

; 04/10/13 3:38 PM
;
;;;LEARNINGTIMES C2C 4/10/2013

WELCOME TO THE WEBINAR.

>>

I WANT TO TALK ABOUT A FEW THINGS IF I CAN BEFOREHAND. IF FOR WHATEVER REASON YOUR AUDIO IS -- YOU'RE EXPERIENCING AN ECHO THERE'S A LIKELIHOOD YOU ARE LOGGED IN TWICE.

PLEASE MAKE SURE THAT YOU ONLY HAVE ONE SESSION OPEN AND IF YOU DO HAVE AN EXTRA ONE OPEN YOU CAN GO AHEAD AND CLOSE THAT.

IT WILL CLEAR UP THE AUDIO. IF FOR SOME REASON YOU ARE NOT EXPERIENCING ANY AUDIO AT THE MOMENT, OBVIOUSLY YOU'RE NOT GOING TO HEAR MY MESSAGE BUT YOU CAN PROBABLY DEDUCT FROM THE CLOSED CAPTIONING THAT IS GOING THROUGH THAT YOU'RE NOT HEARING ANYTHING.

TRY TURNING OFF YOUR -- OR CLOSING YOUR BROWSER SESSION AND COMING BACK IN.

THAT TYPICALLY WILL HELP RESOLVE THE ISSUE IN ADDITION IF AT ANY POINT IN TIME SLIDES OR THINGS ENTIRELY STOP CLOSE YOUR BROWSER AND COME BACK IN AND LET US KNOW IN THE CHAT WINDOW IF YOU NEED ANY ADDITIONAL SUPPORT AND WE WILL DO WHAT WE CAN TO ASSIST. THERE WILL BE A MODERATED CHAT IN THIS SESSION WHICH MEANS THAT ONCE WE BEGIN THE MODERATION ANY QUESTION THAT YOU SUBMIT WILL NOT IMMEDIATELY BE VISIBLE BUT ONCE WE DO MODERATE IT THROUGH YOU'LL SEE IT COME THROUGH TWICE ON YOUR SCREEN.

DON'T WORRY, IT ONLY WILL BE VISIBLE TWICE TO YOU.

EVERYONE ELSE WILL ONLY SEE IT ONCE.

THE FIRST TIME IT BECOMES POSTED IT JUST INDICATES TO YOU THAT YOUR MESSAGE HAS GONE THROUGH AND YOU'LL SEE IT AGAIN WHEN WE DO RESPOND.

WE WILL BEGIN IN A FEW MOMENTS HERE.

I'M JUST GOING TO GO AHEAD AND START THE RECORDING FOR TODAY'S SESSION.

I WOULD LIKE TO INVITE OUR HOST KRISTIN KRISTEN FROM HERITAGE PRESERVATION AND OUR PRESENTERS DANIELLE PLUMER AND JEFFERSON BAILLY TO PLEASE GO AHEAD AND BEGIN AS SOON AS YOUR READY.

>> THANK YOU, MIKE, THIS IS KRISTEN AT HERITAGE PRESERVATION IN WASHINGTON, D.C.

IT'S ALMOST 90 DEGREES AND WE'VE BEEN HEARING WEATHER REPORTS ACROSS THE WORLD IN THE HELLO BOX, THANK YOU FOR THAT.

I WANT TO JUST VERY QUICKLY DO AN INTRODUCTION BECAUSE WE HAVE A LOT TO COVER TODAY AND I DON'T WANT TO TAKE AWAY FROM THAT THERE IS CARING FOR DIGITAL MATERIALS, PREVENTING A DIGITAL DARK AGE.

WE WANT TO THANK LEARNING TIMES FOR PRODUCING THIS TO US AND ESPECIALLY FOR THE MUSEUM AND LIBRARY SERVICES FOR MAKING FUNDING POSSIBLE FOR THESE COURSES.

WE HAVE OVER 200 PEOPLE LOGGED IN ALREADY AND ALMOST 400 ON THE LAST FEW WEBINARS SO WELCOME EVERYONE TO THAT WE HAVE JEFFERSON BAILEY ABOUT PRACTICING SAFE ARCHIVING BACKUP COPIES AND WHAT CAN GO WRONG AND THEN WE WILL HAVE OUR LAST CLASS ON MONDAY, APRIL 15 AND WE LOOK

FORWARD TO THAT.

IF YOU HAVEN'T YET DONE SO,
PLEASE DO LOOK AT THE COURSE WEB
PAGE, WE HAVE PUT LINKS TO THE
POWERPOINT SLIDES WHERE YOU CAN
FIND HOME WORK IN YOU MISS AN
ASSIGNMENT AND YOU WANT TO LINK
TO IT YOU'LL FIND THAT THERE.

LOTS OF RESOURCES THAT OUR
SPEAKERS HAVE PULLED TOGETHER
FOR YOU SO PRETTY MUCH ANYTHING
IN THEIR POWER POINT THEY'VE
ALREADY GOT LINK TO ON THE SITE.
AND FURTHER INFORMATION.

WE'VE HAD GREAT CONVERSATION AS
WE'VE BEEN GOING ALONG AND WE'VE
BEEN GOING THROUGH QUESTIONS AND
COMMENTS AND ALL OF THOSE WILL
BE GETTING UP THERE SO KEEP
CHECKING BACK ON THIS PAGE AND
HOPEFULLY YOU'RE GETTING OUR
E-MAILS WITH THE RECORDINGS
AFTER EACH WEBINAR'S BROADCAST.
IF YOU ARE WORKING TOWARDS OUR
CERTIFICATE OR DIGITAL
CREDENTIALS YOU NEED TO REGISTER
TURN IN YOUR PERMISSION FORM,
WATCH EACH WEBINAR IN THE COURSE
EITHER LIVE OR ON THE RECORDING
AND COMPLETE ALL FIVE HOME WORK
ASSIGNMENTS AND MAKE SURE THAT'S
ALL DONE BEFORE MONDAY, APRIL
22.

DANIELLE ON MONDAY WILL GO OVER
TODAY'S ASSIGNMENT AND THE
ASSIGNMENT SHE MADE YESTERDAY TO
GIVE YOU FEEDBACK BUT HOPEFULLY
BY THE END OF TODAY'S SESSION WE
CAN TALK ABOUT WHAT THE HOME
WORK ASSIGNMENT IS FOR THE
SESSION.

WE WELCOME YOU TO JOIN THE
COMMUNITY.

