Document Analysis Issues in Reading Optical Scan Ballots

Daniel Lopresti Computer Science & Engineering Lehigh University Bethlehem, PA, USA

George Nagy

Electrical, Computer, and Systems Engineering Rensselaer Polytechnic Institute Troy, NY, USA

Elisa Barney Smith

Electrical and Computer Engineering Boise State University Boise, ID, USA

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 1 DAS Workshop • 6/10/10

Portion of an op-scan ballot

	FEDERAL OFFICES		STATE OFFICES		COUNTY OFFICES	
	PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM	S	TATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE	SOIL	AND WATER CONSERVATION DISTRICT SUPERVISOR DISTRICT 5 VOTE FOR ONE	
	JOHN MCCAIN AND SARAH PALIN	\bigcirc	CAROLYN MCELFATRICK Republican	\bigcirc	DONNA RAE ASP	
	Republican	G	LOREN A. SOLBERG Democratic-Farmer-Labor		write-in, if any	
	BARACK OBAMA AND JOE BIDEN	\bigcirc	write-in, if any		TOWN OFFICES	
	Democratic-Farmer-Labor	(CONSTITUTIONAL	тс	WN SUPERVISOR SEAT B	
$C_{1}^{(n)}$	CYNTHIA MCKINNEY AND ROSA CLEMENTE	AMENDMENT		TOWN OF NORDLAND VOTE FOR ONE		
	Green	Failure to will have the amendment	vote on a constitutional amendment, he same effect as voting no for the	\bigcirc	RUTH ANN NELSON	
8 - <u>1</u> - 1	RÓGER CALERO AND ALYSON KENNEDY Socialist Workers	To vote for a proposed constitutional amendment, completely fill in the oval next to the word "YES" for that question. To vote against a proposed constitutional amendment, completely fill in the oval next to the word "NO" for that		\bigcirc	JOHN STEEBER	
				(* - .)	WAYNE DAVIS	
\bigcirc	RALPH NADER AND MATT GONZALEZ	fill in the ov question.	fill in the oval next to the word "NO" for that question.		write-in, if any	

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 2 DAS Workshop • 6/10/10

In a nutshell

Reading optical scan ballots presents an interesting and important challenge. In addition to our helping solve it, what can be learned?

- Task is extremely well specified (codified into law) and cannot be altered just to make pattern recognition problems easier.
- Like most pattern recognition problems, 100% accuracy may not be attainable, but in this case it is at least a reasonable goal.

Primary contribution of this work:

- Assembling and making openly available a large collection of real hand-marked ballots from a major election. (A "first"?)
- Proposing a paradigm for ground-truthing the collection.

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 3 DAS Workshop • 6/10/10

Ballot reading vs. forms processing

Similarities to forms processing, but also some key differences:

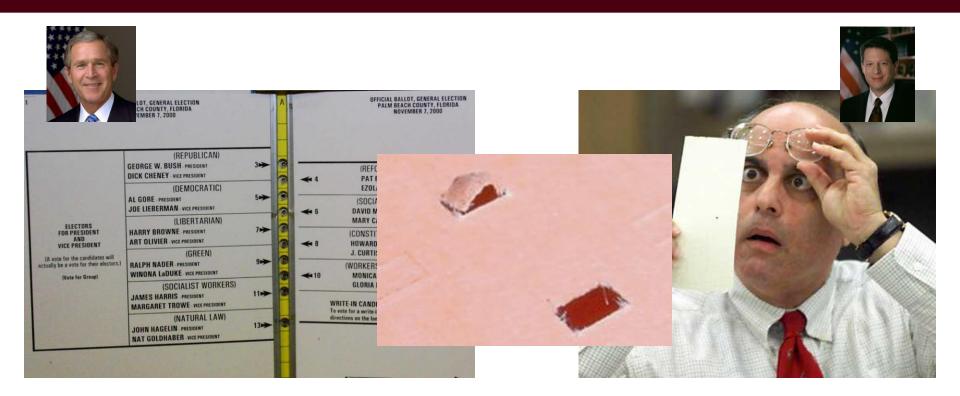
- Much broader range of users (education level, literacy, etc.) than for traditional forms applications.
- Ballots must preserve a voter's anonymity.
- Requirements legally mandated cannot be circumvented.
- Demand to count votes and report results quickly.
- Elections are held infrequently, so voting equipment sits unused for long periods in storage.
- Poll workers often lack technical expertise.
- No *financial* interest in making sure votes are counted accurately, but there is tremendous *public* interest.







Misstep #1: Florida circa 2000



Bad ballot design combined with votomatic technology prone to paper jams. This led to hanging and dimpled chads, making it hard to determine *voter intent*, which provides the legal standard.

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 5 DAS Workshop • 6/10/10

Misstep #2: Rush to DREs ...



