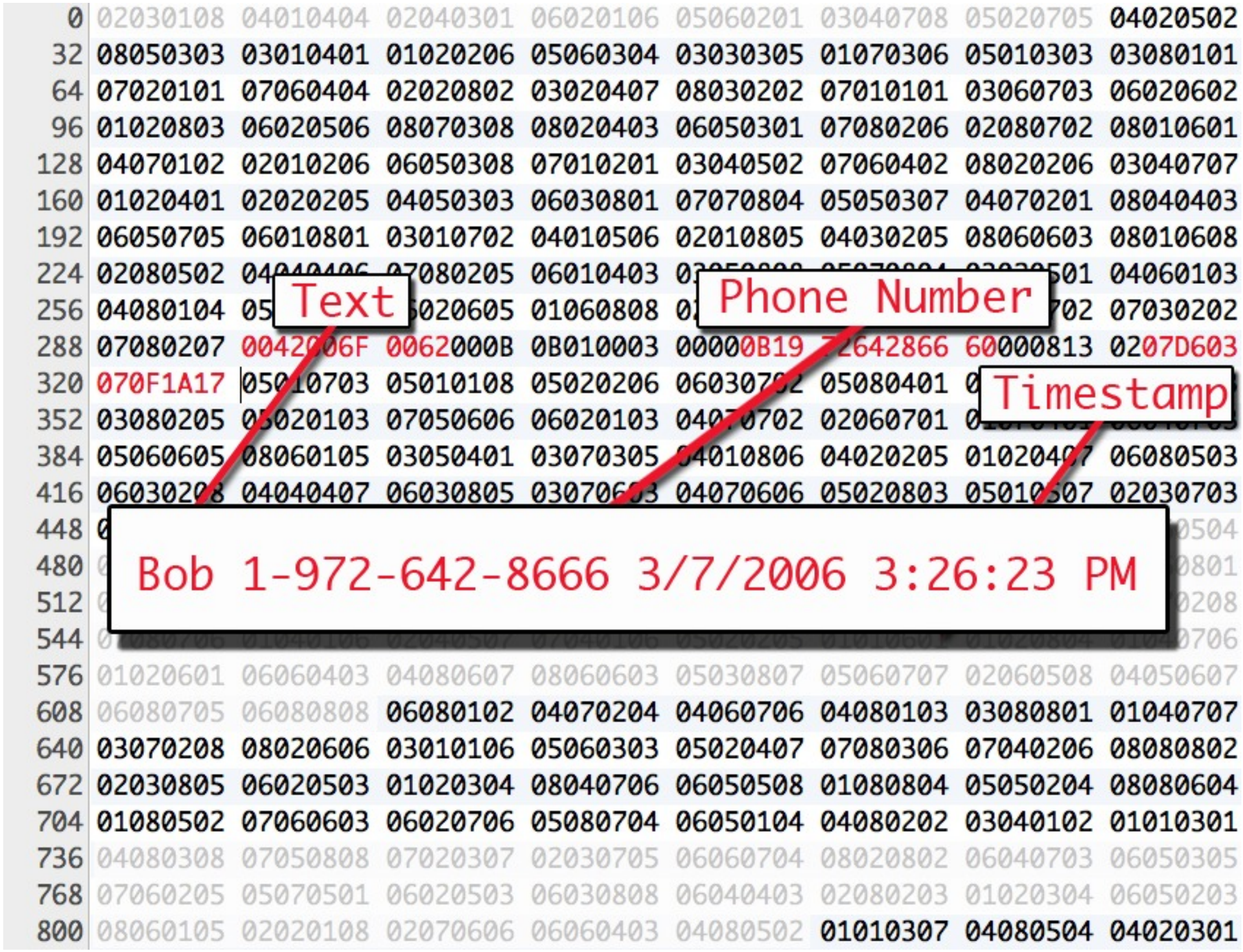
```
  0 02030108 04010404 02040301 06020106 05060201 03040708 05020705 04020502
 32 08050303 03010401 01020206 05060304 03030305 01070306 05010303 03080101
 64 07020101 07060404 02020802 03020407 08030202 07010101 03060703 06020602
 96 01020803 06020506 08070308 08020403 06050301 07080206 02080702 08010601
128 04070102 02010206 06050308 07010201 03040502 07060402 08020206 03040707
160 01020401 02020205 04050303 06030801 07070804 05050307 04070201 08040403
192 06050705 06010801 03010702 04010506 02010805 04030205 08060603 08010608
224 02080502 04040406 07080205 06010403 03050808 05070804 03020501 04060103
256 04080104 05040207 06020605 01060808 02030804 01060403 01010702 07030202
288 07080207 0042006F 0062000B 0B010003 00000B19 72642866 60000813 0207D603
320 070F1A17 05010703 05010108 05020206 06030702 05080401 07050304 04020403
352 03080205 05020103 07050606 06020103 04070702 02060701 01070401 06040703
384 05060605 08060105 03050401 03070305 04010806 04020205 01020407 06080503
416 06030208 04040407 06030805 03070603 04070606 05020803 05010507 02030703
448 08010301 06030108 06020802 03050303 04020308 03080704 03070404 08060504
480 04060502 04060602 06060803 02070108 07020301 08080801 02070608 08060801
512 05050606 08010807 04020205 01040305 06040507 03040404 06050105 08070208
544 07080706 01040106 02040507 07040106 05020205 01010601 01020804 01040706
576 01020601 06060403 04080607 08060603 05030807 05060707 02060508 04050607
608 06080705 06080808 06080102 04070204 04060706 04080103 03080801 01040707
640 03070208 08020606 03010106 05060303 05020407 07080306 07040206 08080802
672 02030805 06020503 01020304 08040706 06050508 01080804 05050204 08080604
704 01080502 07060603 06020706 05080704 06050104 04080202 03040102 01010301
736 04080308 07050808 07020307 02030705 06060704 08020802 06040703 06050305
768 07060205 05070501 06020503 06030808 06040403 02080203 01020304 06050203
800 08060105 02020108 02070606 06060403 04080502 01010307 04080504 04020301
```

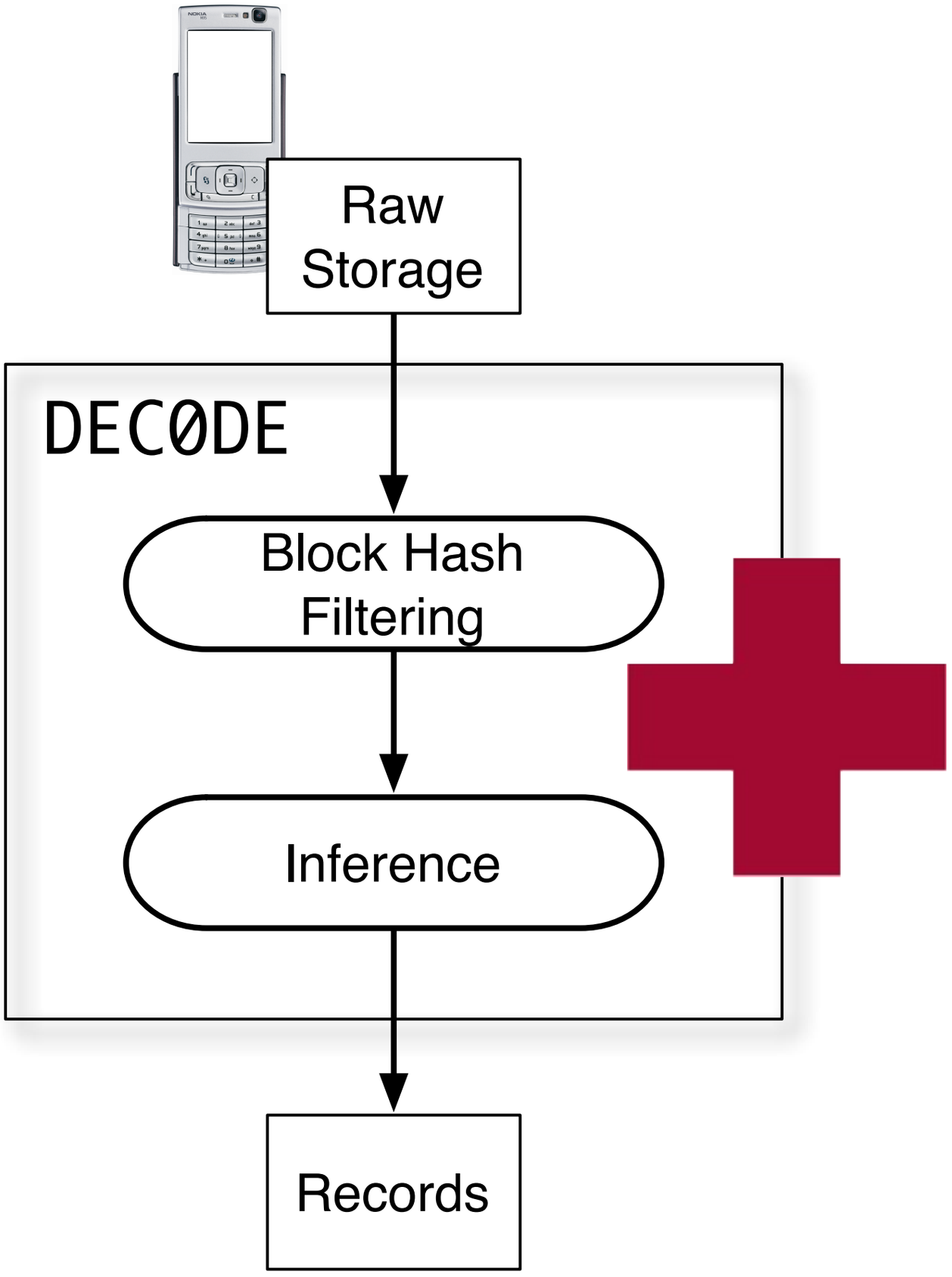
```
  0 02030108 04010404 02040301 06020106 05060201 03040708 05020705 04020502
 32 08050303 03010401 01020206 05060304 03030305 01070306 05010303 03080101
 64 07020101 07060404 02020802 03020407 08030202 07010101 03060703 06020602
 96 01020803 06020506 08070308 08020403 06050301 07080206 02080702 08010601
128 04070102 02010206 06050308 07010201 03040502 07060402 08020206 03040707
160 01020401 02020205 04050303 06030801 07070804 05050307 04070201 08040403
192 06050705 06010801 03010702 04010506 02010805 04030205 08060603 08010608
224 02080502 04040406 07080205 06010403 03050808 05070804 03030501 04060103
256 04080104 05020605 01060808 02020702 07030202
288 07080207 0042006F 0062000B 0B010003 00000B19 72642866 60000813 0207D603
320 070F1A17 05010703 05010108 05020206 06030702 05080401 0
352 03080205 05020103 07050606 06020103 04070702 02060701 0
384 05060605 08060105 03050401 03070305 04010806 04020205 01020407 06080503
416 06030208 04040407 06030805 03070603 04070606 05020803 05010507 02030703
448 08010301 06030108 06020802 03050303 04020308 03080704 03070404 08060504
480 04060502 04060602 06060803 02070108 07020301 08080801 02070608 08060801
512 05050606 08010807 04020205 01040305 06040507 03040404 06050105 08070208
544 07080706 01040106 02040507 07040106 05020205 01010601 01020804 01040706
576 01020601 06060403 04080607 08060603 05030807 05060707 02060508 04050607
608 06080705 06080808 06080102 04070204 04060706 04080103 03080801 01040707
640 03070208 08020606 03010106 05060303 05020407 07080306 07040206 08080802
672 02030805 06020503 01020304 08040706 06050508 01080804 05050204 08080604
704 01080502 07060603 06020706 05080704 06050104 04080202 03040102 01010301
736 04080308 07050808 07020307 02030705 06060704 08020802 06040703 06050305
768 07060205 05070501 06020503 06030808 06040403 02080203 01020304 06050203
800 08060105 02020108 02070606 06060403 04080502 01010307 04080504 04020301
```