THE CONNECTING TO COLLECTIONS
ONLINE COMMUNITY IF YOU BECOME A
MEMBER YOU CAN JOIN THE
DISCUSSION BOARDS AND ASK

QUESTIONS OF YOUR PEERS.
WE'RE ABOUT SIX PEOPLE OFF OF
HITTING OUR 3,000 MEMBER MARK SO
IF YOU CAN BE THAT THREE
THOUSANDTH MEMBER WE HAVE A
SPECIAL PRIZE THAT WE'RE LINING
UP FOR THAT.

SO IF YOU ARE A MEMBER WE INVITE
YOU ASK YOUR COLLEAGUES TO JOIN
YOU ON THE ONLINE COMMUNITY.
AND AS ALWAYS IF YOU HAVE ANY
QUESTIONS CONNECT HERITAGE
PRESERVATION AND WE'LL DO OUR
BEST TO HELP.

I WANT TO INTRODUCE TODAY'S
SPEAKER, JEFFERSON BAILEY.
HE IS WORKING WITH THE
METROPOLITAN NEW YORK LIBRARY
COUNCIL, OR METRO, ON THEIR
STRATEGIC INITIATIVE WHICH
INCLUDE PROGRAM DEVELOPMENT,
RESEARCH AND PUBLICATION, NEW
TECHNOLOGIES, MEMBER SERVICES
AND OTHER EVENTS AND PROGRAMS.
BEFORE 2012 WHEN HE JOINED METRO
HE WORKED ON DIGITAL
PRESERVATION PROGRAMS FOR THE
NATIONAL DIGITAL INFORMATION
INFRASTRUCTURE AND PRESERVATION
PROGRAM AND THE OUTREACH AND
EDUCATION PROGRAM, PEOPLE CALLED
THAT DEPOT AND THAT'S AT THE
LIBRARY OF CONGRESS.

HE RECEIVED HIS MASTERS IN
LIBRARY INFORMATION SCIENCE WITH
A SPECIALTY IN ARCHIVAL STUDIES
FROM THE UNIVERSITY OF
PITTSBURGH AND HIS UNDERGRADUATE
DEGREES FROM OBERLIN COLLEGE.

I WANT TO THANK DANIELLE PLUMER
FOR BEING WITH US TODAY.
SHE'S OUR COURSE COORDINATOR AND
SHE WILL BE ASSISTING JEFFERSON
AND ANSWERING YOUR QUESTIONS IN
THE CHAT BOX AND WHEN HE TAKES
HIS BREAK FOR QUESTIONS SHE'LL
BE ANSWERING THAT AS WELL.

SO I'M GOING TO CLOSE MY POWER POINT, BRING UP JEFFERSON'S AND CLOSE OUR HELLO BOX AND WE'LL START OUR MODERATED CHAT.

AS MIKE MENTIONED, YOU WILL NOTICE IF YOU POST ANYTHING YOU'LL SEE IT TWICE, ONCE WHEN YOU LET US KNOW AND ONCE WHEN WE PUBLISH IT TO THE GROUP.

SO I'LL MOVE THAT AWAY AND IF YOU HAD ANY TECH ISSUES THAT YOU MENTIONED IN THE HELLO BOX WE'LL MAKE SURE TO DEAL WITH THAT.

SO WE'LL BRING YOU BACK TO THE BEGINNING HERE.

JEFFERSON?

SLIDE ONE?

>> YUP.

>> GREAT, I'LL TURN IT OTHER TO YOU.

THANK YOU.

>> OKAY, GREAT, THANKS, KRISTEN. I ALSO WANT TO SAY THANKS TO DANIELLE AND SAY THANKS TO HERITAGE PRESERVATION WHO'S PUTTING ON THIS GREAT SERIES ON IMLS SERIES.

SO WHAT I AM TALKING ABOUT TODAY IS PRACTICE SAFE ARCHIVING, BACKUPS, COPIES AND WHAT CAN GO WRONG.

SO WE'LL BE TALKING ABOUT DIGITAL MATERIALS, OBVIOUSLY DIGITAL PRESERVATION AS WELL AS HOW IT TIES IN TO PHYSICAL MEDIA WHICH IS, OF COURSE, HOW WE STORE ALL OUR DIGITAL INFORMATION.

WHAT CAN GO WRONG AND WHAT ACTIONS WE CAN TAKE TO ADDRESS THOSE ISSUES.

SO THIS IS TO TALK ABOUT THE WHOLE SERIES WRIT LARGE, THE CARING FOR DIGITAL MATERIALS GOALS.

PARTICIPANTS HAVE A BETTER UNDERSTANDING OF THE INHERENT

FRAGILITY OF DIGITAL OBJECTS,
WE'LL DEFINITELY BE TALKING
ABOUT THAT TODAY.

PARTICIPANTS WILL ACQUIRE
INFORMATION TO HELP AND SELECT
PRESERVATION FORMAT, METADATA
AND BACKUP SYSTEMS FOR DIGITAL
OBJECTS AND SOME PREVIOUS
PRESENTERS IN THE SERIES HAVE
TALKED ABOUT THAT.

WE'LL TALK ABOUT BACKUP MORE
TODAY BUT, OF COURSE, ALL THESE
THINGS WILL BE TOUCHED UPON.
AND THE THIRD IS PARTICIPANTS
WILL BE ABLE TO IDENTIFY ONE OR
MORE ACTIONS THAT CAN BE TAKEN
TO IMPROVE THEIR INSTITUTION'S
DIGITAL PRESERVATION EFFORTS.
WE'LL DEFINITELY BE TALKING
ABOUT THAT TODAY.

>>

AND THERE YOU SEE THE PREVIOUS
ONES SO THOSE WILL BE AVAILABLE
ONLINE AND WE'LL PUT THIS ONE UP
AND LIZ AND TOM WILL BE TALKING
NEXT WEEK ABOUT DIGITAL
PRESERVATION AND COLLABORATION.
SO OUR OUTLINE FOR TODAY'S
SESSION IS BASICALLY GOING TO BE
IN THREE PARTS.

THE FIRST PART I'LL TALK ABOUT
PHYSICAL MEDIA AND DIGITAL
INFORMATION.

THIS WILL TALK ABOUT SOME OF THE
CHALLENGES AS WELL AS THE
ACTIONS THAT CAN BE TAKEN TO
PRESERVE DIGITAL INFORMATION
THAT LIVES ON PHYSICAL MEDIA.

PART TWO WILL BE TALKING ABOUT
BACKUP AND STORAGE.