Document Analysis Issues in Reading Optical Scan Ballots

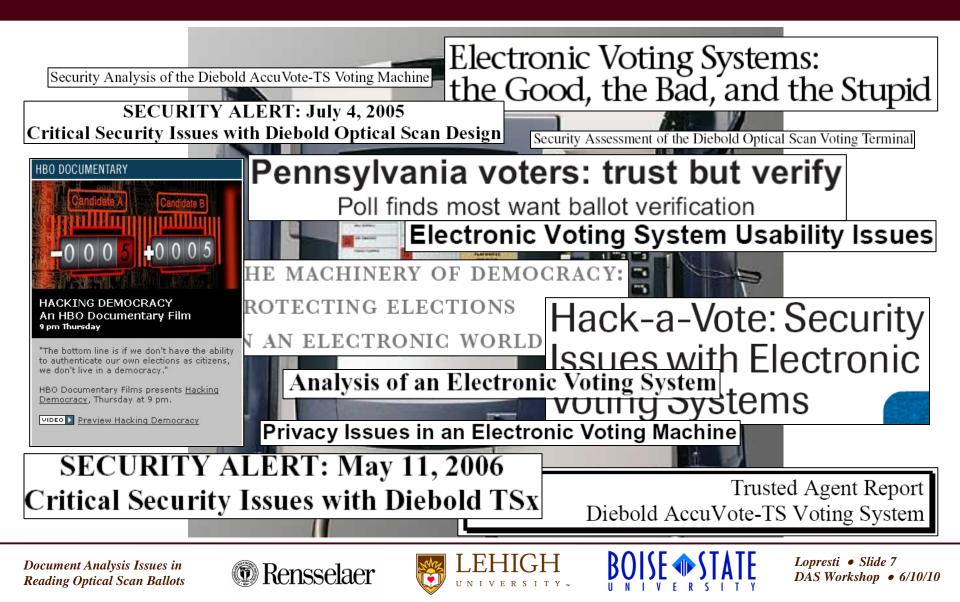




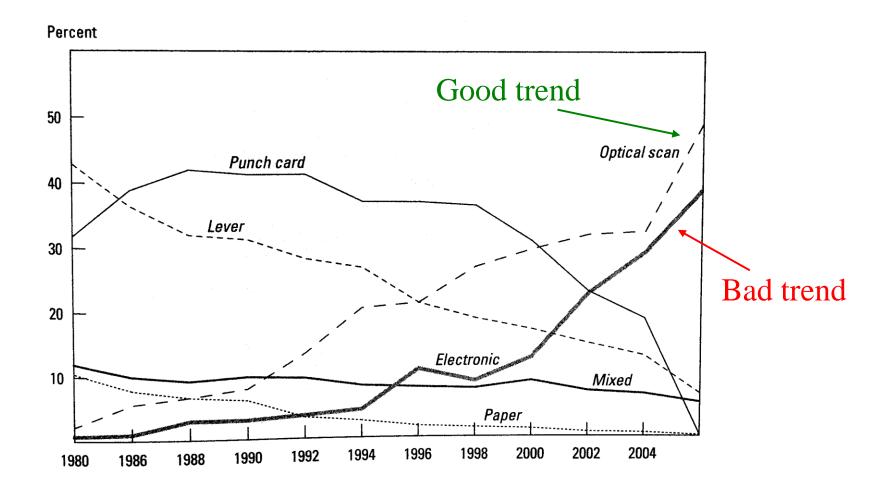


Lopresti • Slide 6 DAS Workshop • 6/10/10

... Misstep #2: Rush to DREs



Voting system use in the U.S.



From Voting Technology: The Not-So-Simple Act of Casting a Ballot, by Paul S. Herrnson, et al, Brookings Institution Press, 2008.

Document Analysis Issues in Reading Optical Scan Ballots

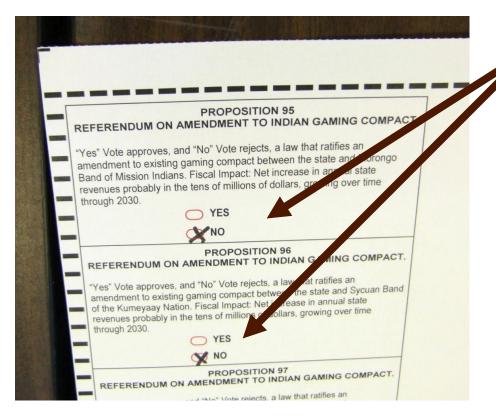




Lopresti • Slide 8 DAS Workshop • 6/10/10

Counting op-scan votes is not so easy

Real ballot from an election in CA:



One of these votes was counted correctly by the op-scan equipment, the other wasn't.

Note: this does <u>not</u> mean voting on paper ballots is bad, just (1) manual audits should be mandatory, and (2) more research is needed.

"Improving California's 1% Manual Tally Procedure," Joseph Lorenzo Hall, UC Berkeley School of Information, EVT Workshop 2008.

Document Analysis Issues in Reading Optical Scan Ballots







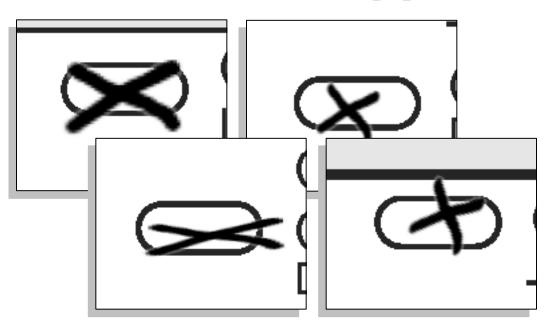
Lopresti • Slide 9 DAS Workshop • 6/10/10

It's all about the data

OFFICIAL BALLOT STAT	TE GENERAL ELECTION BA COUNTY NAME NOVEMBER 7, 2006	ALLOT
To vote, c	INSTRUCTIONS TO VOTERS ompletely fill in the oval(s) next to your choi	ce(s) like this:
FEDERAL OFFICES	STATE OFFICES	COUNTY OFFICES
UNITED STATES SENATOR VOTE FOR ONE	SECRETARY OF STATE VOTE FOR ONE	COUNTY AUDITOR VOTE FOR ONE
CANDIDATE INDEPENDENCE	CANDIDATE INDEPENDENCE	CANDIDATE
CANDIDATE DEMOCRATIC FARMER-LABOR		wite-it, f any
		COUNTY TREASURER VOTE FOR ONE
CANDIDATE Party or Principle	- ward tak	CANDIDATE
UNITED STATES REPRESENTATIVE	VOTE FOR ONE	
DISTRICT [NUMBER]	CANDIDATE INDEPENDENCE	unterio, Fary
	CANDIDATE REPUBLICAN	COUNTY RECORDER VOTE FOR ONE
	CANDIDATE DEMOCRATIC-FARMER-LABOR	CANDIDATE
CANDIDATE DEMOCRATIC FARMER-LABOR		
		urbein, Fary
atteis, if any	VOTE FOR ONE	COUNTY SHERIFF VOTE FOR ONE
STATE OFFICES		CANDIDATE
STATE SENATOR	CANDIDATE	
DISTRICT [NUMBER]	INSTRUCTIONS TO VOTERS COMPLETE IN THE oval(s) next to your choice(s) like the STATE OFFICES STATE OFFICES COUNTY OFFI SECRETARY OF STATE OVER TO ROLE COUNTY ADD OVER TO ROLE COUNTY THEAD COUNTY THEA	O yelevis, fary
		COUNTY ATTORNEY
CANDIDATE REPUBLICAN	CONSTITUTIONAL	
CANDIDATE DEMOCRATIC FARMER-LABOR	AMENDMENT	CANDIDATE
witeit. I avy	Failure to vote on a constitutional amendment, will have	S attain. Lary
STATE REPRESENTATIVE DISTRICT [NUMBER] VOTE FOR ONE	To units for a responsed constitutional amendment	COUNTY SURVEYOR VOTE FOR ONE
CANDIDATE INDEPENDENCE	question. To vote against a proposed constitutional amendment, completely fill in the oval next to the word "NO" for that question.	0
CANDIDATE REPUBLICAN		
CANDIDATE DEMOCRATIC-FARMER-LABOR		
JEMOCRATIC-PARMER-DABOR	-	[CITY NAME OPTIONAL]
GOVERNOR AND		COUNCIL MEMBER
LIEUTENANT GOVERNOR VOTE FOR ONE TEAM		VOTE FOR UP TO TWO
CANDIDATE AND CANDIDATE	COUNTFOFFICES	
CANDIDATE AND	COUNTY COMMISSIONER DISTRICT INUMBER1	
CANDIDATE REPUBLICAN	VOTE FOR ONE	A start, fay
CANDIDATE AND CANDIDATE DEMOCRATIC-FARMER-LABOR		srite-it, if any
steel, EAV	VOTE FRONT AND BACK OF BALLOT	
	Shelor	