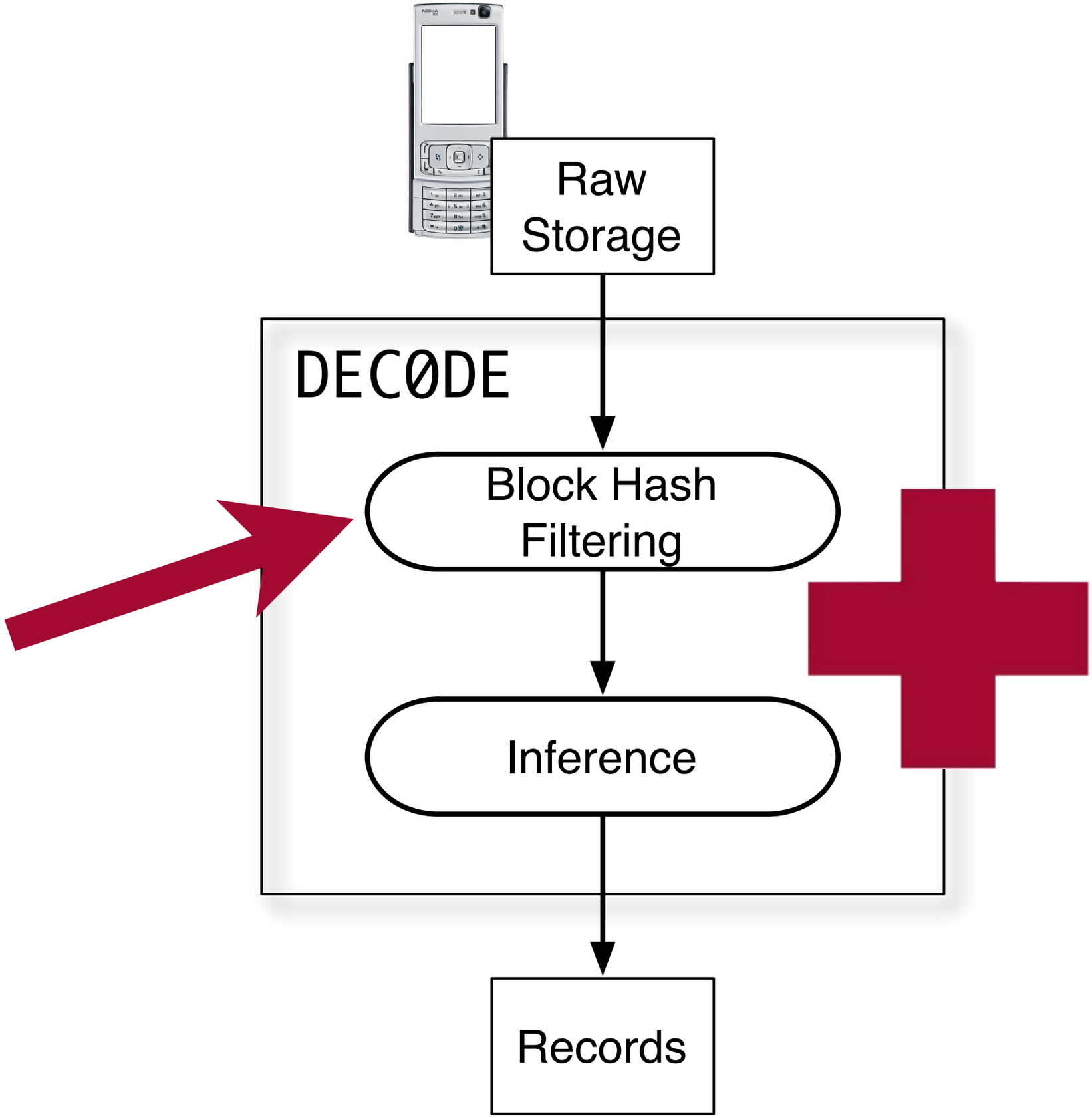
Text

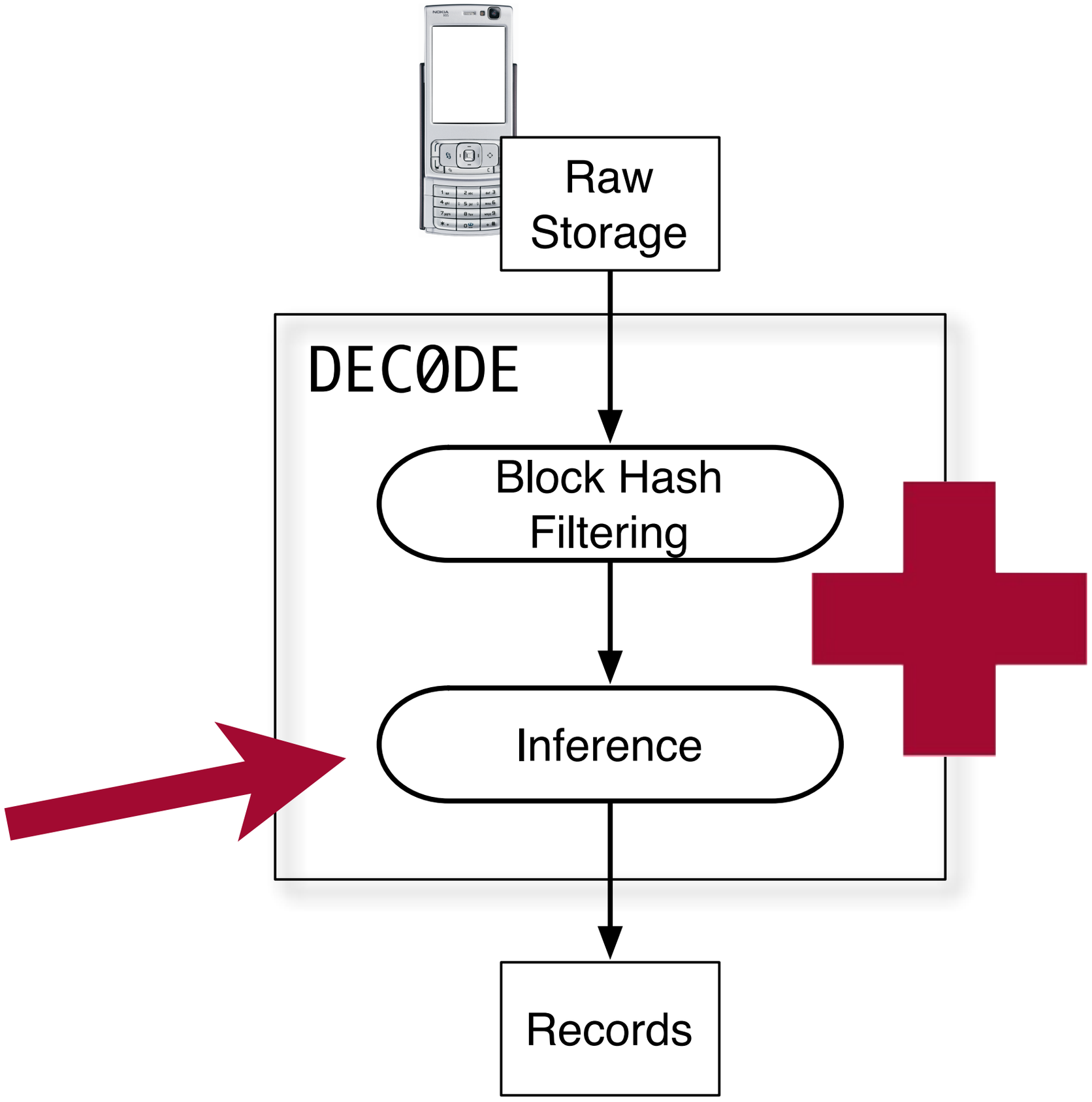
Phone Number

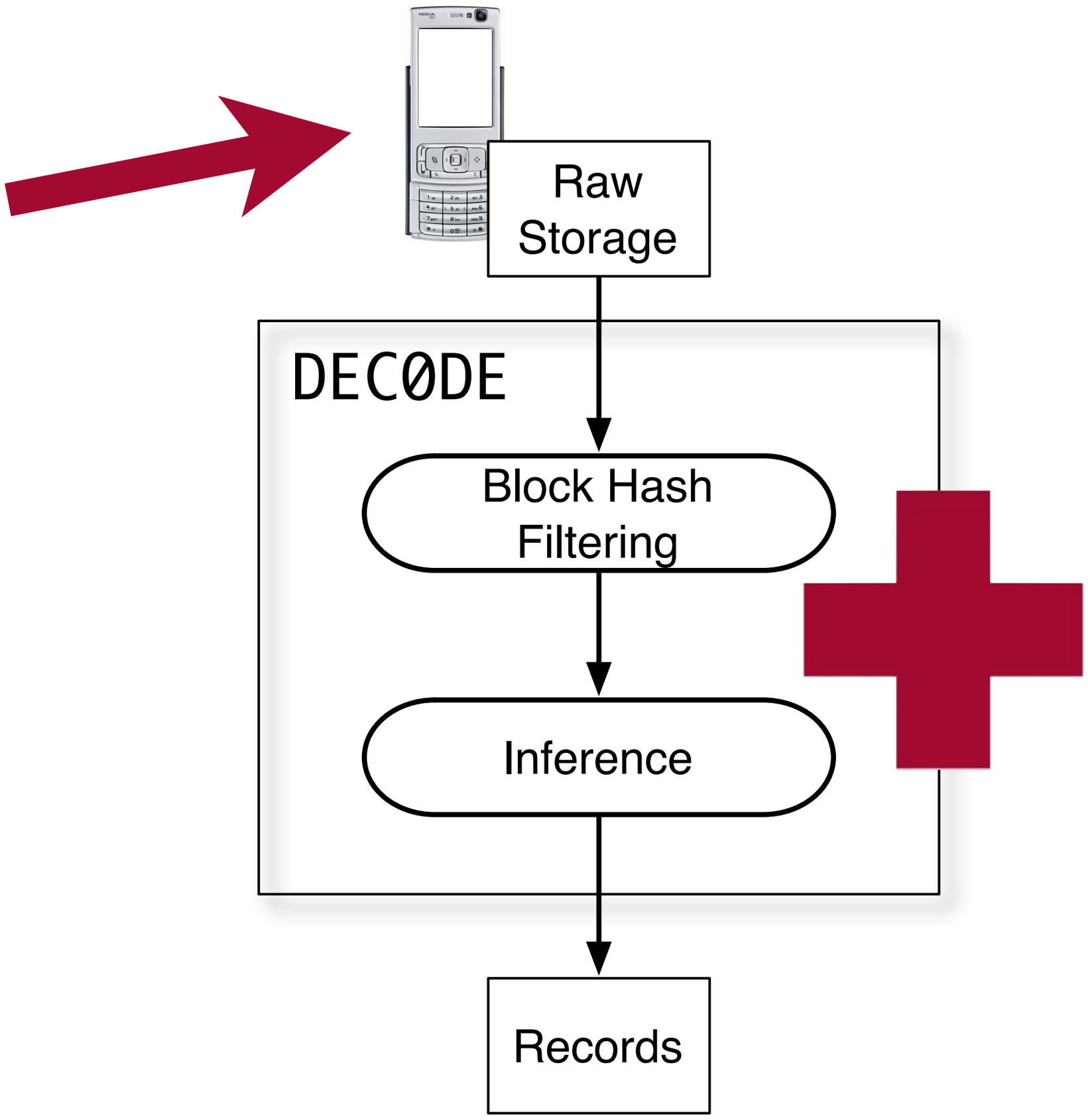
Timestamp

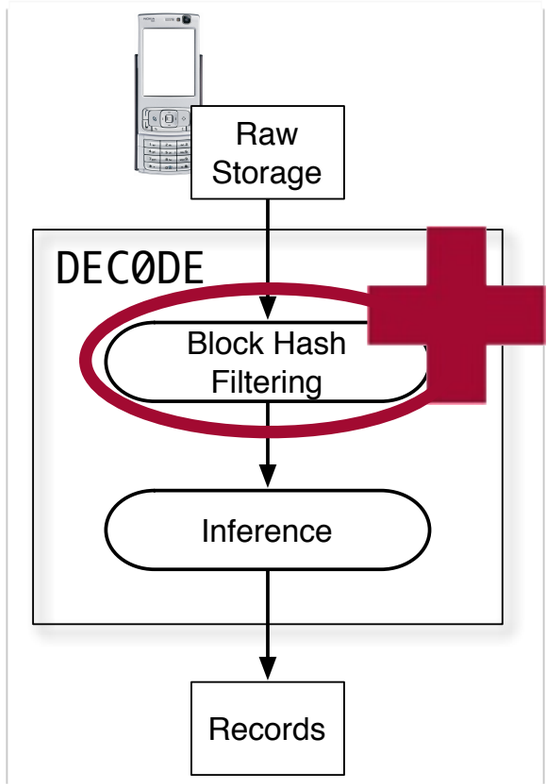

Text
Phone Number
Timestamp
Bob 1-972-642-8666 3/7/2006 3:26:23 PM


Raw
Storage
DEC0DE
Block Hash
Filtering
Inference