AND THEN PART THREE WE'RE GOING
TO TALK ABOUT THE LEVELS OF
DIGITAL PRESERVATION PROJECT AND
THIS IS A PROJECT THAT GOT
STARTED WHEN I WAS AT THE
LIBRARY OF CONGRESS OUT OF INDSA
WHICH IS THE NATIONAL DIGITAL

STEWARTSHIP ALLIANCE.
THAT'S A CONSORTIUM THAT ELSIE
OVERSEAS IN PEOPLE INTERESTED IN
DIGITAL PRESERVATION.
AND THE GOALS OF THE LEVELS OF
DIGITAL PRESERVATION PROJECT,
I'LL TALK ABOUT IT MORE WHEN WE
COME TO THAT PART OF THE
PRESERVATION.
BUT IT'S GENERALLY INTENDED TO
BE AN ACCESSIBLE GUIDE FOR
INSTITUTIONS OF ANY TYPE TO TRY
TO UNDERTAKE DIGITAL
PRESERVATION.
PRACTICES AND ACTIONS.
SO THE NEXT SLIDE.
SO I JUST WANT TO START OFF WITH
A COUPLE SORT OF HIGH LEVEL
SLIDES BUT I'M GOING TO TRY TO
KEEP ALL OF THIS VERY
ACCESSIBLE.
BUT I THINK THROUGHOUT THE WHOLE
SERIES OF THESE PRESENTATIONS WE
SORT OF ALL NEED TO KEEP IN MIND
LIKE WHAT IS A DIGITAL OBJECT.
WHAT IS DIGITAL INFORMATION.
AND THE MANY DIFFERENT FORMS
THAT IT TAKES AND MANY DIFFERENT
WAYS THAT WE INTERACT WITH IT.
SO YOU'LL SEE SOME PICTURES
HERE.
WE HAVE FLOPPYS AND DRIVE,
DIGITAL INFORMATION IS DEPENDENT
ON PHYSICAL OBJECTS AND STORAGE.
WE SEE SOME 1s AND 0s, THAT'S A
FAMILIAR WAY WE THINK ABOUT
DIGITAL INFORMATION AS BEING BUY
THEIR IS WHAT THAT'S REFERRED TO
IF YOU OPEN SOME IMAGES, SAY A
JPEG IN A TEXT EDITOR WHICH IS
THE THIRD IMAGE, IT BASICALLY
LOOKS LIKE GIBBERISH BECAUSE A
TEXT EDITOR DOESN'T KNOW WHAT TO
DO WITH AN IMAGE.
OBVIOUSLY WE HAVE IMAGES OF
GREAT THINGS LIKE BABY WALRUS
AND THAT CAN BE REPRESENTED IN

SOFTWARE, IT CAN BE REPRESENTED
IN CODE.

SO -- THE POINT OF THIS SLIDE IS
JUST SORT OF TO REMEMBER THAT A
DIGITAL OBJECT IN PRESERVING
THAT DIGITAL OBJECT REQUIRES US
TO INTERACT WITH ALL OF THESE
DIFFERENT LAYERS.

AND SO THE PHYSICALITY AND WHAT
WE'LL BE TALKING ABOUT TODAY IS
THE DIGITAL INFORMATION IS
STORED ON FLOPPYS AND HARD
DRIVES IN OUR COMPUTERS AND
SERVERS SOMEWHERE SO THEY'RE ALL
VERY DEPENDENT ON PHYSICAL
MATERIAL AND THIS IS ACTUALLY A
VERY LARGE MICROSCOPE SHOT OF
THE SURFACE OF THE HARD DRIVE
AND YOU CAN SEE THESE ARE
ACTUALLY WHAT WE CALL BITS.

A BIT IS SORT OF THE COMPONENT
PIECE OF DIGITAL OBJECT.

IT'S A 1 OR 0 AND YOU CAN SEE
THE TRENCHES AND VALLEYS AND THE
PEAKS ARE BASICALLY 1s AND THE
VALLEYS ARE 0s AND THAT'S
HOW DIGITAL INFORMATION IS READ
BY YOUR COMPUTER AND IT GOES
THROUGH MANY DIFFERENT
TRANSLATIONAL STATES UNTIL IT
COMES UP ESSENTIALLY ON YOUR
SCREEN TO AN OBJECT THAT YOU CAN
LOOK AT AND UNDERSTAND.

AND SO CLEAR AND FAR A GOT
TOGETHER TO THINK ABOUT THESE
SAME ISSUES THAT I'M TALKING
ABOUT AND THEY CAME UP WITH THIS
ON THEOLOGY OF DIGITAL OBJECTS.

SO ONTOLOGY MEANS WHAT ARE THE
PARTS OF A DIGITAL OBJECT AND
THIS SOUNDS A LITTLE FANCY IN
HIGH CONCEPT BUT IT'S IMPORTANT
TO REMEMBER THAT WHEN YOU'RE
DOING DIGITAL PRESERVATION YOU
ARE ADDRESSING EACH OF THESE
DIFFERENT PARTS OF AN OBJECT.

SO YOU HAVE THE PHYSICAL OBJECT

WHICH IS, OF COURSE, THE FLOPPY DRIVE OR HARD DRIVE, THE LOGICAL OBJECT, WHICH WOULD BE CODE AND SOFTWARE AND APPLICATIONS AND PROGRAMS THAT HELP YOU UNDERSTAND WHAT THAT -- WHAT IS ON THAT PHYSICAL OBJECT AND THEN YOU HAVE A CON IS UP WHICH YOU WILL OBJECT AND THAT, IS OF COURSE, AN IMAGE THAT YOU CAN LOOK AT OR A DOCUMENT THAT YOU CAN READ.

AND THEN THE THIRD-- AND I'VE SORT OF ADDED THIS-- IS THE CONDITIONAL OBJECT AND TO DO DIGITAL PRESERVATION AND THIS WILL COME OUT THROUGHOUT THE PRESENTATION.

YOU NEED TO STAY MINDFUL OF THE FACT THAT EACH INSTITUTION HAS ITS OWN DIFFERENT LEVEL OF RESOURCES, LEVELS OF EXPERTISE, LEVEL OF I.T. SUPPORT AND WHAT WE TRY TO DO AS DIGITAL PRESERVATIONISTS IS MAKE SURE THAT WE CAN DO THE BEST THAT WE CAN WITH WHAT WE HAVE.

BUT WHAT WE HAVE WILL ALWAYS BE VERY DIFFERENT.