To do research, we need data. We could:

- Hire students to fill out sample ballots.
- Generate synthetic data (see examples below from our DAS 2008 paper).



Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 10 DAS Workshop • 6/10/10

Fast forward to November 2008

While useful, neither of these alternatives is completely satisfying. What we would like to have is actual ballots from a real election. Even better, the ballots would be from an important election where the voter markings present serious pattern recognition challenges.



Extremely close U.S. Senate race in State of Minnesota: six days after election, unofficial results showed Republican Norm Coleman leading Democratic challenger Al Franken by 206 votes out of nearly 3 million cast, a difference of less than 0.01%.



Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 11 DAS Workshop • 6/10/10

Fortunate series of events

- Minnesota uses op-scan ballots.
- Closeness of election triggers a manual recount.
- Both sides are allowed to challenge validity of "questionable" ballots.
- Openness laws make challenged ballots a matter of public record.
- Ballot images made available on MN public radio website.
- PDF files contain 300 dpi TIF images!



http://minnesota.publicradio.org/features/2008/11/19_challenged_ballots/

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 12 DAS Workshop • 6/10/10

Minnesota Statutes

Remember that the guiding principle is *voter intent*. Here are a few key points to keep in mind when interpreting ballot markings:

- "A ballot shall not be rejected for a technical error that does not make it impossible to determine the voter's intent."
- "If a mark (X) is made out of its proper place, but so near a name or space as to indicate clearly the voter's intent, the vote shall be counted."
- "Misspelling or abbreviations of the names of write-in candidates shall be disregarded if the individual for whom the vote was intended can be clearly ascertained from the ballot."

https://www.revisor.mn.gov/statutes/?id=204C.22







Lopresti • Slide 13 DAS Workshop • 6/10/10

Minnesota Statutes

... and ...

- "If a voter uniformly uses a mark other than (X) which clearly indicates an intent to mark a name or to mark yes or no on a question, and the voter does not use (X) anywhere else on the ballot, a vote shall be counted for each candidate or response to a question marked.
 - If a voter uses two or more distinct marks, such as (X) and some other mark, a vote shall be counted for each candidate or response to a question marked, unless the ballot is marked by distinguishing characteristics that make the entire ballot defective ..."

https://www.revisor.mn.gov/statutes/?id=204C.22

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 14 DAS Workshop • 6/10/10

Minnesota Statutes

... and ...

- "If the names of two candidates have been marked, and an attempt has been made to erase or obliterate one of the marks, a vote shall be counted for the remaining marked candidate."
- "A ballot shall not be rejected merely because it is slightly soiled or defaced."
- "If a ballot is marked by distinguishing characteristics in a manner making it evident that the voter intended to identify the ballot, the entire ballot is defective."

Goal here is to prevent coercion or vote selling.

https://www.revisor.mn.gov/statutes/?id=204C.22

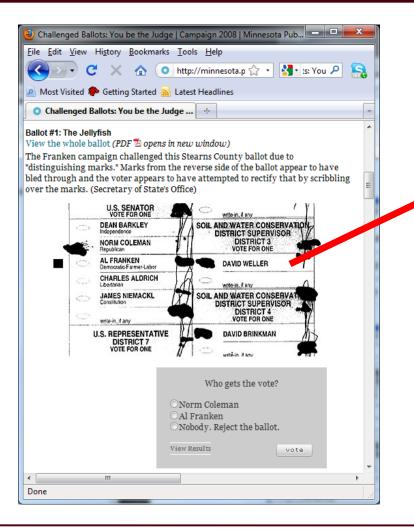
Document Analysis Issues in Reading Optical Scan Ballots

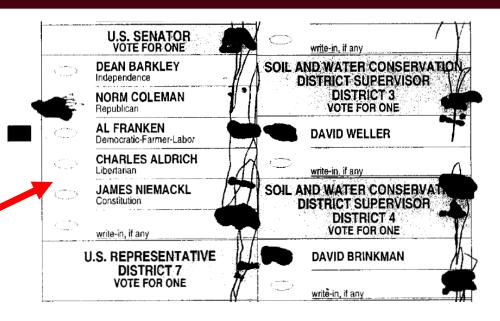






Lopresti • Slide 15 DAS Workshop • 6/10/10





Who gets vote? Public opinion:

- Norm Coleman: 63% (7,626 votes)
- Al Franken: 4% (474 votes)
- Nobody: 33% (4,050 votes)

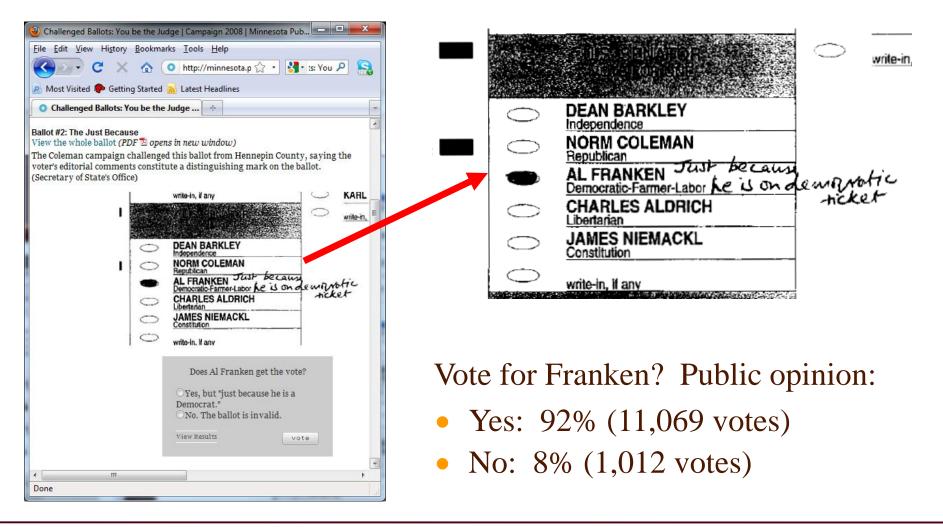
Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 16 DAS Workshop • 6/10/10



Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 17 DAS Workshop • 6/10/10

🕑 Challenged Ballots: You be the Ju	udge Ca	impaign 2008 Minnesota Put	olic Radio NewsQ	2 - Mozilla F 💻 💷 🗾 📕			U.S. SENATOR VOTE FOR ONE	write-i
	• htt	p://minnesota.publicradio.org	/feature 🏠 🔹	🛐 🖬 Is: You be thi 🔎 ි		\bigcirc	DEAN BARKLEY	SOIL AND W
Most Visited P Getting Started Challenged Ballots: You be the Ballot #5: The Yes						\bigcirc	NORM COLEMAN Republican	v
View the whole ballot (PDF 2 ope The Coleman campaign challenge to have written the word 'yes' insi	ed this B	enton County ballot becaus val next to Al Franken's nar			_	Ð	AL FRANKEN Democratic-Farmer-Labor	🗇 WAC
	0 0	U.S. SENATOR VOTE FOR ONE DEAN BARKLEY Independence NORM COLEMAN	SOIL AND DIS1			\bigcirc	CHARLES ALDRICH	write-i
	0	Republican AL FRANKEN Democratic-Farmer-Labor CHARLES ALDRICH Libertaria				\bigcirc	JAMES NIEMACKL Constitution	SOIL AND W
	0 0	JAMES NIEMACKL Constitution write-in, if any	SOIL AND DIST	Ξ		0	write-in, if any	v v
	⊙Ye		?			U	S. REPRESENTATIVE	Der Ber
	ON0	D. Results vo	te	-	Vo	ote fo	or Franken? Public	c opinion:
< Done	III				•	Yes:	: 96% (11,250 vot	es)

• No: 4% (452 votes)

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 18 DAS Workshop • 6/10/10

	,	
📎 Challenged Ballots: You be the Judge Campaign 2008 Minnesota Public Radio NewsQ - Mozilla F., 💶 💷 🔀	U.S. SENATOR VOTE FOR ONE	, <u>write-in, if a</u> ny
File Edit View History Bookmarks Iools Help , C × ☆ ○ http://minnesota.publicradio.org/feature ☆ ・ 🛐・Is: You be th P		
	DEAN BARKLEY	
A Most Visited P Getting Started 🔊 Latest Headlines	NORM COLEMAN	NORM CULEMAN
Challenged Ballots: You be the Judge		
Ballot #7: The Write-out	Democratic Farmer-Labor	
View the whole ballot (PDF 🔁 opens in new window) The Franken campaign challenged this Hennepin County ballot, saying the ballot for U.S. Senate is an	CHARLES ALDRICH	
undervote and not a vote for Norm Coleman. (Secretary of State's Office)	JAMES NIEMACKL Constitution	
U.S. SENATOR write in, if any	write-in, if any	
DEAN BARKLEY	U.S. REPRESENTATIVE	
NORM COLEMAN NORM CULENT	DISTRICT 5 VOTE FOR ONE	
AL FRANKEN Demogratic Farmer Labor		
Charles Aldrich	Independence	
JAMES NIEMACKL Constitution	BARB DAVIS WHITE Republican	
write-in, if any	KEITH ELLISON Democratic-Farmer-Labor	
	write-in, if any	
Does Norm Coleman get the vote?		
○Yes. The voter's intent is clear. ○No. The ballot is an undervote.		
View Results vote	Voto for Colomon'	Dublic opinion
View Results vote	Vote for Coleman'	. Fublic opinion.
Cone Cone Cone Cone Cone Cone Cone Cone	• V_{22} = $540/(600)$	(0, x)
	• Yes: 54% (6,08	su votes)

• No: 46% (5,203 votes)

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 19 DAS Workshop • 6/10/10

MN Challenged Ballot Collection

How the ballot collection was generated and harvested:

- Ballots photocopied and originals stored in a secure location.
- Copies scanned to PDF using auto-feeder flatbed scanner.
- Ballot was two-sided, with both sides scanned simultaneously.
- We wrote a simple web "crawler" that automatically downloaded all the files and extracted TIF images from PDF.
- A total of 6,737 ballots in the set.
- Examination of the TIF suggests that ballots were scanned at 300 dpi bitonal, and that lossy compression was never used.
- Hence, they form an ideal dataset for research purposes.