Records


Raw
Storage
DEC0DE
Block Hash
Filtering
Inference
Records


Raw
Storage
DEC0DE
Block Hash
Filtering
Inference
Records


Raw
Storage
DEC0DE
Block Hash
Filtering
Inference
Records

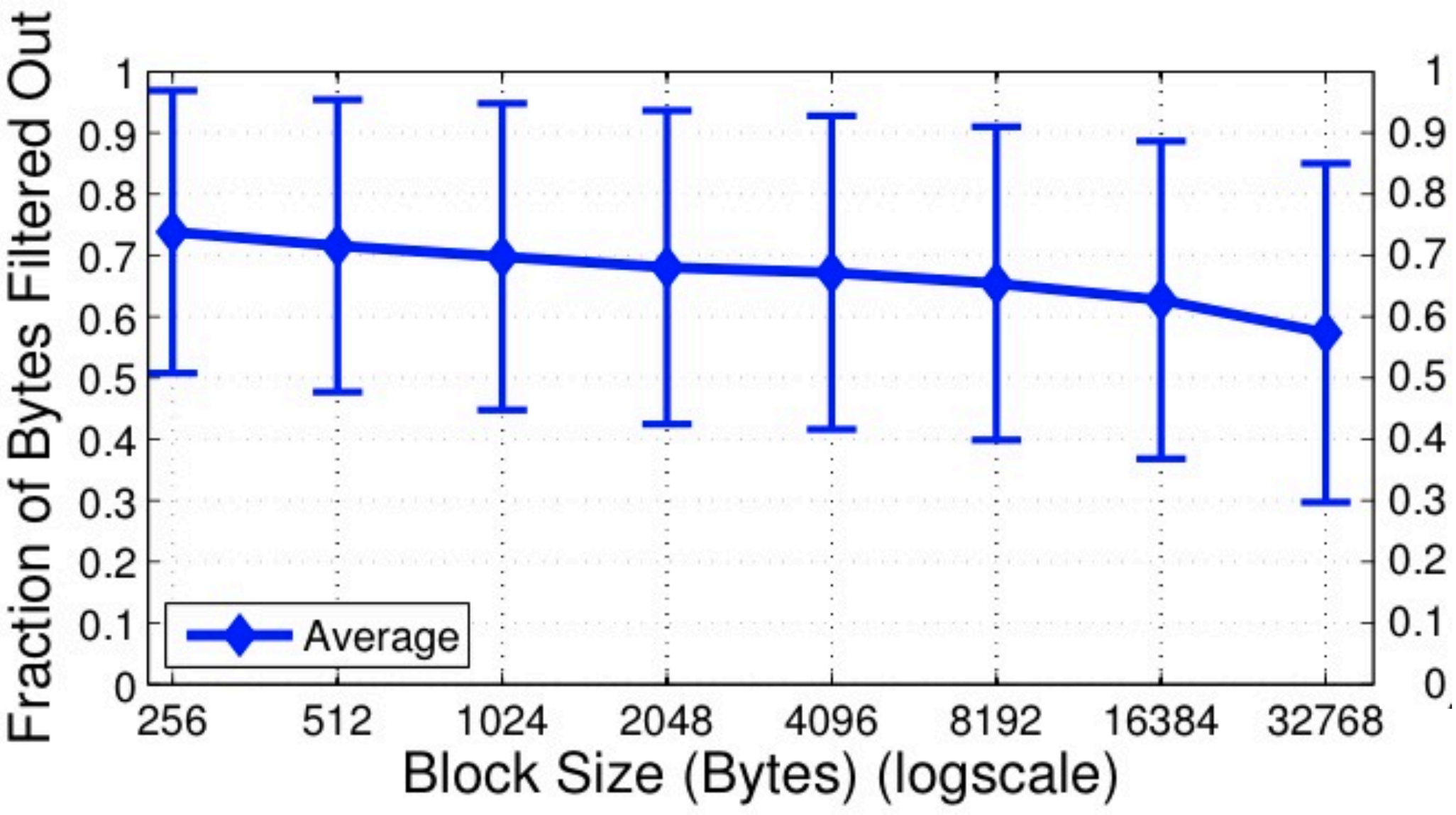


Component 1:
Block Hash Filtering

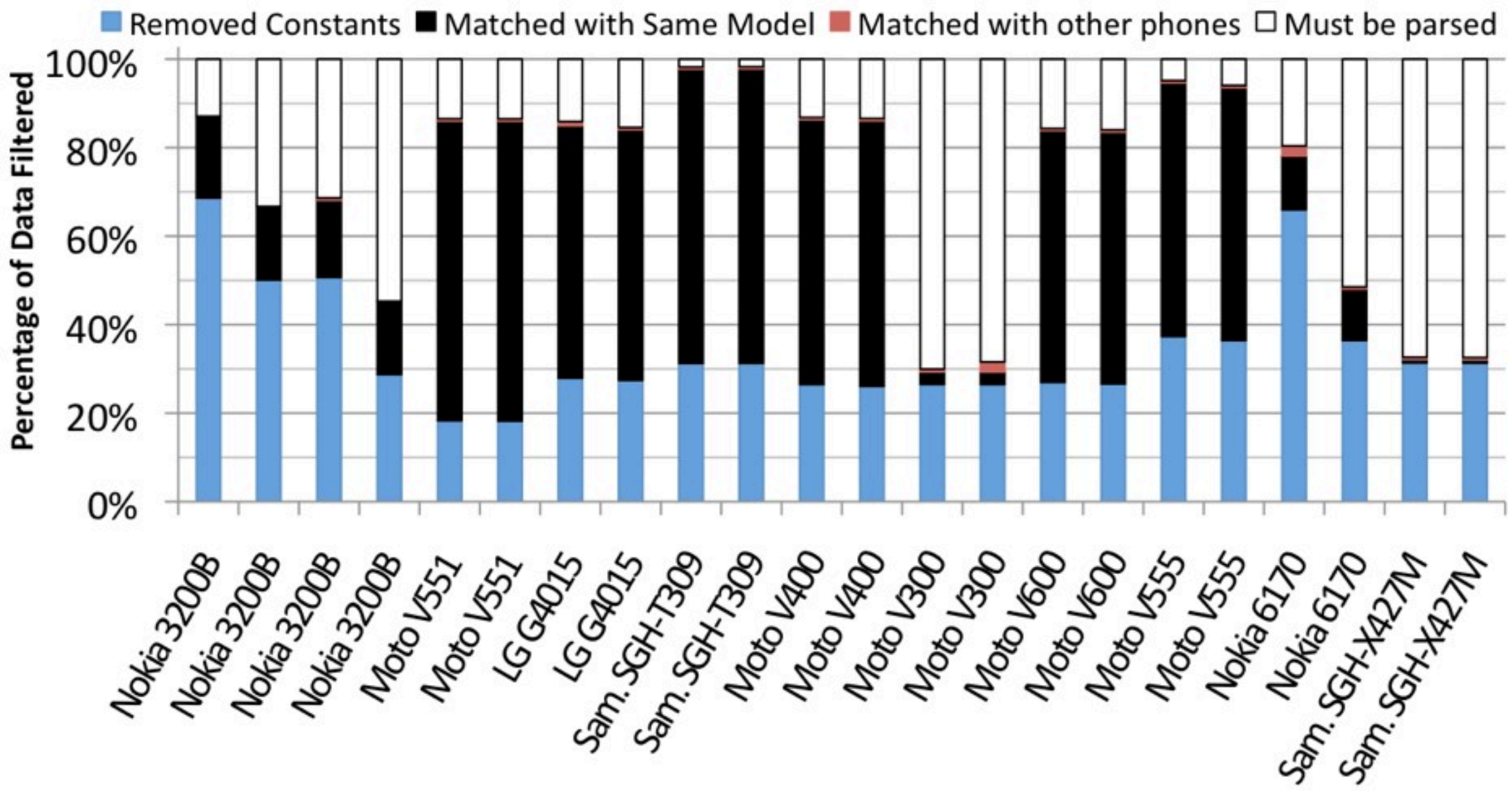
Process:

- Divide storage into blocks
- Compare block hash to library
- Filter duplicates

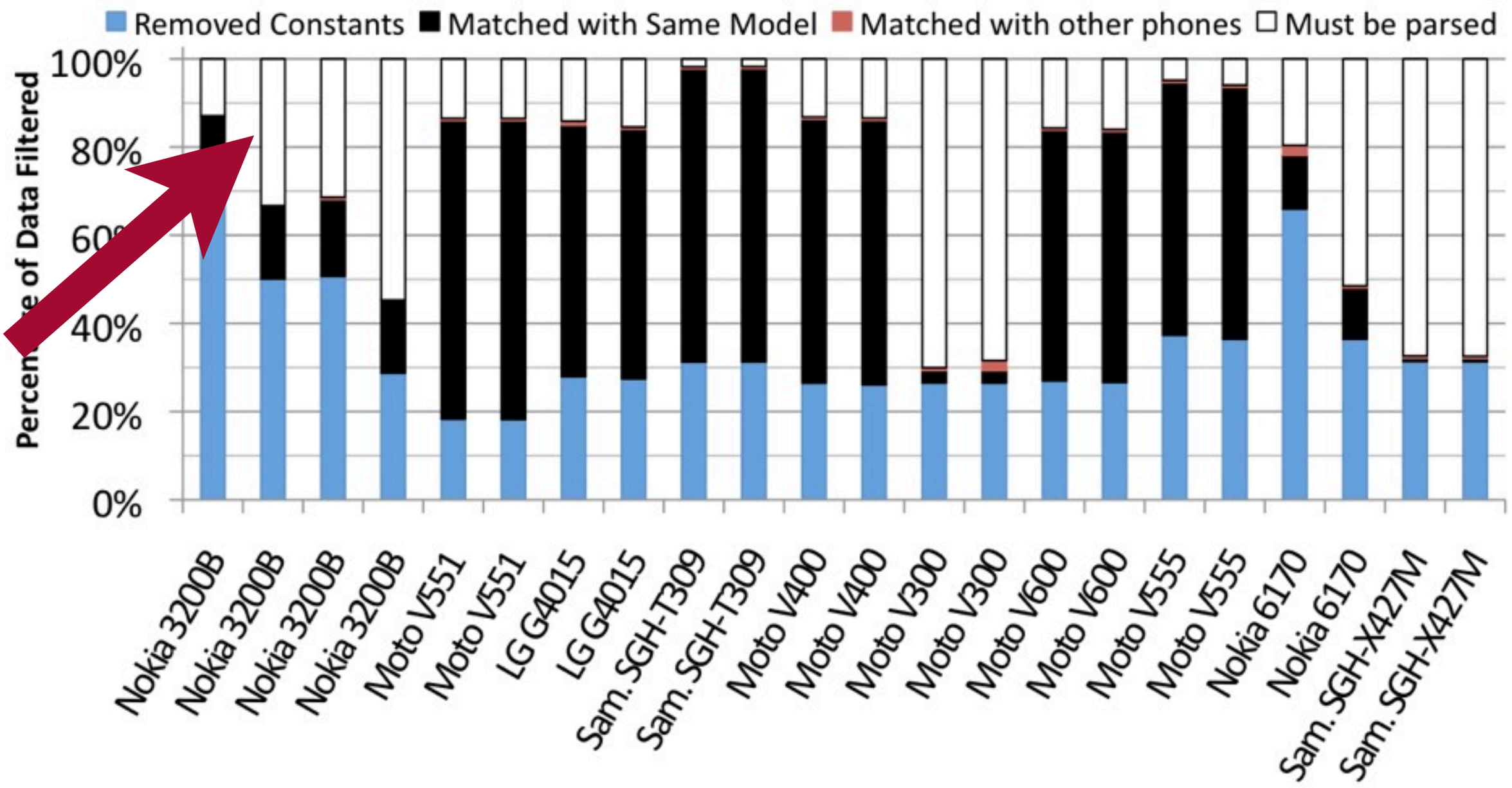
Evaluation: BHF



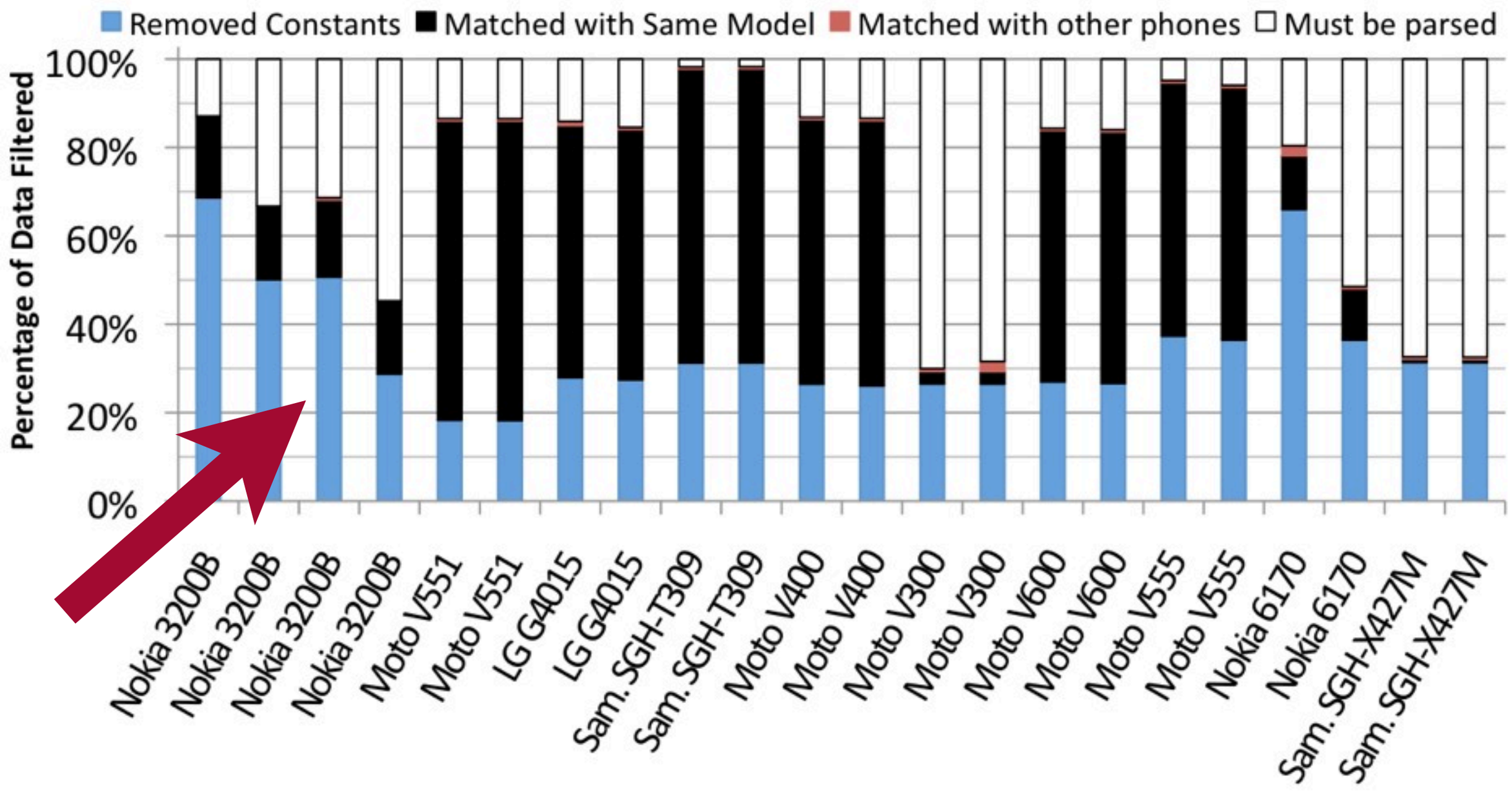
Evaluation: BHF



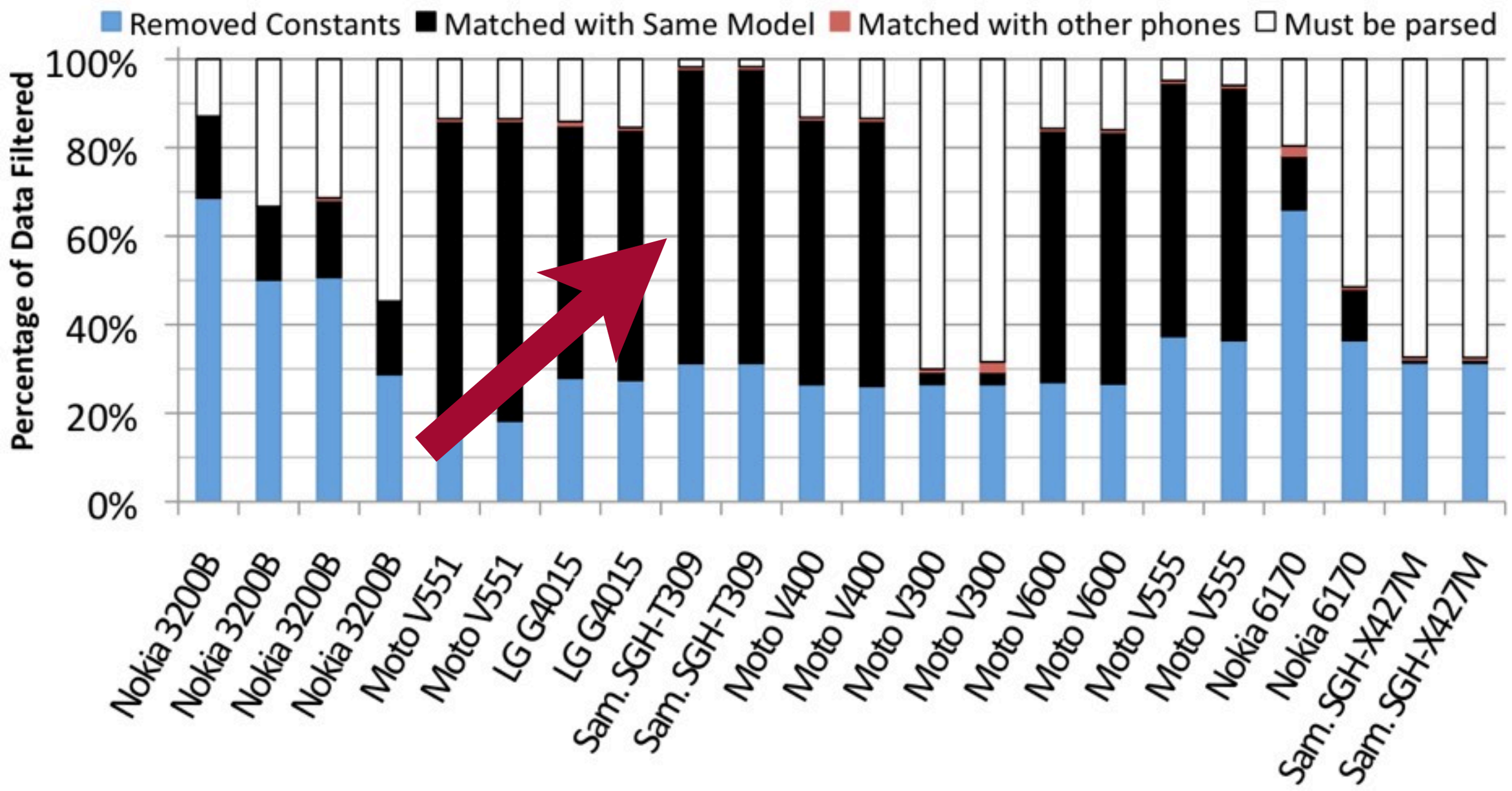
Evaluation: BHF



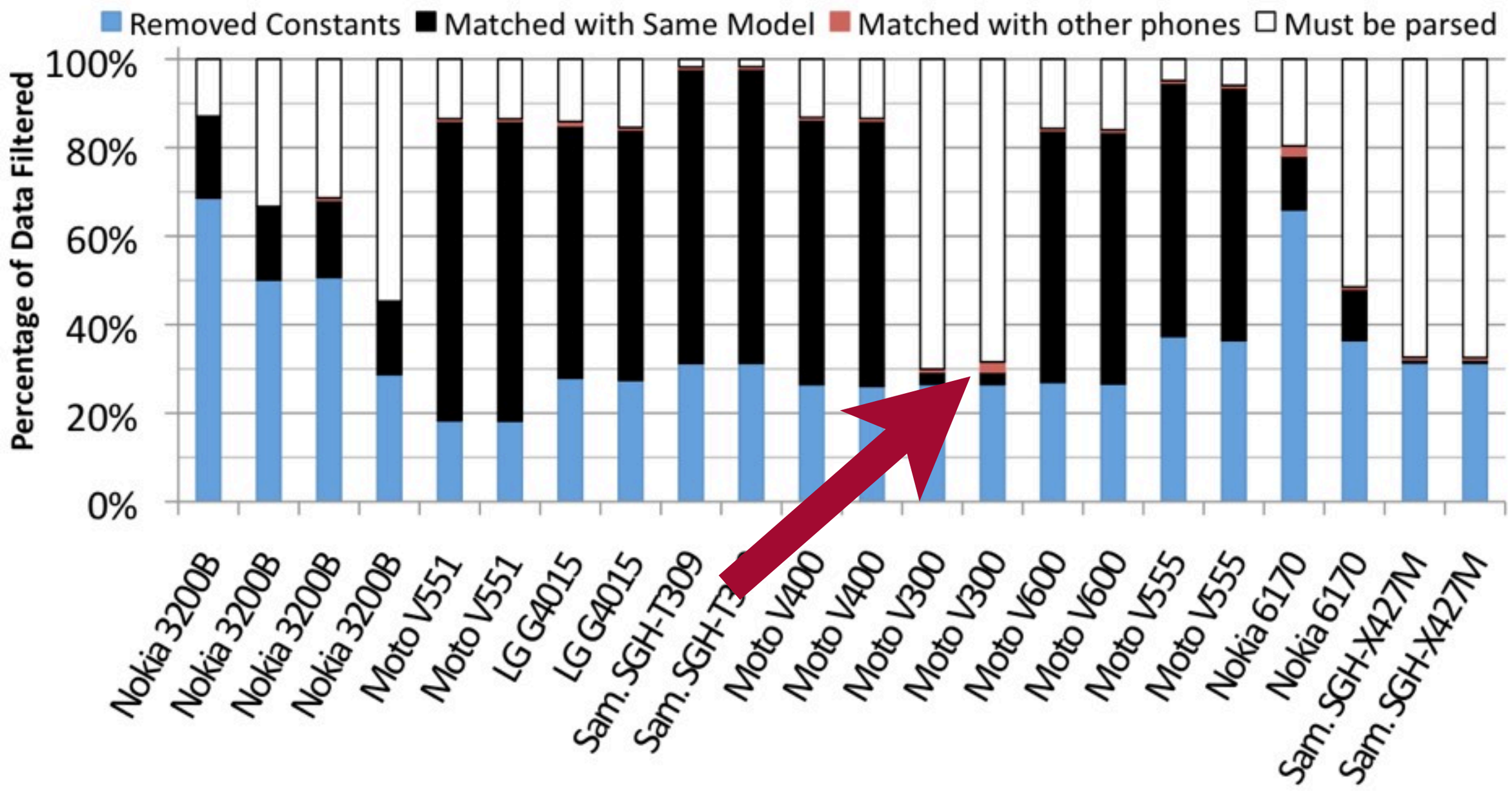
Evaluation: BHF



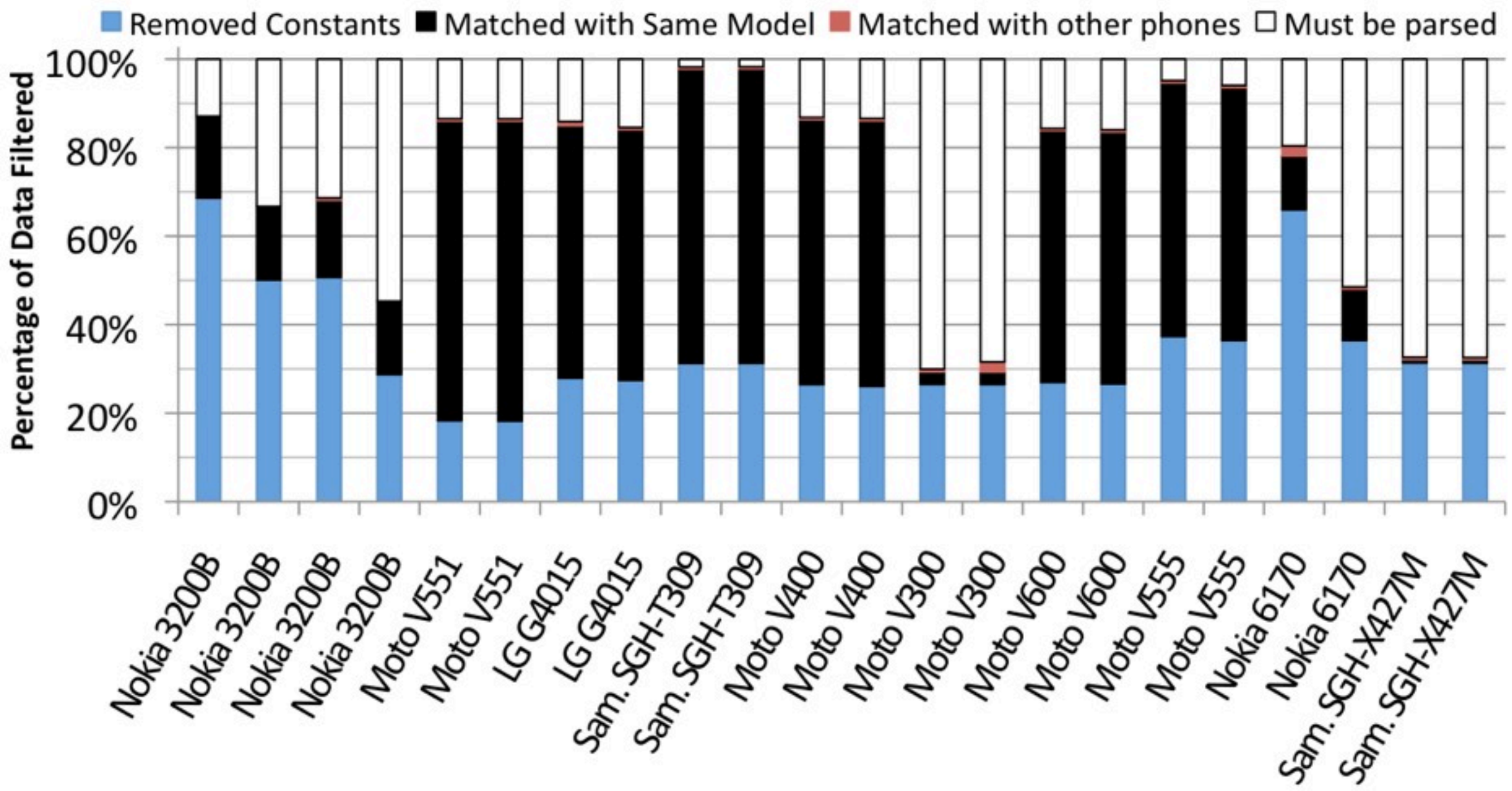
Evaluation: BHF

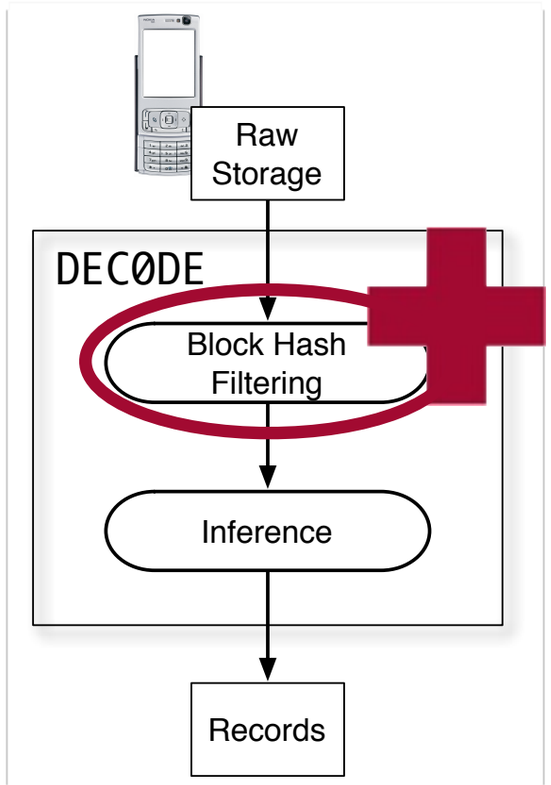


Evaluation: BHF



Evaluation: BHF



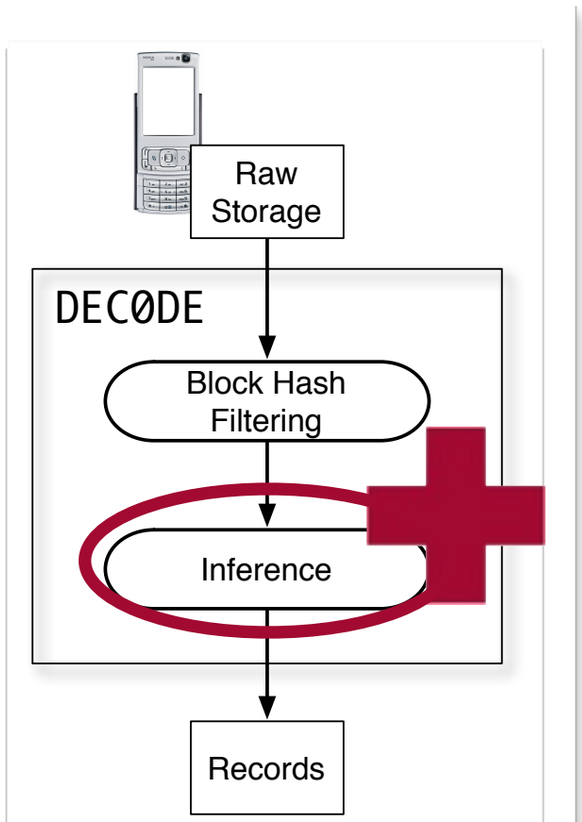


Component 1:
Block Hash Filtering

Evaluation Summary:

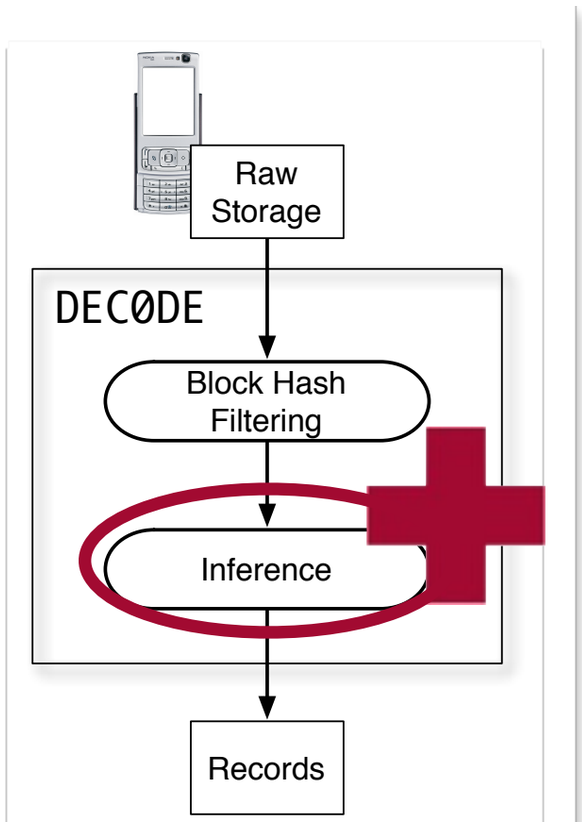
> Filtered 69% on average

> Lot of overlap between phones of same model



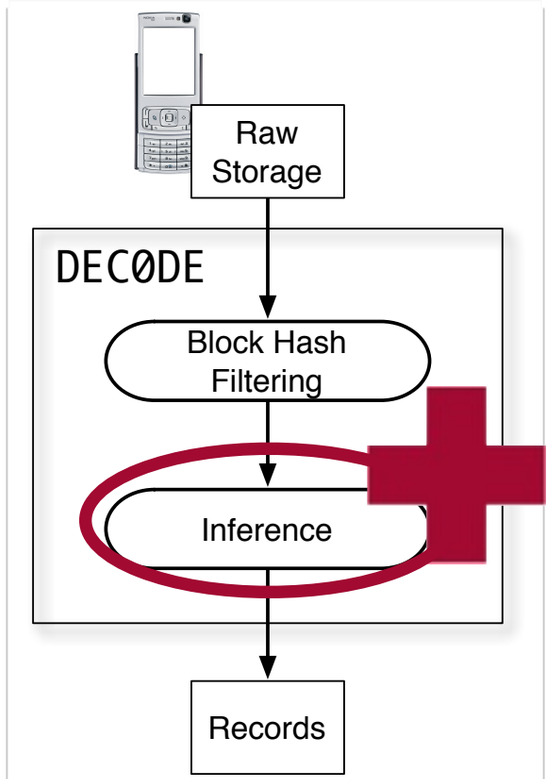
Inference?

Simple, just use regular expressions.



Inference?

~~Simple, just use regular expressions.~~



Component 2:

Inference

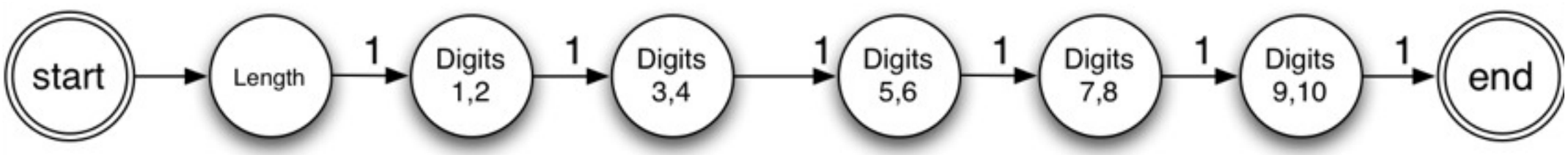
Process:

> Encode formats using Probabilistic Finite State Machines (PFSM)

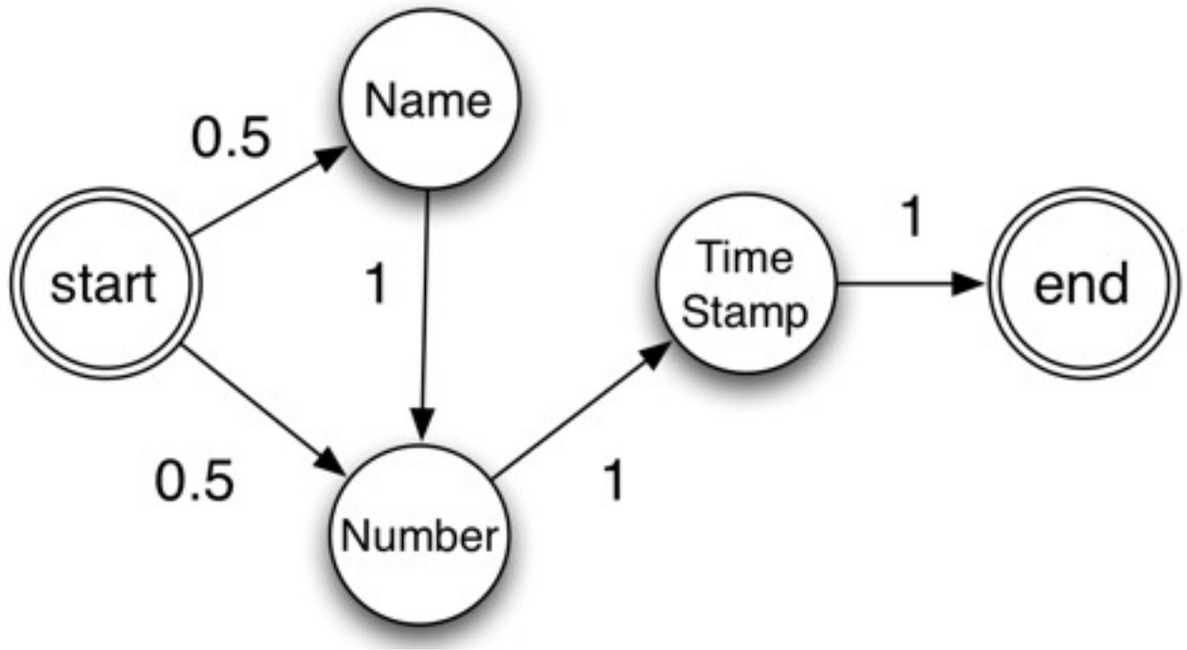
> Parse using Viterbi's Algorithm

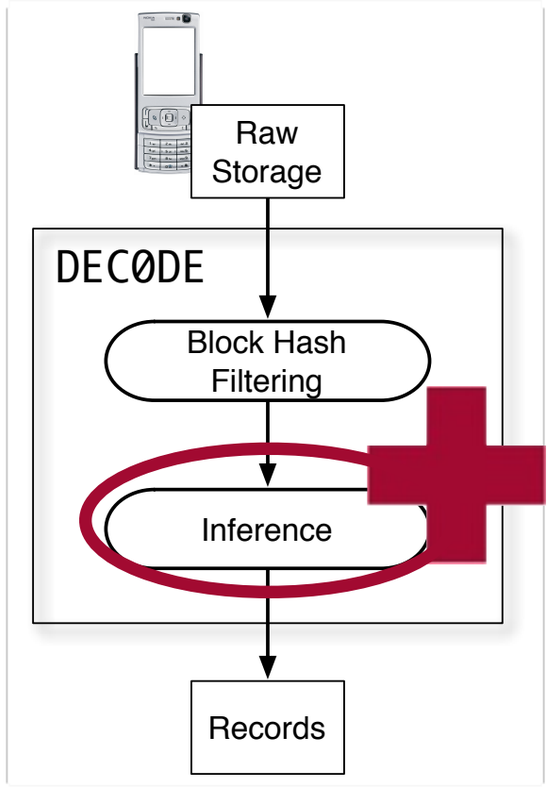
> Remove false positives using decision tree.