THAT REALLY DRIVES A LOT OF THE DIGITAL PRESERVATION CONVERSATION AND BEST PRACTICES. SO WE'LL START WITH PART ONE AND TALK ABOUT WHAT CAN GO WRONG WITH THE PHYSICAL OBJECT AND SO THESE ARE, YOU KNOW, THE DISK IN THE DRIVES, AS I MENTIONED.

THERE'S A NUMBER OF THINGS THAT CAN GO WRONG WITH THEM.

OBSOLESCENCE.

THAT JUST MEANS SOMETHING IS OBSOLETE.

IT'S NO LONGER BEING PRODUCED OR THE MACHINERY REQUIRED TO ACCESS IT IS NO LONGER BEING PRODUCED SO THIS IS VERY EASY TO THINK ABOUT WITH A 5.25 FLOPPY DISK,

THE OLD BIG FLOPPY DISK.
OBVIOUSLY THOSE ARE NO LONGER
BUILT INTO COMPUTERS.
THEY'RE NO LONGER IN OUR LAPTOPS
SO IF YOU HAVE A COLLECTION ON
5.25 FLOPPY DISKS IT'S VERY
DIFFICULT FOR YOU TO GET TO THAT
INFORMATION.
ACCESS, SO OBVIOUSLY PHYSICAL
ITEMS CAN DEGRADE OVER TIME JUST
LIKE ANALOG OR PAPER CAN FALL
APART.
DIGITAL MEDIA HAS EVEN SHORTER
LIFE SPAN THAN MOST PHYSICAL
DOCUMENTS.
APPRAISAL, OBVIOUSLY IT'S VERY
DIFFICULT TO UNDERSTAND WHAT IS
ON A PIECE OF PHYSICAL MATERIAL
WITHOUT ACCESSING IT.
AND THEN AUTHENTICITY IS JUST
THE IDEA THAT YOU NEED TO ENSURE
THAT WHAT IS ON THAT MEDIA STAYS
THE SAME EVEN WHEN IT COMES OFF
OF THAT MEDIA.
AND THIS IS TRUE FOR PAPER, TOO.
I MEAN, XEROX COPIES, MICROFICHE
THESE ARE ATTEMPTS TO PRESERVE
AUTHENTICITY OF CONTENT
REGARDLESS OF WHAT FORMAT IT'S
ON.
THAT'S TRUE FOR DIGITAL
MATERIALS AS WELL.
SO I WANT TO HAVE A QUICK POLL
WHEN WE TALK ABOUT PHYSICAL
FORMATS AND JUST GET A SENSE OF
WHAT PEOPLE EITHER ARE
COLLECTING OR WHAT THEY HAVE
WITHIN THEIR INSTITUTIONS AND
I'VE SORT OF ADDED A RANGE OF
OPTIONS HERE, A NUMBER OF
DIFFERENT TYPES OF FLOPPY DISKS,
ZIP AND JAZZ DISKS WHICH ARE
BOTH ZIP DISKS ESSENTIALLY.
AND SD CARDS WHICH YOU PUT IN
YOUR COMPUTER AND THEN TAPES.
OPTICAL IS A VERY POPULAR ONE,
SO C.D.-s AND D.V.D.s AND THEN

EXTERNAL AND INTERNAL HARD
DRIVES WE'RE FAMILIAR WITH.
THEN ONLINE SERVERS.
SO WHEN YOU SAVE IT TO YOUR
NETWORK DRIVE THAT'S ACTUALLY
JUST ANOTHER DRIVE THAT HAPPENS
TO BE ATTACHED TO A NETWORK.
SO,, WOW, A LOT, THIS IS A GREAT
DIVERSITY.
A LOT OF THREE AND A HALFS.
WE ALL PROBABLY HAVE V.H.S. AND
AUDIOTAPES.
AND DEFINITELY A LOT OF OPTICAL
WHICH WE WERE CERTAINLY
EXPECTING.
AND A VERY LARGE NUMBER OF HARD
DRIVES.
WE'LL GIVE PEOPLE ANOTHER
MINUTE.
IT'S GREAT TO SEE SUCH A
DIVERSITY IN THE OLDER FORMATS
LIKE 5.25, 3.5 AND ZIPS.
THAT'S SURPRISING, ESPECIALLY
FOR A.V. MATERIAL.
SO MOVING ON THIS IS JUST
LOOKING ON AT THE DIFFERENT
MEDIA TYPES AND THEIR LONGEVITYS
JUST A QUICK NOTE TO QUANTIFY
THE LONGEVITY OF A PIECE OF
MEDIA IT DEPENDS ON HOW OFTEN
IT'S ACCESSED AND HOW T TYPE OF
EQUIPMENT IT PLAYS ON AND TAPE
WHICH PEOPLE HAD DOES NOT
SURVIVE WELL WITH REPEATED
REPLAY BECAUSE IT'S DEPENDENT ON
TENSION AND TENSION IS OBVIOUSLY
MOSTLY TO RIP THINGS APART, OF
COURSE, RIGHT?
SO THERE ARE JUST SOME NUMBERS
HERE THAT PEOPLE CAN LOOK
THROUGH TO AT LEAST GET AN
UNDERSTANDING OF WHEN I NEED TO
MIGRATE OFF OF THESE PHYSICAL
CARRIERS AND OPTICAL IS REALLY
THE BIG ONE BECAUSE THERE WAS A
LOT OF DIGITIZATION DONE AND
OPTICAL WAS OFTEN USED AS