Document Analysis Issues in Reading Optical Scan Ballots

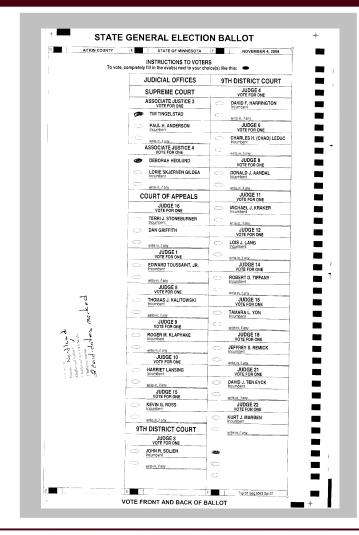






Minnesota ballot front and back

-A			C NOVEMBER 4, 2008	
•		GENERAL ELECTION INSTRUCTIONS TO VOTERS apletely fill in the oval(s) next to your choice(s)		
FEDERAL OFFICES		STATE OFFICES	COUNTY OFFICES	
	PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM	STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE	SOIL AND WATER CONSERVATIO DISTRICT SUPERVISOR DISTRICT 5 VOTE FOR ONE	
n 🔳	JOHN MCCAIN AND SARAH PALIN Republican	CAROLYN MCELFATRICK	DONNA RAE ASP	
	riepub idan	LOREN A. SOLBERG Democratic Farmer-Labor	C write-in, if any	
•	BARACK OBAMA AND JOE BIDEN	C wile-in . I any	TOWN OFFICES	
	CYNTHIA MCKINNEY AND	CONSTITUTIONAL AMENDMENT	TOWN SUPERVISOR SEAT B TOWN OF NORDLAND VOTE FOR ONE	
1	ROSA CLEMENTE Grean	Failure to vote on a constitutional amendment, will have the same effect as voting no for the amendment.	RUTH ANN NELSON	
	RÓGER CALERO AND ALYSON KENNEDY			
	Socialist Workers	To vote for a proposed constitutional amendment, completely fill in the oval next to the word "VES" for that question. To vote against a proposed constitutional amendment, completely fill in the oval inext to the word "NO" for that question.	WAYNE DAVIS	
`	RALPH NADER AND MATT GONZALEZ		C within Kany	
	BOB BARR AND WAYNE A. ROOT	CLEAN WATER, WILDLIFE, CULTURAL HERITAGE AND NATURAL AREAS	TOWN SUPERVISOR SEAT C TOWN OF NORDLAND VOTE FOR ONE	
-	Ubertarian	Shall the Minnesota Constitution be amended to dedicate funding to protect our drinking water sources; to protect, enhance, and restore our	STANLEY W NIX	
-	CHUCK BALDWIN AND DARRELL CASTLE Constitution	webands, prairies, forests, and fish, game, and wildlife habitat; to preserve our arts and cultural heritage; to support our parks and trails; and to protect, enhance, and restore our lakes, rivers.		
	·······	Shall the Whrnesots Constitution be amended to declare function to protect out tricking water wearancy, primes horess, and that, pame, and writing the tablest, to preserve our arts and outward heritage to export our pairs and that is, and to heritage to export our pairs and that is, and to heritage to export our pairs and that is, and to the state state arts for the beginning (b) = 2004. The tables ard use for the beginning (b) = 2004. The tables are used as the beginning (b) = 2004. The tables are used as the beginning (b) = 2004. The tables are used as the beginning (b) = 2004. The pairs (b) = 2004. The table beginning (b) = 2004. The t	TOWN TREASURER TOWN OF NORDLAND VOTE FOR ONE	
1	wile-in, if any	I YES	JUDY PEYSAR	
	U.S. SENATOR VOTE FOR ONE	○ N0	with in, if any	
	DEAN BARKLEY Independence	COUNTY OFFICES	SCHOOL DISTRICT OFFICES	
Xe	NORM COLEMAN Republican	COUNTY COMMISSIONER	SCHOOL BOARD MEMBER	
	AL FRANKEN Democratic-Farmer-Labor	DISTRICT 2 VOTE FOR ONE	INDEPENDENT SCHOOL DISTRICT NO. 1 (AITKIN) VOTE FOR UP TO THREE	
- 1950) 1	CHARLES ALDRICH Libertarian	DALE K. LUECK		
1.000	JAMES NIEMACKL Constitution	LAURIE A. WESTERLUND	PETER A. WELSHONS	
	wile in, if any	🗢 wile-n, itany	DENNIS HASSKAMP	
	J.S. REPRESENTATIVE DISTRICT 8 VOTE FOR ONE	SOIL AND WATER CONSERVATION DISTRICT SUPERVISOR DISTRICT 3	JOHN CHUTE	
	MICHAEL CUMMINS Republican	VOTE FOR ONE	CHARLES WIKELIUS	
	AMES L. OBERSTAR Democratic-Farmer-Labor	G FRANKLIN TURNOCK	WILLIAM E. STIMAC	
100		SOIL AND WATER CONSERVATION	write-it, it any	
0145 TOW	write-in, if any	DISTRICT SUPERVISOR DISTRICT 4 VOTE FOR ONE	write-in_if any	
100		BOB JANZEN	witten, it any	
		witte-in, if any		



Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 21 DAS Workshop • 6/10/10

More about the collection

Reason for challenge is recorded in stamp added to each ballot:

1 S Senator Precinct: Shamrock Challenged ballot #: 1 Challenged by COLEMAN Challenge reason: Red stray mark

- Challenges as a result of manual recount not machine rejects.
- Ballot is challenged for one race others may be "normal."
- Challenges are extremely aggressive (both sides want to win).

Franken ultimately declared winner by 312 votes: machine-reading got the results of the election wrong.

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 22 DAS Workshop • 6/10/10

Ground-truthing protocol

Guiding principles:

- "Ground-truthing" can refer to identifying raw markings on ballot and/or interpreting their meaning ("vote" / "no vote").
- We have built a GUI tool that captures truth at both levels.
- Ideally, a system built for reading op-scan ballots should replicate same understanding of voter intent possessed by a knowledgeable human judge.
- Legitimate differences of opinion exist, so "truth" is relative.
- Users provided with 15-page guideline, as well as MN statutes.

What here generalizes to other document analysis applications?