Phone number:



Call log:



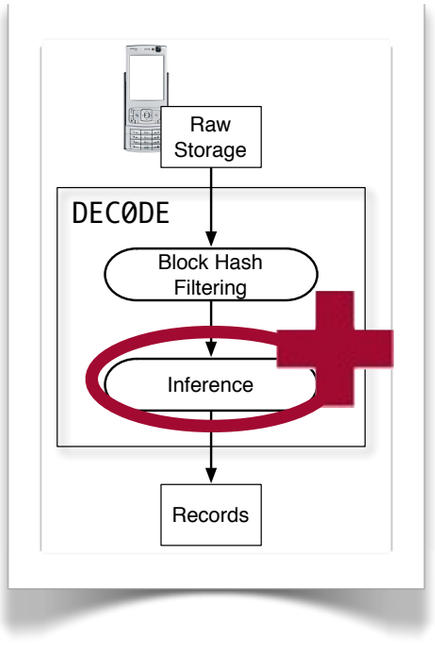


Component 2:

Inference

Post Processing:

- Simpler to encode certain features
- Reduces complexity of state machines
- Increases precision



Component 2:
Inference Evaluation

Process:

Step 0 > Pick phone set

Step 1 > Pick target record types

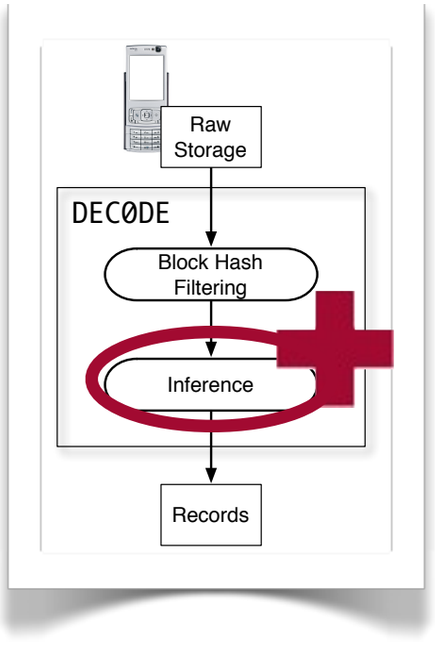
Step 2 > Manually create state machines

Step 3 > Acquire Raw Storage

Step 4 > Run DEC0DE

Step 0 > Pick phone set

Make	Model	Count	MB
PFSM Development Set			
Nokia	3200b	4	1.4
Motorola	V551	2	32.0
Samsung	SGH-T309	2	32.0
LG	G4015	2	48.0
Evaluation Set			
Motorola	V400	2	32.0
Motorola	V300	2	32.0
Motorola	V600	2	32.0
Motorola	V555	2	32.0
Nokia	6170	2	4.9
Samsung	SGH-X427M	2	16.0



Component 2:
Inference Evaluation

Process:

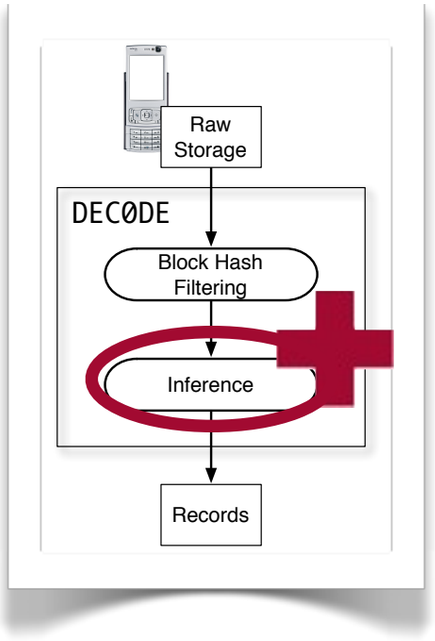
Step 0 > Pick phone set

Step 1 > Pick target record types

Step 2 > Manually create state machines

Step 3 > Acquire Raw Storage

Step 4 > Run DEC0DE



Component 2:
Inference Evaluation

Process:

Step 0 > Pick phone set

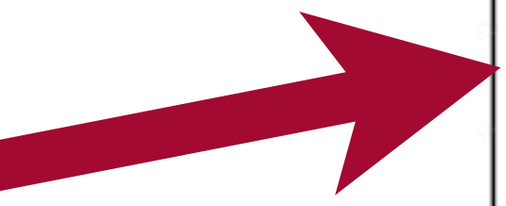
Step 1 > Pick target record types

Step 2 > Manually create state machines

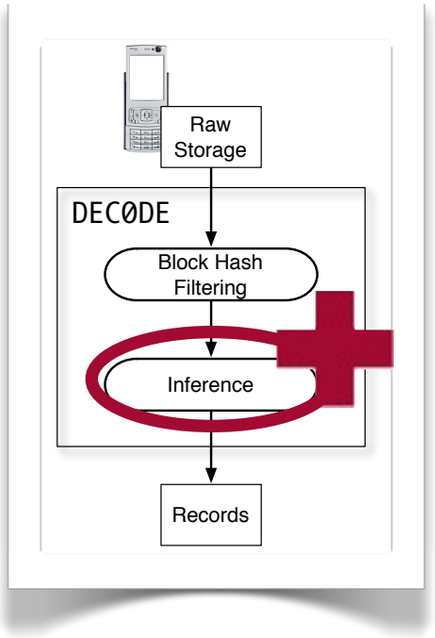
Step 3 > Acquire Raw Storage

Step 4 > Run DEC0DE

Step 2 > Manually create state machines



Make	Model	Count	MB
PFSM Development Set			
Nokia	3200b	4	1.4
Motorola	V551	2	32.0
Samsung	SGH-T309	2	32.0
LG	G4015	2	48.0
Evaluation Set			
Motorola	V400	2	32.0
Motorola	V300	2	32.0
Motorola	V600	2	32.0
Motorola	V555	2	32.0
Nokia	6170	2	4.9
Samsung	SGH-X427M	2	16.0



Component 2:
Inference Evaluation

Process:

Step 0 > Pick phone set

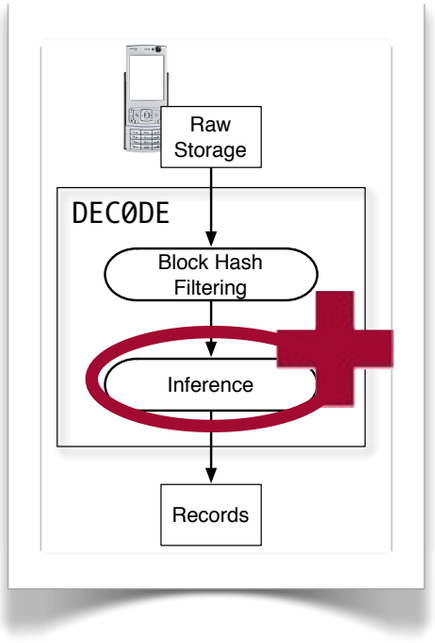
Step 1 > Pick target record types

Step 2 > Manually create state machines

Step 3 > Acquire Raw Storage

Step 4 > Run DEC0DE

.XRY



Component 2:
Inference Evaluation

Process:

Step 0 > Pick phone set

Step 1 > Pick target record types

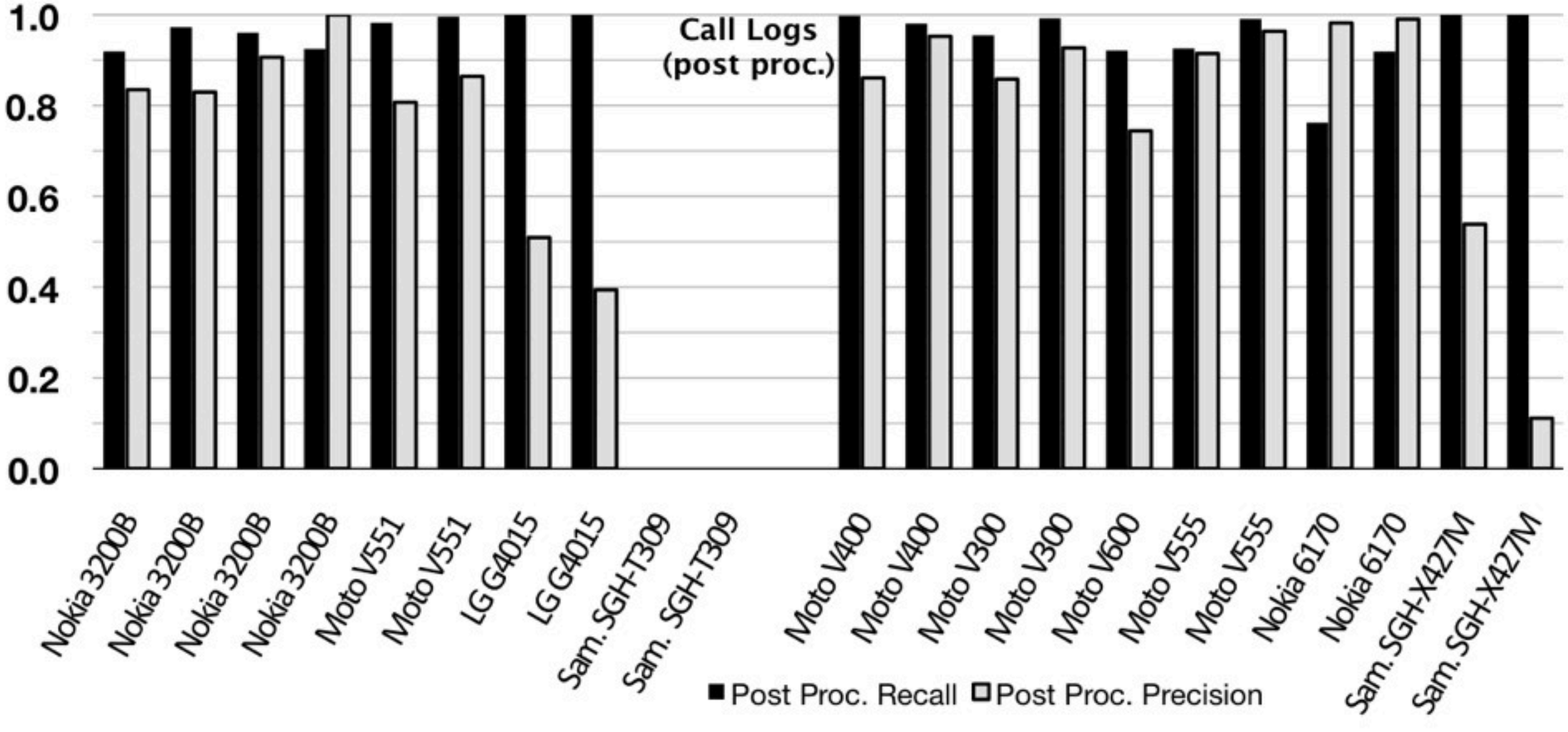
Step 2 > Manually create state machines

Step 3 > Acquire Raw Storage

Step 4 > Run DEC0DE

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

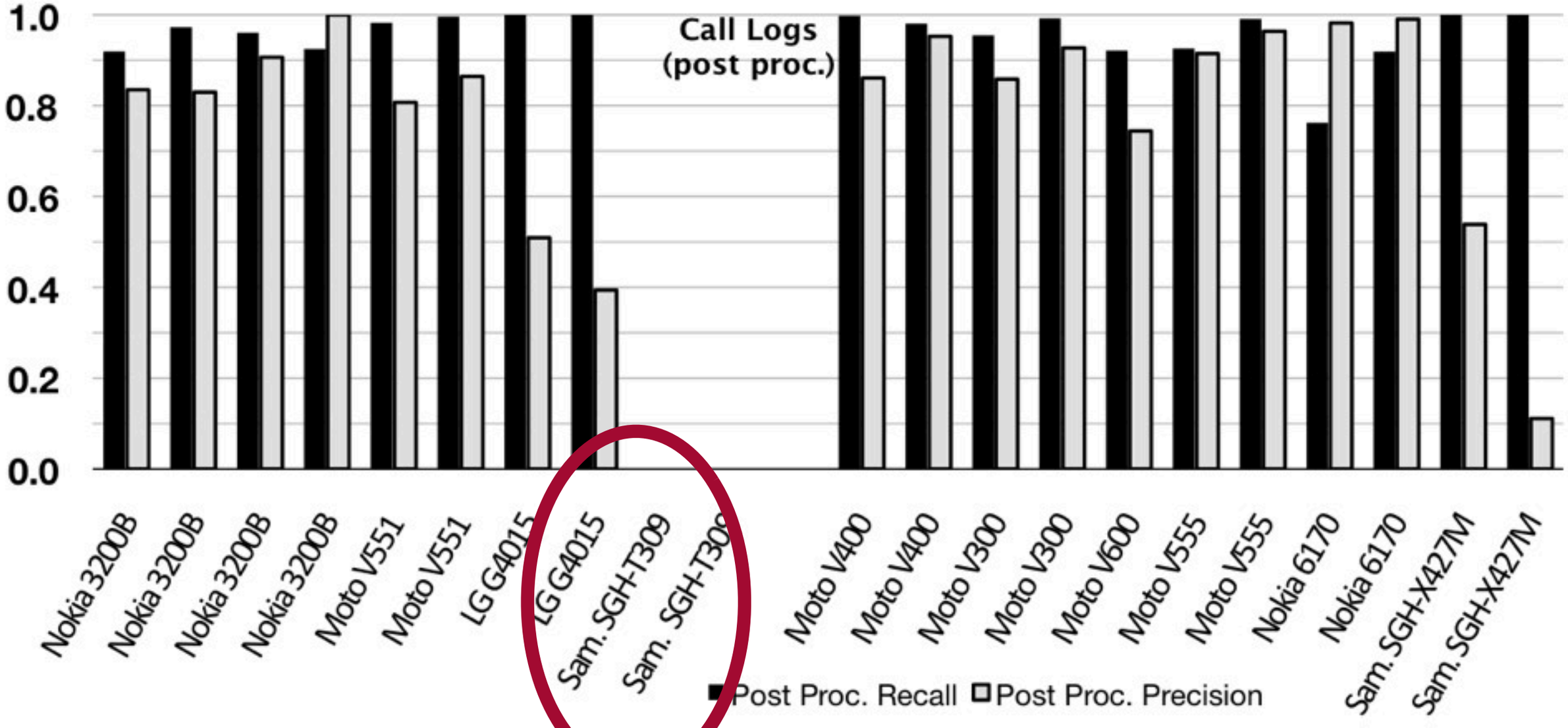


Development Set

Evaluation Set

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

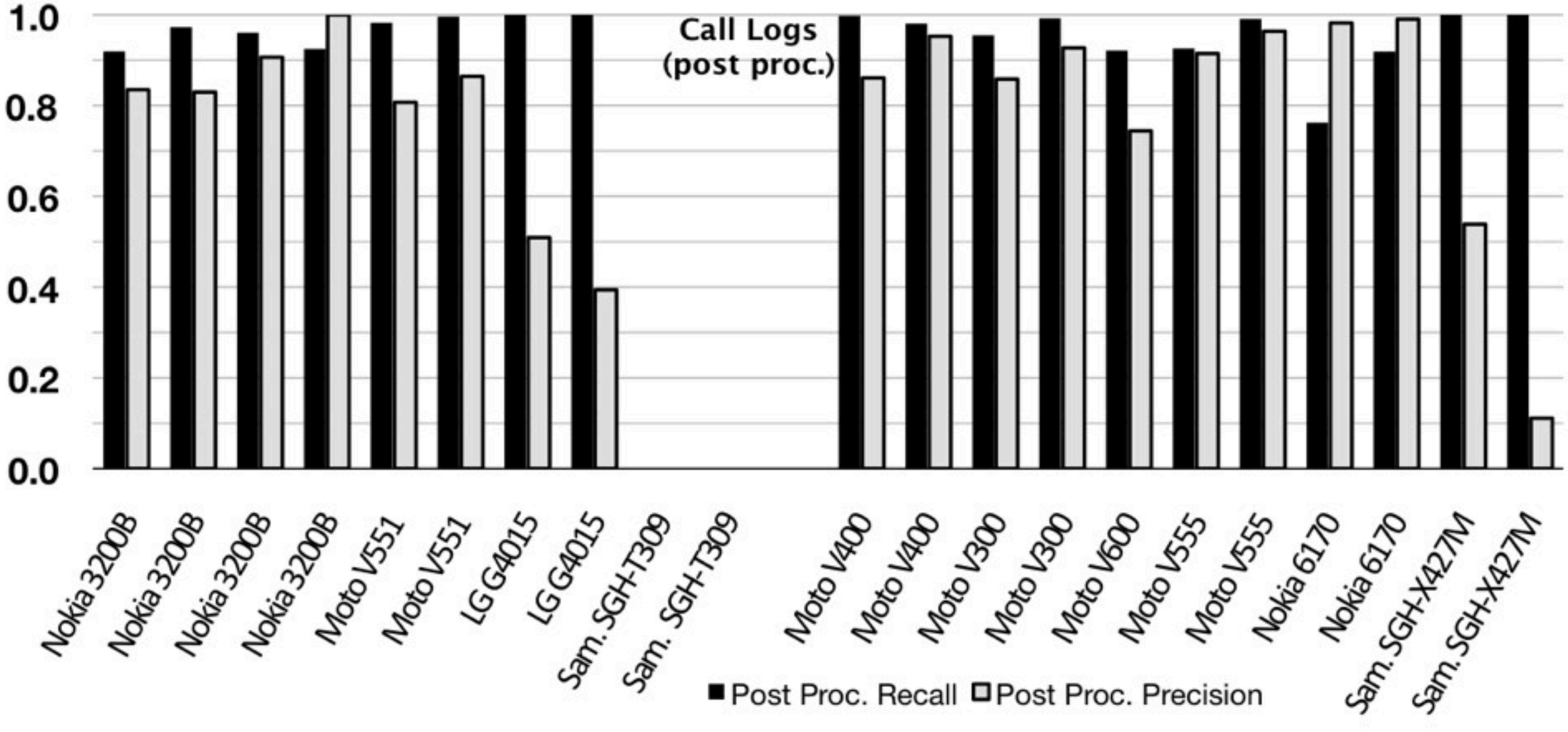


Development Set

Evaluation Set

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

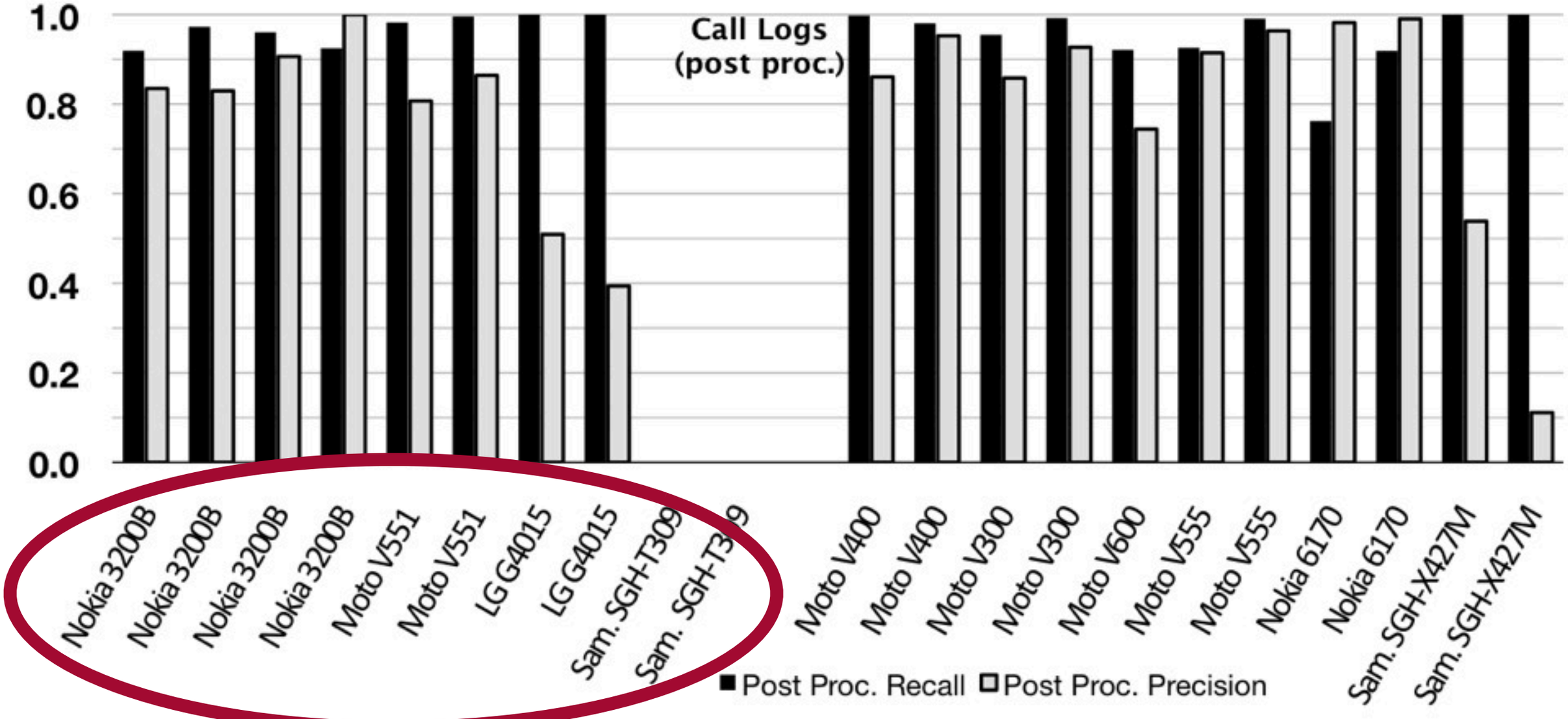


Development Set

Evaluation Set

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

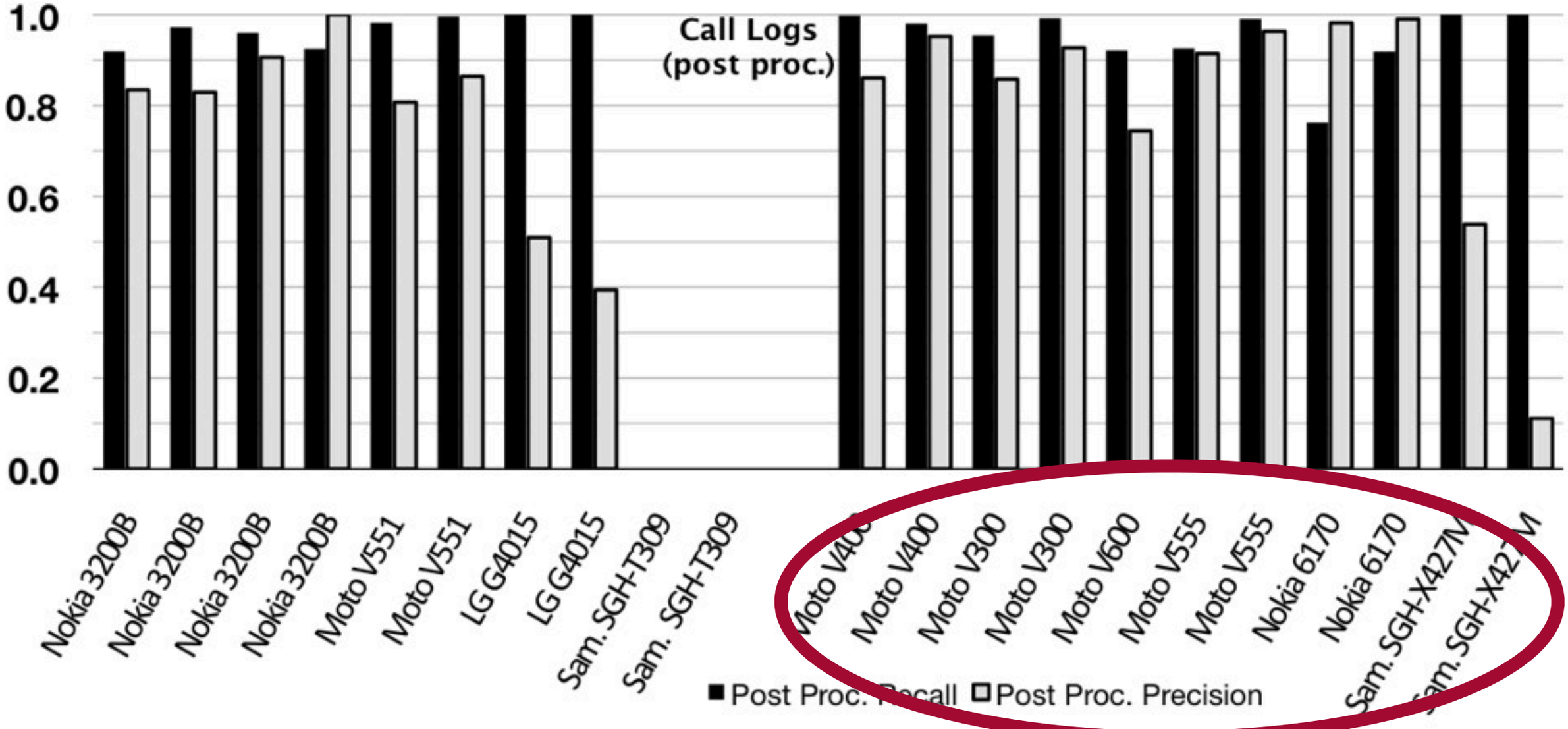


Development Set

Evaluation Set

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

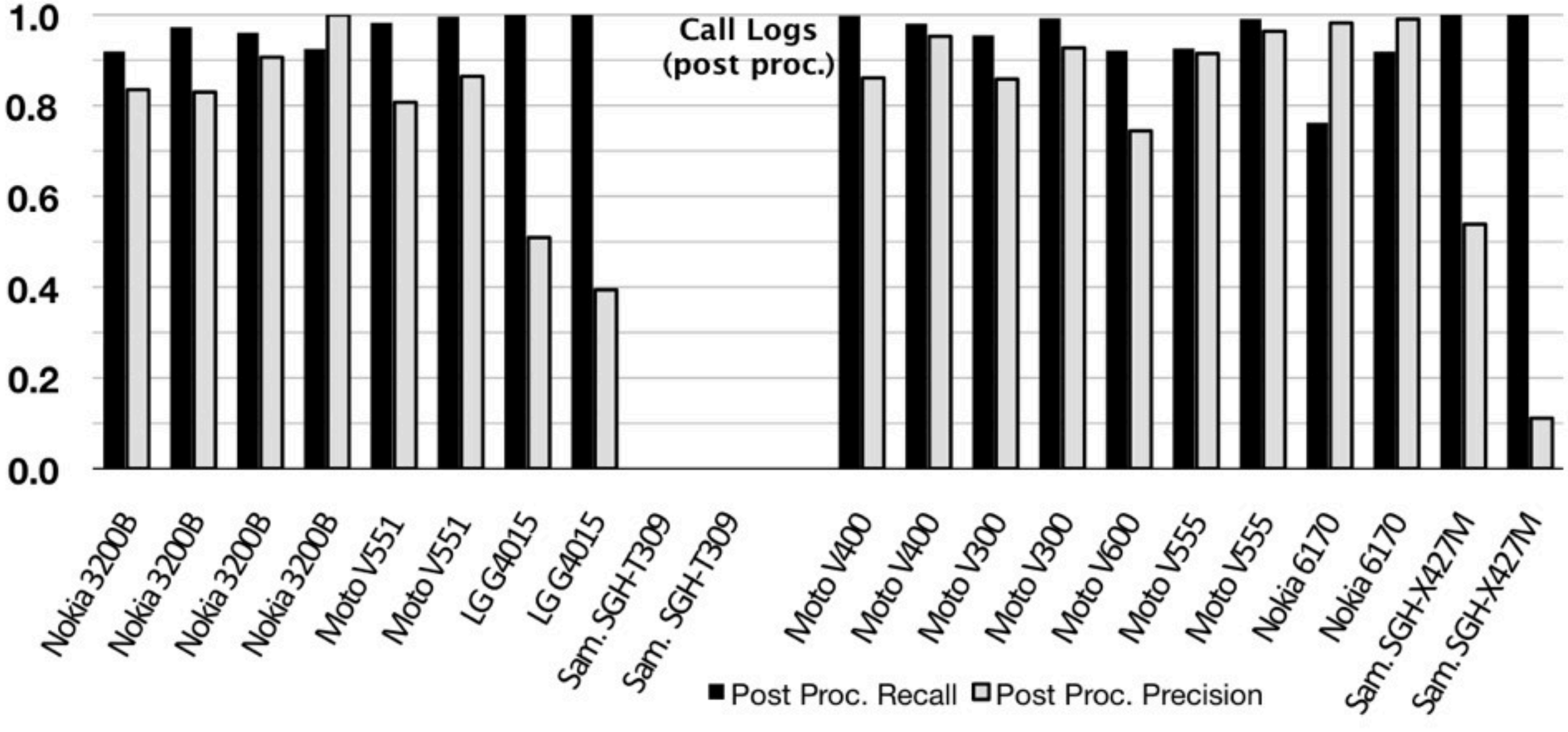


Development Set

Evaluation Set

Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.

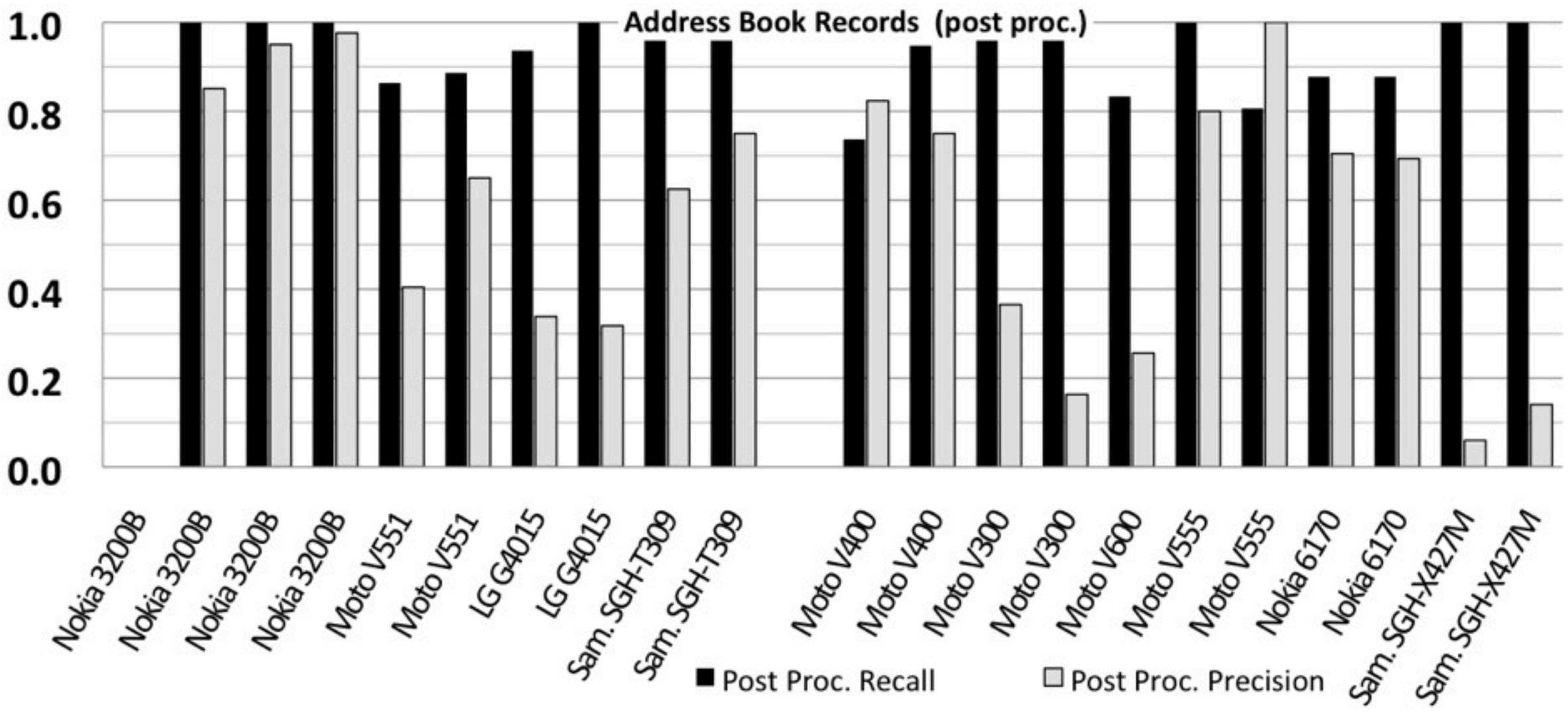


Development Set