STORAGE IN THE '90s SO A LOT OF INFORMATION PUT ON C.D.-Rs AND D.V.D.s FROM THOSE PROJECTS OF THE '90s ARE STARTING TO REACH THE END OF THEIR LONGEVITY. SO MIGRATING OFF OF OPTICAL IS SOMETHING I SEE A BIG ISSUE IN THE NEXT COUPLE YEARS. AND I'VE ALSO ADDED TWO LINKS TO EXCELLENT REPORTS AND THEN ALSO IN THE RESOURCES DOCUMENTS THAT ARE PART OF THE BROADCAST. SO YOU CAN LOOK AT THOSE AND THEY CAN THERE CAN BE A BETTER SENSE OF HANDLING AND STORAGE. SO WHAT ARE THE ACTIONS TO GET THINGS OFF OF PHYSICAL OBJECTS? VINTAGE DRIVES IN MACHINES ARE NOT TOO COMMON BUT I THINK WE'LL START SEEING AT LEAST COLLABORATIVE EFFORTS TO MAINTAIN AND HOUSE OLD COMPUTER SYSTEMS AND MACHINERY AND REPLAY DEVICES AS FAR AS ALLOWING ACCESS TO THOSE SORTS OF MEDIUM. AND THEN A COUPLE WE WANTED TO FOCUS NONTHIS PRESENTATION AND THESE ARE ALL TIED TOGETHER, CONTROL CARDS AND WRITE BLOCKERS. A CONTROLLER CARD IS ESSENTIALLY A DEVICE THAT ALLOWS YOUR CURRENT COMPUTER TO ACCESS AN OLD DRIVE. SO CONTROLLER CARDS ARE MOSTLY USED IN ORDER TO ACCESS 3.5 AND 5.25 FLOPPY DRIVE WHICH IS YOU CAN STILL BUY ON EBAY FOR ABOUT 50 BUCKS FOR A 5.25 DRIVE BUT OBVIOUSLY THERE'S NO WAY TO PLUG IT INTO YOUR COMPUTER SO THESE ARE BASICALLY THESE ARE INTERFACES THAT HAVE A U.S.B. ON ONE SIDE AND A CONTROLER TO THE FLOPPY DRIVE ON THE OTHER. A WRITE BLOCKER PROHIBITS YOU FROM ACTUALLY WRITING TO THE

DEVICE THAT YOU ARE TRYING TO GETTING ACCESS TO.

SO IT IS WHAT IT SAYS IT IS: IT BLOCKS WRITING TO THE OLD DRIVE AND THAT GOES BACK TO THE AUTHENTICITY POINT THAT WE WERE TALKING ABOUT EARLIER.

SO FORENSIC SOFTWARE, WE'LL TALK ABOUT THIS IN A COUPLE OF SLIDES.

THESE JUST SOFTWARE TOOLS TO HELP YOU DO SOME OF THESE ACTIVITIES.

A PHOTO STATION IS INTERESTING. WHEN PEOPLE GET DONATED PHYSICAL MEDIA THEY OFTEN OBVIOUSLY WANT TO HAVE A PICTURE OF IT BECAUSE PEOPLE WRITE ON THE OUTSIDE OF FLOPPYS, IT'S A VERY COMMON THING.

SO THERE'S CONTEXTUAL INFORMATION THAT CAN BE GAINED THERE.

SO TAKING PICTURE OF STORAGE IS NOT USUALLY THE CASE BUT WHEN NECESSARY CAPTURING AN IMAGE OF THE DEVICE IS OFTEN IMPORTANT. AND THEN A FRED IS A FORENSIC RECOVERY OF EVIDENCE AND IT BASICALLY COMBINES ALL THESE TOOL KITS INTO ONE BIG TOWER YOU SEE AT THE BOTTOM.

SO THESE ARE ALL METHODS OF TRYING TO ACQUIRE DATA OFF OF OLD PHYSICAL OBJECTS.

AND THEN I TALK ABOUT BACK LINE HERE AND WE'LL TALK ABOUT THIS MORE IN THE BACKUP AND STORAGE SECTION BUT STORAGE TRANSFER AND INFRASTRUCTURE WILL BE THINGS THAT WE'LL TALK ABOUT MORE.

SO A COUPLE OF OTHERS IS THE BIT CURATOR PROJECT AND THIS IS A PROJECT TO DEVELOP SOME OF THE FORENSIC SOFTWARE TOOLS THAT WE TALKED ABOUT.

SO IT IS VERY EMERGING AND NOT

REALLY USABLE IN AN EFFORT LEVEL AT THIS POINT BUT IT'S WORTH KEEPING AN EYE ON THAT PROJECT. AND WHAT IT DOES IS ESSENTIALLY ONE OF THE STAGES IS HELP CREATE DISK IMAGE.

SO A DISK IMAGE IS AN EXACT COPY OF A PIECE OF MEDIA.

AND THAT CAN BE C.D., C.D.-R, D.V.D., IT COULD BE A FLOPPY DRIVE OR AN ENTIRE COMPUTER SO IT JUST MEANS YOU ARE COPYING EVERYTHING THAT'S ON THERE AND IT'S A GOOD PRESERVATION METHOD BECAUSE YOU'RE NOT ACTUALLY TRYING TO ACCESS THE FILES.

YOU'RE ACTUALLY JUST MAKING A FULL COMPLETE COPY AND YOU CAN SORT OF PUT IT AWAY AND DECIDE WHAT TO DO WITH IT LATER.

VIRUS CHECKING IS SOMETHING THAT WILL BE DONE ON ANYTHING AND IT'S PROBABLY SOMETHING YOUR I.T. DEPARTMENT IS ALREADY DOING IN MANY CASES.

THERE ARE FREE TOOLS FOR THESE AND THEY'RE LISTED IN THE RESOURCES DOCUMENT AND THEY'RE GOOD TO USE.

AND FOR PEOPLE THAT ARE MORE INTERESTED IN THIS CHARGER AREA OF DIGITAL FORENSICS IS WHAT IT'S CALLED WE'VE ADDED A COUPLE LINKS TO GREAT REPORTS THAT ARE VERY USEFUL AS FAR AS WRAPPING YOUR HEAD AROUND HOW DO WE GET THINGS OFF OF OLD MEDIA.

SO THOSE ARE THE CHALLENGES OF PHYSICAL OBJECT.

ACCESS, OBVIOUSLY METADATA, DANIELLE TALKED ABOUT YESTERDAY, I WON'T GO INTO TOO MUCH DETAIL.

BUT, YOU KNOW, SOME KEY PRINCIPLES ARE AN INVENTORY.

THAT'S MOST COLLECTING INSTITUTIONS ARE ALREADY DOING THAT BUT IT'S IMPORTANT TO ALSO

DO IT FOR PIECES OF PHYSICAL MEDIA SO THAT YOU KNOW WHAT YOU HAVE.

HOW OLD IT IS, WHEN YOU MIGHT NEED TO MIGRATE OFF OF IT.

UNIQUE UNIVERSAL IDENTIFIERS. THIS JUST MEANS AN IDENTIFIER FOR AN OBJECT THAT CAN BE A COLLECTION OR IT CAN BE A FLOPPY DISK AND THAT IDENTIFYER WILL STAY THE SAME THROUGH TIME. THIS ALSO APPLIES TO SPECIFIC FILE NAMES SO IDENTIFIER FOR A FILE SHOULD REMAIN IN SAME REGARDLESS OF WHERE THAT FILE IS STORED SO IF IT'S STORED ON YOUR LAPTOP OR IF IT'S STORED ON THE NETWORK OR ON A D.V.D. THE IDENTIFIER SHOULD ALWAYS BE CONSISTENT.