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 23 DAS Workshop • 6/10/10

BallotTool GUI

OFFICIAL BALLOT Judge J	
AITKIN COUNTY STATE OF MINNESOTA C NC User D: 900 Last earlby: Amotation: Addit, Nortand challenged, Date STATE GENERAL ELECTION BALLOO INSTRUCTIONS TO VOTERS To vote, completely fill in the oval(s) next to your choice(s) like this: FEDERAL OFFICES STATE OFFICES COUNT PRESIDENT AND VICE PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM VICE PRESIDENT VOTE FOR ONE TEAM VICE PRESIDENT VICE FOR ONE TEAM VICE PRESIDENT VICE FOR ONE TEAM VICE FOR ONE TEAM VIE ON AND SARAH PALIN Republican VIE of dama A SOLBERG VIE of dama A SOLBERG VIE of dama A SOLBERG VIE of dama A AND JOE BIDEN DEMOGRATIC Farmer Labor CONSTITUTIONAL TOWN SUPER VIE	
STATE GENERAL ELECTION BALLOT INSTRUCTIONS TO VOTERS To vote, completely fill in the oval(s) next to your choice(s) like this: RESIDENT AND VICE PRESIDENT Vote, completely fill in the oval(s) next to your choice(s) like this: FEDERAL OFFICES STATE OFFICES PRESIDENT AND VICE PRESIDENT AND VICE PRESIDENT VOTE FOR ONE STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE Soil AND WATE DISTRICT 3B VOTE FOR ONE JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK Republican DONNA Caroeled wete Note: 21 Y 1 Vaid Mark Stray Mark Caroeled wete No Yole No Yole Y 1 DOUL JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK Bemocraitic Farmer-Labor DONNA Republican Republican Vaid Mark Stray Mark Caroeled wete No Yote No Yote Y 1 Y 1 JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK Bemocraitic Farmer-Labor DONNA Republican Republican Vaid Mark Stray Mark Caroeled wete No Yote No Yote Y 1 Y 1 JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK Bemocraitic Farmer-Labor DONNA Republican Republican Vaid Mark Stray Mark Caroeled wete No You C	alo Cickmode: L
INSTRUCTIONS TO VOTERS To vote, completely fill in the oval(s) next to your choice(s) like this: FEDERAL OFFICES FEDERAL OFFICES STATE OFFICES COUNT PRESIDENT AND VICE PRESIDENT VOTE FOR ONE JOHN MCCAIN AND SARAH PALIN Republican Vald (87,53) BARACK OBAMA AND JOE BIDEN Damocratic Farmer-Labor Vald Mark ConstitutionAL Mathematic Stray Mark ConstitutionAL TOWN SUPER Substance COUNTINE Solution CONSTITUTIONAL TOWN SUPER Solution Town Super Solution ConstitutionAL Solution S	_1-000.tf
To vote, completely fill in the oval(s) next to your choice(s) like this: C vald Mark: C Bray Mark: C Canceled Vate: P Na Yaze: X1 Y1 FEDERAL OFFICES COUNT FEDERAL OFFICES STATE OFFICES COUNT PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE TEAM STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE TEAM SoliL AND WATE DISTRICT 3B VOTE FOR ONE TEAM JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK DONNAF DONNAF C Vaid Mark: C Bray Mark: C Canceled Vale: C No Vace: X1 or 1 Y1 cool Vaid (97,533 CAROLYN MCELFATRICK Democratic-Farmer-Labor DONNAF C Vaid Mark: C Bray Mark: C Canceled Vale: C No Vace: X1 or 1 Y1 cool Vaid (97,533 CAROLYN MCELFATRICK Democratic-Farmer-Labor DONNAF C Vaid Mark: C Bray Mark: C Canceled Vale: C	
PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE TEAM SOIL AND WATE DISTRICT DISTRICT SOIL AND WATE DISTRICT Curkle Mark Canceled Vate N Vate XI I III JOHN MCCAIN AND SARAH PALIN Republican CAROLYN MCELFATRICK Republican DONNA f Vaid Mark Stray Mark Canceled Vate M Agentited View M Agentited View Vaid (87,53) CAROLYN MCELFATRICK BARACK OBAMA AND JOE BIDEN Damocratic Farmer-Labor DONNA f Vaid Mark Stray Mark Canceled Vate W Magentited View Write-n d any TOWN Bob Bar Constitution All ToWn Super Quit	X2 Y2
PRESIDENT AND VICE PRESIDENT VOTE FOR ONE TEAM STATE REPRESENTATIVE DISTRICT 3B VOTE FOR ONE SOIL AND WATE DISTRICT DISTRI	
VICE PRESIDENT VOTE FOR ONE TEAM USTRICT 3B USTRICT 3B USTRICT USTRICT USTRICT 3B USTRICT USTRICT USTRICT 3B USTRICT USTRICT 3B USTRICT USTRICT USTRICT 3B USTRICT U	x2 <u>410</u> v2
VOTE FOR ONE TEAM VOTE FOR ONE DIS Vote FOR ONE TEAM VOTE FOR ONE VOTE VOTE VOTE VOTE VOTE VOTE VOTE VOT	
JOHN MCCAIN AND SARAH PALIN Republican valid (87,53) BARACK OBAMA AND JOE BIDEN Democratic-Farmer-Labor CONSTITUTIONAL TOWN SUPER Care Valid Mark C Bray Mark C Care Mark C Ca	x2 Y2
SARAH PALIN Republican Republican valid (97,53) Image: Concell to the second s	
Valid (97,53) BARACK OBAMA AND JOE BIDEN Democratic: Farmer-Labor Cvald Mark C Brey Mark C Cancel Democratic: Farmer-Labor Bob Bar CVAld Mark C Brey Mark C Cancel Bob Bar CONSTITUTIONAL TOWN SUPER Quit	
Image: Second	
Democratic-Farmer-Labor CONSTITUTIONAL TOWN SUPEr	
CYNTHIA MCKINNEY AND AMENDMENT TOWN OF STATE OF THE STATE	
Green Failure to vote on a constitutional amendment, will have the same effect as voting no for the AUTH ANN NELSON	
RÓGER CALERO AND	-
Socials Workers word "VEST for being fill in the oval next to the	
processed constitutional amendment, completely	
debautin, debautin, construction of the second se	
Independent	

Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 24 DAS Workshop • 6/10/10

Ground-truth

Physical level:

- Each target annotated with bounding box as valid mark, stray mark, cancelled vote, or no vote.
- Marking style is recorded as filled oval, partially filled oval, ex (X), or check mark.
- Other marks on ballot face annotated (challenge stamp, bleedthrough, voter or official handwriting, fiducial, skew line, etc.).

Logical level:

Note two are not equivalent.

• User specifies who gets legal vote(s) in each race (i.e., intent).