Evaluation Set

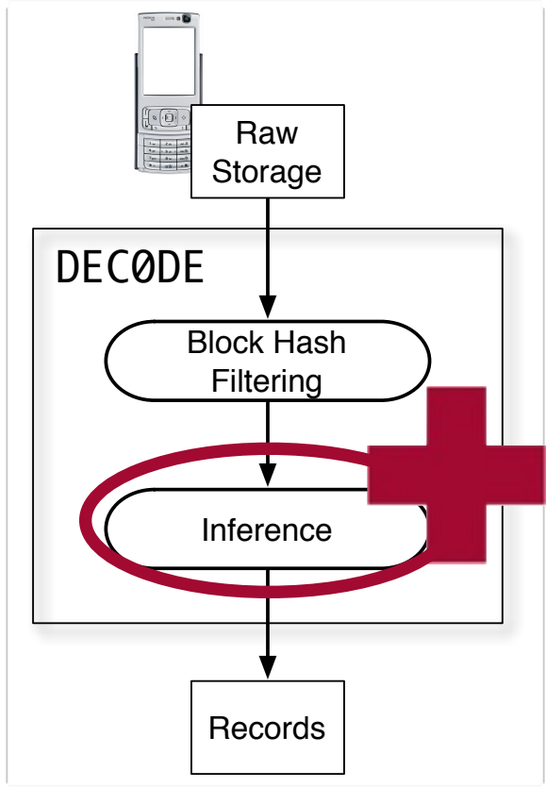
Evaluation: Inference

Recall: Fraction of records recovered. *Precision*: Fraction of results that are actual records.



Development Set

Evaluation Set



Component 2:

Inference

Evaluation Summary:

> Recovered over 93% of records

> Post-processing improves precision by 10-20%

Limitations

- Challenging to acquire raw storage
- Success dependent on PFSM quality
- Small fields are tough

Instrument the binary?

Related Work: Polyglot (Caballero et al. 2007), Tupni (Cui et al. 2008), and Dispatcher (Caballero et al. 2009).

Instrument the binary?

Related Work: Polyglot (Caballero et al. 2007), Tupni (Cui et al. 2008), and Dispatcher (Caballero et al. 2009).

Too slow, too much work, and too difficult to implement for new phone models.

By leveraging our knowledge of a small set of phones,

we can quickly, accurately, and effectively

recover information from previously unexamined phones.