LOCATE, THIS SORT OF GOES BACK TO THE INVENTORY PART.

KNOW WHAT YOU HAVE AND WHERE IT LIVES AND THIS IS APPLICABLE BOTH WITHIN YOUR PHYSICAL INSTITUTION BUT ALSO WITHIN YOUR DIGITAL ENVIRONMENT.

SO I'VE CALLED THAT CYBERSPACE AND MEAT SPACE.

ONE IS ONLINE AND ONE IS ON SITE SO THE LOCATION OF BOTH OF THOSE IS IMPORTANT.

DESCRIPTIVE INFORMATION.

YOU KNOW, WE'RE ALL PROBABLY FAMILIAR WITH THAT.

AND THEN THE PHOTOGRAPHING OBJECTS I TALKED ABOUT A GOOD BIT.

SO APPRAISAL AND AUTHENTICITY, THESE ARE ALSO KEY CONCEPTS WHEN TALKING ABOUT THE ACCESSIONING OR PRESERVATION OF PHYSICAL MEDIA THAT HAS DIGITAL INFORMATION.

YOU KNOW, WORK WITH YOUR DONORS IF YOU'RE WORKING IN AN INSTITUTION THAT IS ACCESSIONING

THINGS FROM DONORS.

DESCRIBE WHAT YOU HAVE WHEN IT COMES IN.

BACK IT UP AND JUST REMEMBER TO ALIGN WHAT YOU'RE ACQUIRING OR PRESERVING WITH WHAT YOUR COLLECTION POLICY IS AND WHAT YOUR INSTITUTIONAL ABILITIES ARE.

AND NOBODY CAN SAVE EVERYTHING, THIS IS A POINT WE OFTEN TRY TO MAKE IN DIGITAL PRESERVATION. PEOPLE FEEL VERY OVERWHELMED BY THE VOLUME OF INFORMATION THAT THEY HAVE AND THE KEY POINT IS TO TRY TO DO AS BEST YOU CAN IN SOME STANDARD BEST PRACTICES AND WE'LL TALK ABOUT THOSE IN THE LEVELS OF PRESERVATION SECTION WITH THE RESOURCES THAT ARE AVAILABLE TO YOU.

SO I'LL SORT OF STOP THERE TO SEE IF THERE ARE ANY QUESTIONS SO FEEL FREE TO PUT THEM IN THE Q&A BOX AND WE'LL TRY TO COMBINE SOME OF THEM SO THAT I CAN ADDRESS THEM.

>> SURE, WE DON'T HAVE TOO MANY QUESTIONS YET.

WE HAVE A LOT OF PEOPLE WHO ARE TALKING ABOUT THE DIFFERENT DIVERSE TYPES OF MATERIALS THEY HAVE IN THEIR COLLECTIONS. ONE OF THE COMMENTS OF OTHERS IS LPS, PHONOGRAPH GLASS SLIDES, WAX CYLINDERS IN SOME CASES. A FEW PEOPLE HAVE LASER DISKS. SO IT'S REALLY A LOT OF DIFFERENT KINDS OF THINGS SOME OF THOSE ARE NOT DIGITAL BUT NONETHELESS IT'S GOOD TO KNOW.

>> YEAH, WELL I THINK WHEN YOU'RE DIGITIZE A LOT OF THINGS LIKE CERTAINLY GLASS PLATES AND PRINT PHOTOGRAPHS AND LPS, AFTER YOU'RE DIGITIZING THEY'RE GOING THAT DIGITAL

OBJECT IS GOING TO GO TO SOME FORM OF MEDIA SO THOSE AREN'T CARRIERS OF DIGITAL INFORMATION WHICH IS MORE OF WHAT WE'RE FOCUSING ON BUT THEY WILL PLAY A ROLE IN DIGITAL PRESERVATION, ESPECIALLY AS PEOPLE DIGITIZE.

>> GREAT.

SO ONE OF THE QUESTIONS-- AND THERE ARE A COUPLE OF DIFFERENT THINGS THAT I'M SEEING RELATED TO THIS-- WOULD BE ABOUT MEDIA THAT'S BEEN ADVERTISED AS ARCHIVAL.

I'M SPECIFICALLY TALKING ABOUT GOLD C.D.s HERE, WHAT ITS EXPECTED LIFETIME IS.

AND THEN SOMEBODY ASKED IF ANYONE HAD ANY EXPERIENCE WITH THE M-DISK WHICH IS ANOTHER MEDIA THAT'S BEING SOLD AS ARCHIVAL.

WHAT ARE YOUR THOUGHTS ON THAT?

>> YEAH, I'M -- CERTAINLY ARCHIVAL GOLD STANDARD IS BASICALLY A SLIGHTLY LONGER LONGEVITY.

NOW, WHETHER IT'S WORTH THE ADDED COST, YOU KNOW, PEOPLE WILL HAVE DIFFERENT OPINIONS ABOUT THESE THINGS.

YOU KNOW, THE THING TO REMEMBER WITH ANY PIECE OF PHYSICAL MEDIA IS THAT IT IS EPHEMERAL.

SO IT'S -- I MEAN, AN ARCHIVAL GOLD C.D. MIGHT LAST A LITTLE BIT LONGER THAN A REGULAR C.D. BUT YOU WILL STILL HAVE TO GET YOUR INFORMATION OFF OF THAT AND ON TO A DIFFERENT MEDIA AT SOME POINT IN THE FUTURE.

SO, YOU KNOW, I TEND TO SHY AWAY FROM RECOMMENDING ARCHIVAL GOLD. IT DOES HAVE A LONGER LON JEFFTY BUT IT ALSO HAS A HIGHER PRICE AND THOSE ARE KIND OF INSTITUTIONAL DECISIONS AND THEY

NEED TO BE WEIGHED AGAINST EACH OTHER.

SO I THINK, YOU KNOW, FOR SOME INSTITUTIONS IT WILL PROBABLY BE THE RIGHT DECISION AND YOU WILL PROBABLY STORE IT LONG PAST ITS ACTUAL LONGEVITY, WHICH IS OKAY. AND FOR OTHERS THEY MIGHT BE MIGRATING TO AN ONLINE STORAGE ENVIRONMENT SOONER RATHER THAN LATER IN WHICH CASE THE COST IS NOT NECESSARILY JUSTIFIED.