Document Analysis Issues in Reading Optical Scan Ballots

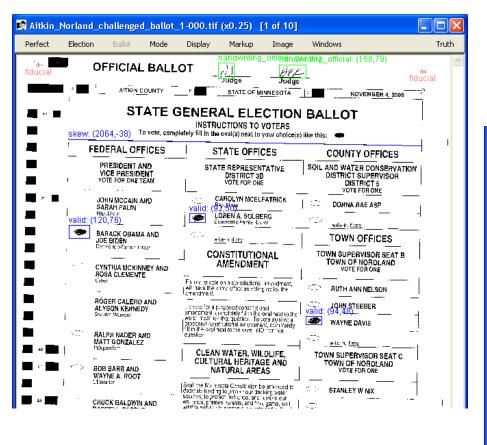


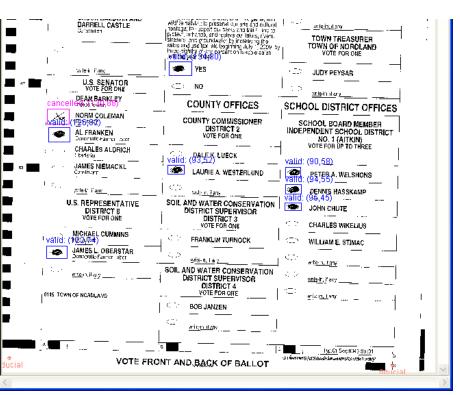




Lopresti • Slide 25 DAS Workshop • 6/10/10

Example of a Fully Annotated Ballot





Document Analysis Issues in Reading Optical Scan Ballots

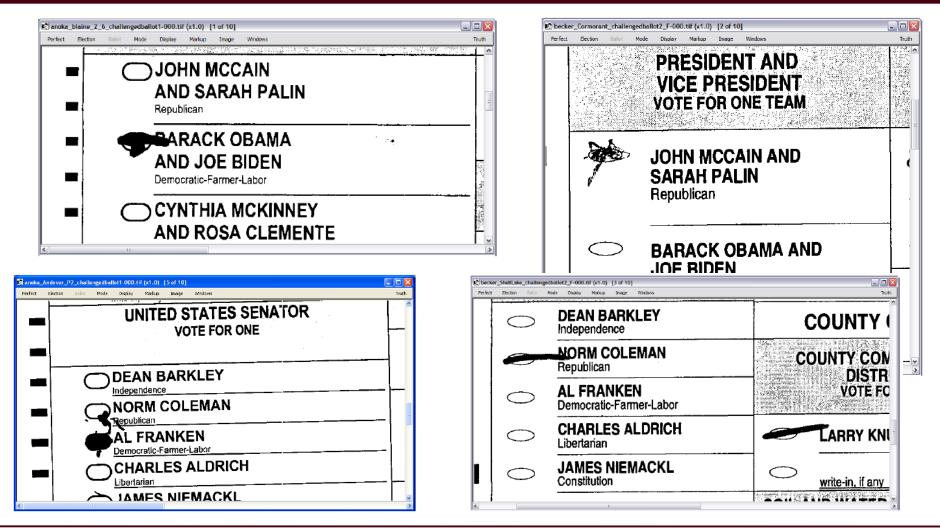






Lopresti • Slide 26 DAS Workshop • 6/10/10

Sloppy-but-valid marks



Document Analysis Issues in Reading Optical Scan Ballots

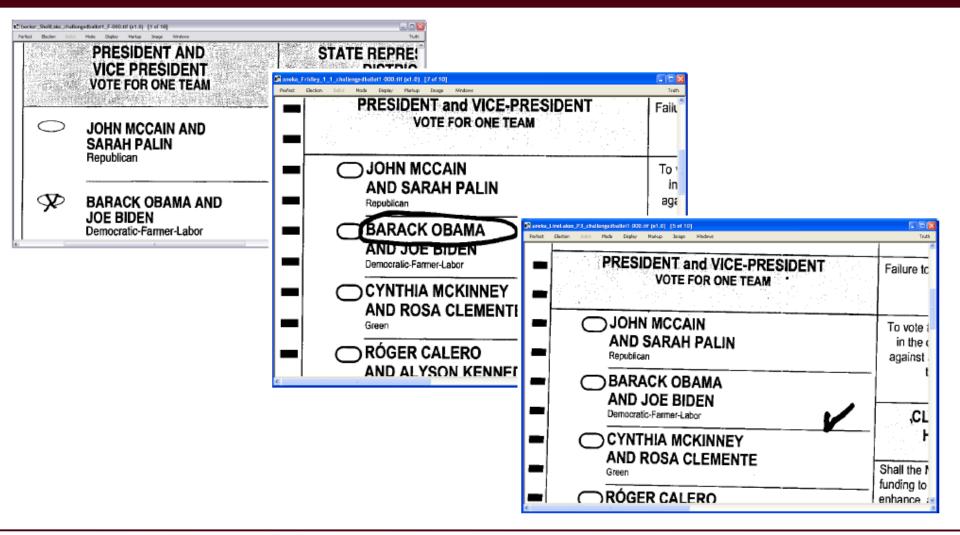






Lopresti • Slide 27 DAS Workshop • 6/10/10

Non-confirming marking styles



Document Analysis Issues in Reading Optical Scan Ballots

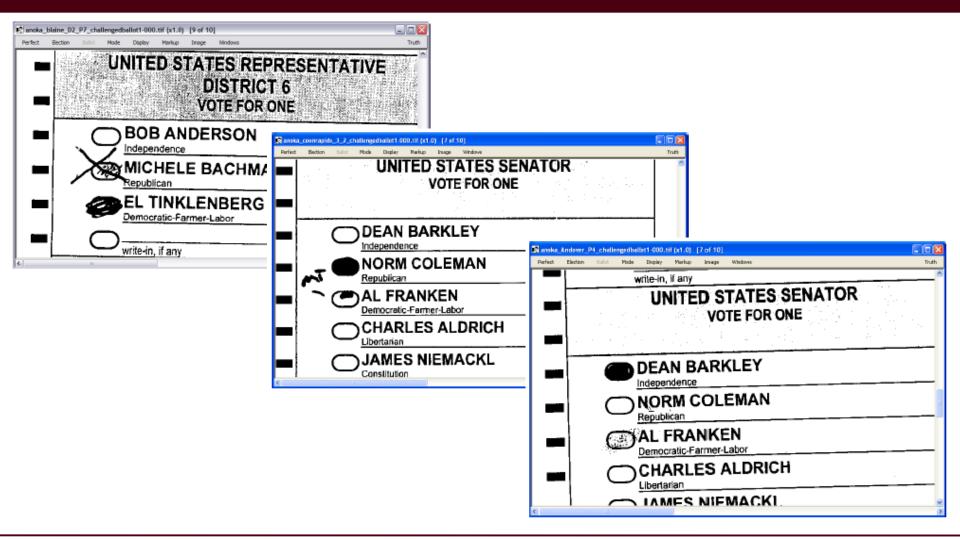






Lopresti • Slide 28 DAS Workshop • 6/10/10

Attempts to cancel a vote



Document Analysis Issues in Reading Optical Scan Ballots

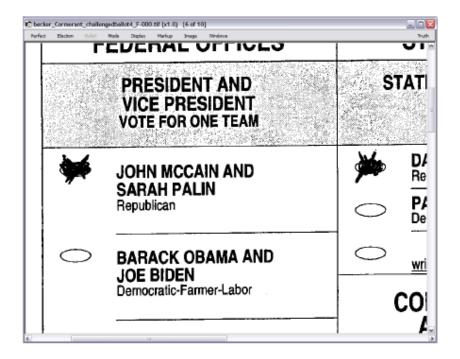


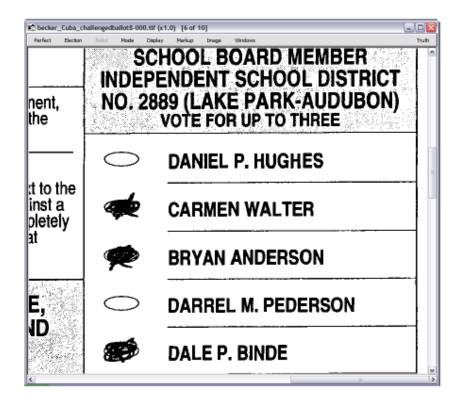




Lopresti • Slide 29 DAS Workshop • 6/10/10

Intended votes that look cancelled





Document Analysis Issues in Reading Optical Scan Ballots

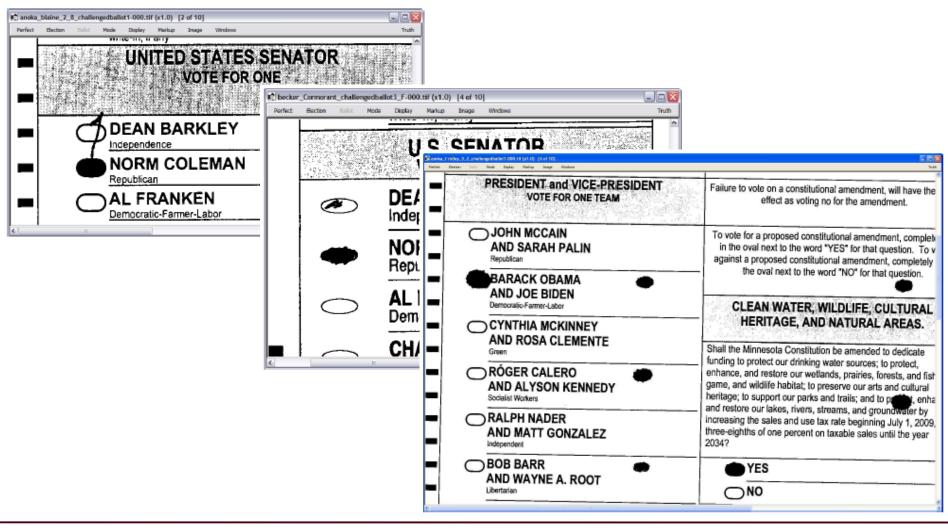






Lopresti • Slide 30 DAS Workshop • 6/10/10

Stray marks and bleedthrough



Document Analysis Issues in Reading Optical Scan Ballots