M-DISK I'M LESS FAMILIAR WITH. I'VE HEARD OF IT AND KNOW IT HAS HIGHER STORAGE CAPACITY BUT I THINK WITH ANY OPTICAL MEDIA THE LOVE GENERALITY IS NOT THAT MUCH GREATER.

>> YEAH.

AND SOMEONE HAS JUST RESPONDED THAT UTAH STATE ARCHIVES IS ADOPTING IT AND WE'LL FIND OUT AS MORE PEOPLE START USING IT HOW THAT ONE DOES.

>> AND I GUESS ONE THING ALWAYS TO REMEMBER IS THAT IT STILL HAS TO PLAY ON SOMETHING.

SO EVEN IF THE DISK ITSELF LASTS LONGER, WELL, YOU STILL, YOU KNOW, OBVIOUSLY WE DON'T HAVE FLOPPY DRIVES IN OUR COMPUTERS ANYMORE AND IT'S VERY EASY TO THINK THAT AT SOME POINT IN THE FUTURE-- AND THIS IS OBVIOUSLY LOOKING VERY FAR DOWN THE LINE-- BUT THAT'S SOMETHING YOU NEED TO DO IN DIGITAL PRESERVATION, IT'S EASY TO THINK THAT ACCESS TO ANY SORT OF MEDIA HAS ITS OWN LIFE SPAN.

>> MOVING ON, WE HAD A COUPLE QUESTIONS ABOUT BIT CURATOR AND FORENSIC WORKSTATIONS IN GENERAL.

SO BRIEFLY, SOMEONE WANTED TO KNOW HOW A BIT CURATOR IMAGE IS DIFFERENT FROM JUST A NORMAL

BACKUP.

>> IT'S -- IT IS ACTUALLY ITS OWN FILE AND DISK IMAGE HAVE THEIR OWN FILE FORMAT. SO I GUESS YOU COULD THINK OF IT IN A WAY AND THIS IS -- I SHOULDN'T BE SAYING THIS, BUT AS A ZIP FILE.

SO IT'S AN ENTIRE PIECE OF INFORMATION IN -- COMPACTED INTO ONE DISK IMAGE.

THAT WILL BE A WHOLE FLOPPY DRIVE, IT COULD BE AN ENTIRE TWO TERABYTE DRIVE.

THE VOLUME DOESN'T MATTER IT'S JUST ESSENTIALLY MAKING AN EXACT COPY OF THE WHOLE THING.

SO A BACKUP WOULD JUST BE A COPY IT WOULD BE THE ENTIRE FILE SYSTEM, NOT NECESSARILY PACKAGED INTO ONE QUOTE/UNQUOTE OBJECT.

>> ALL RIGHT, SO ONE WAY TO THINK OF IT IS THAT NORMAL BACKUPS THINK ABOUT FILES AND THE BIT CURATOR IS ACTUALLY COPYING THE MEDIA, THE BITS ON THE MEDIA.

>> THAT'S A GOOD WAY TO PUT IT, YEAH.

>> OKAY, THERE'S SOME OTHER THINGS THAT WE'LL RESPOND TO AFTER THIS SESSION BUT WE'LL JUST KEEP GOING FOR NOW AND IF WE HAVE TIME AT THE END WE'LL COME BACK OR WE'LL ANSWER THESE IN WRITING AFTER THE SESSION.

>> OKAY.

SO MOVING ON IS I THINK DANIELLE TALKED ABOUT THIS AS PEOPLE WERE ON HER PEREZ STATION YESTERDAY BUT SIXTY.

THAT MEANS YOUR DIGITAL OBJECT HAS NOT CHANGED AND THAT IT IS FIXED AND IT'S CONSISTENT THROUGH TIME THROUGH WHEN YOU GOT IT THROUGH WHEN YOU LAST MADE A CHANGE TO IT IF THERE ARE

CHANGES AND THIS IS VERY IMPORTANT IN PRESERVATION BECAUSE, YOU KNOW, WHEN AN OBJECT IS SITTING ON A SERVER OR ON YOUR LAPTOP THERE'S NO VISUAL WAY TO TELL NECESSARILY THAT IT HAS CHANGED.

SOMETIMES THINGS WON'T OPEN OR THEY MIGHT LOOK FUNNY BUT THERE CAN BE CHANGES TO A DIGITAL OBJECT THAT AFFECT IT THROUGH TIME THAT ARE NOT FIZZABLE TO THE HUMAN EYE AND SO FIXITY IS ESSENTIALLY AS I PUT HERE A DIGITAL FINGERPRINT OF A FILE. AND WHAT IT IS IS AN ALGORITHM, WHICH IS JUST A LONG COMPUTER PROGRAM THAT COMPUTES AN IDENTIFIER.

AND SO THIS WILL BE A SORT OF LONG NUMBER -- IT'S A LITTLE BIT TO VIEW ON THE SLIDE BUT IT WILL BE NUMBERS AND LETTERS OF 24, 16 38 CHARACTERS AND SO WHAT YOU DO A SOFTWARE PROGRAM WILL RUN THIS ALGORITHM ON A FILE OR IT CAN RUN IT ON A WHOLE COMPUTER AND GENERATE A WHOLE LIST OF CHECK SUMS.

SO THE ALGORITHM PRODUCES ITS NUMBER AND IT'S OFTEN REFERRED TO AS A CHECK SUM.

SOMETIMES REFERRED TO AS A HASH, SOMETIMES REFERRED TO AS A MESSAGE DIGEST.

THESE ALL MEAN THE SAME THING AND IT'S BASICALLY JUST THIS DIGITAL FINGERPRINT AND SO THE UTILITY OF IT IS THAT WHEN YOUR OBJECT HAS CHANGED OR BECOME CORRUPTED THIS ALGORITHM WILL PRODUCE A DIFFERENT NUMBER.

SO THE NUMBER STAYS THE SAME ONLY AS LONG AS YOUR DIGITAL ITEM STAYS THE SAME.

AND SO YOU CAN RUN THESE ALGORITHMS AND THEN THEY WILL

PRODUCE CHECK SUM AND THEN YOU CAN RUN CHECK SUM AGAINST A PREVIOUSLY PRODUCED ONE LIKE ONE A MONTH AGO OR A YEAR AGO AND IF THEY'RE THE SAME YOU KNOW YOUR OBJECT IS STILL THE SAME AND IF THEY'RE DIFFERENT YOU KNOW IT'S CORRUPTED OR SOMEHOW CHANGED AND YOU CAN DO THIS WITHOUT ACTUALLY ACCESSING THE ITEM ITSELF.