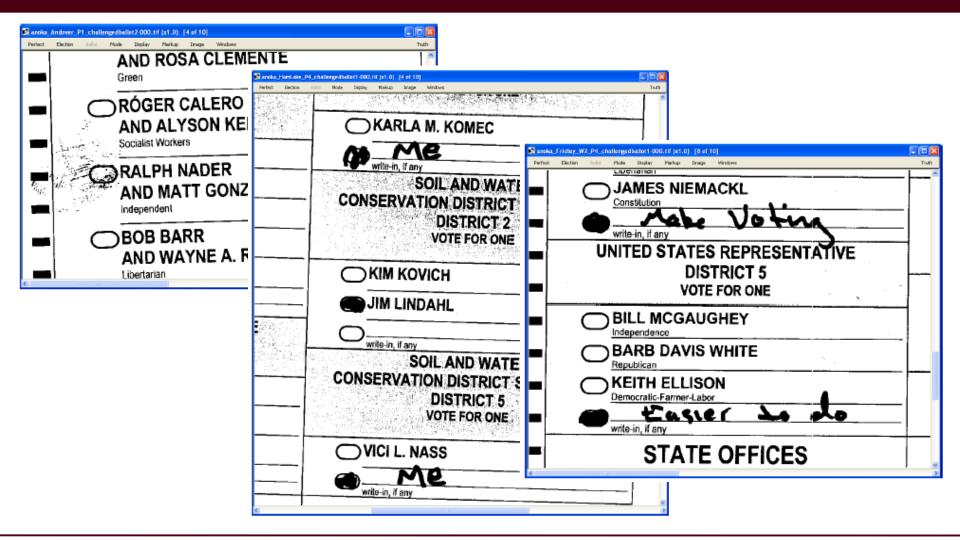






Lopresti • Slide 31 DAS Workshop • 6/10/10

Invalidating markings



Document Analysis Issues in Reading Optical Scan Ballots







Lopresti • Slide 32 DAS Workshop • 6/10/10

Why isn't this an easy problem?

After all, ballots are just a simple type of form. We must read the votes correctly, but we aren't expected to recognize write-in names. Can't we just push up reject rate until the accuracy reaches 100%?

Remember, we can't change rules in ways that violate the law. VOTER INTENT is the definition we must always follow.

To do this right, we must be prepared to:

- Reject any ballot that may contain "identifying marks."
- Recognize intent when mark is atypical and/or far from target.
- Accurately identify when a vote has been cancelled.

Document Analysis Issues in Reading Optical Scan Ballots



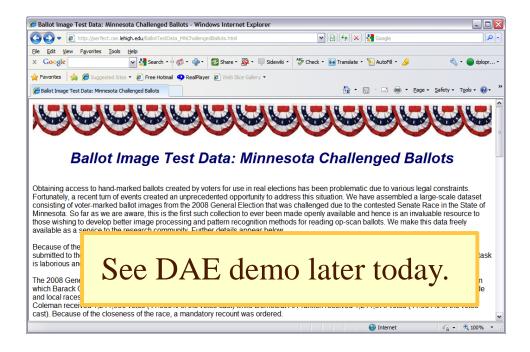




Lopresti • Slide 33 DAS Workshop • 6/10/10

Status

- All 6,737 ballots now online at: http://perfect.cse.lehigh.edu/BallotTestData_MNChallengedBallots.html
- Ground truth collected from 8 test subjects, 980 ballot sides.





Coming soon:

- Add to TC-11 website.
- Ground-truthing continues – to be made available online.







Lopresti • Slide 34 DAS Workshop • 6/10/10