SO THIS PROCESS OF AUDITING THEM IS OFTEN CALLED A FIXITY CHECK OR FIXITY AUDIT AND YOU CAN SORT OF UNDERSTAND CLEARLY WHY IT'S ESSENTIAL TO DIGITAL PRESERVATION IT'S BECAUSE IT'S THE PRESERVATION OF A DIGITAL OBJECT AND YOU NEED A SORT OF KEY WAY TO UNDERSTAND WHETHER SOMETHING HAS CHANGED.

SO OUR FIRST HOME WORK EXERCISE WILL SORT OF BE FOCUSED ON COMPUTING THE CHECK SUM AND THEN CHECKING IT.

BUT WE ALSO WANTED TO RUN THIS POLL HERE EARLY IN THE PEREZ STATION JUST TO GET A SENSE OF HOW MANY PEOPLE WERE ACTUALLY GENERATING THESE CHECK SUMS. THIS IS CALLED FIXITY INFORMATION.

THERE'S LOTS OF TERMS FOR IT BUT IT'S ESSENTIALLY AN IDENTIFIER AT THE POINT OF INGEST OR CREATION OF A DIGITAL FILE.

>>

SO SOMETIMES AN I.T. DEPARTMENT WILL DO THIS BUT IT IS USUALLY AN ARCHIVIST OR A COLLECTION LICE BRARN OR CURATOR OR PERSON IN CHARGE OF MANAGING DIGITAL ASSETS.

SO NOT BAD NUMBERS.

>>

SO IT'S GOOD TO SEE THAT, YOU KNOW, WE HAVE ABOUT 8% OR SO ARE ACTUALLY DOING THIS AND IF

YOU'RE GENERATING IT, YOU'RE MOST LIKELY ALSO AUDITING IT. THERE ARE SOME TOOLS THAT WE LISTED IN THE RESOURCES AS WELL AS SOME DOCUMENTS THAT YOU CAN READ ABOUT AND IT'S NOT AN OVERLY DIFFICULT PROCESS FOR PEOPLE TO GENERATE THIS INFORMATION AND HOW PERIODICALLY YOU RUN THESE FIXITY CHECKS OR FIXITY AUDITS IS AN INSTITUTIONAL DECISION BUT CERTAINLY GENERATING FIXITY INFORMATION IS A KEY COMPONENT OF PRESERVING DIGITAL ITEMS. BECAUSE YOU WILL NOT ALWAYS KNOW WHEN SOMETHING GOES WRONG. ALL RIGHT, WE CAN CLOSE THE POLL.

IT LOOKS GOOD.

AND SO HERE ARE JUST SOME SCREEN SHOTS AND I'LL GO THROUGH THIS QUICKLY AND THIS IS BASICALLY WHAT IT LOOKS LIKE.

IT'S A LITTLE HARD TO SEE BUT YOU CAN SEE THESE ARE EITHER SOFTWARE SYSTEMS OR MACHINE ENVIRONMENTS RUNNING THROUGH A WHOLE FILE SYSTEM AND THEY GENERATE THESE CHECK SUMS AND IT'S JUST A LONG STRING OF DIGITS AND IT CREATES A REPORT AND AT A LATER POINT YOU CAN RUN SOFTWARE ON THE SAME SET OF ITEMS.

IT WILL CREATE A NEW REPORT, COMPARE THEM AND TELL YOU WHEN SOMETHING HAS CHANGED.

AND SO, YOU KNOW, WHAT THIS GUARDS AGAINST, FIXITY, IS OF COURSE BITROT.

YOU SEE IN THIS IMAGE THIS IS SOMETHING THAT THROUGH BEING UPLOADED ON LINE HAS SUFFERED BIT ROT AND TRANSFERRING DIGITAL FILES OFTEN RESULTS IN SOMETHING HAPPENING TO THE FILE THAT CAN

AFFECT IN THE THIS MANNER AND SOMETIMES IT'S NOT QUITE AS APPARENT AS THIS AND SOMETIMES IT'S ENTIRELY INVISIBLE AND THAT'S SORT OF THE IMPORTANT POINT OF FIXITY.

SO I'VE LISTED A COUPLE OF SOFTWARE TOOLS THAT ARE FREE AND CAN BE DOWNLOADED TO PLAY AROUND WITH.

AND BAGGER IS ALSO A LIBRARY OF CONGRESS TOOL FOR TRANSFERRING DATA AND IT ALSO GENERATES CHECK SUMS WHEN USING IT AND THERE ARE SOME GUIDES AND VIDEOS THAT WE LINK TO IN THE RESOURCES THAT ARE WORTH LOOKING AT IF YOU'RE INTERESTED.

THEN I TALK ABOUT FIXITY CHECKING AND AUDIT THERE IN SOME OF THOSE BULLET WHICH IS WE ALREADY DISCUSSED.

SO I'M GOING TO RUN THROUGH THIS PART A LITTLE BRIEFLY AND IT'S MOSTLY ABOUT THE LOGICAL OBJECT SO IF WE THINK ABOUT THAT, THAT ESSENTIALLY MEANS, YOU KNOW, THE SOFTWARE, THE APPLICATION, THE COMPUTER CODE THAT CAN TAKE THAT PHYSICAL MEDIA AND TURN IT INTO SOMETHING LIKE AN IMAGE OR A TEXT THAT CAN BE READ.

SO THERE'S A SORT SO MIDDLE STAGE THERE AND THERE'S A LOT THAT CAN GO WRONG THERE JUST AS THERE IS WITH PHYSICAL MEDIA, OBSOLESCENCE, WE ALL KNOW ABOUT OLD FILE FORMATS AND YOU WILL TRY TO OPEN SOMETHING ON YOUR COMPUTER AND IT DOESN'T OPEN BECAUSE THE SYSTEM DOESN'T KNOW WHAT IT IS.

AND THIS IS A VERY COMPLICATED PART OF DIGITAL PRESERVATION WHICH IS NOT WHY WE'RE GOING TO SKIM THROUGH IT HERE BUT IT'S JUST SOMETHING TO KEEP IN MIND

AS YOU TRY TO ESTABLISH A PRESERVATION PROGRAM WITHIN YOUR INSTITUTION.

SO ACCESS I MENTIONED, APPRAISAL OBVIOUSLY IF YOU ARE CREATING OR TAKING FROM DONORS' DIGITAL MATERIALS, WHAT FORMAT IT'S ON WILL BE VERY IMPORTANT WHETHER YOU CAN ACCESS IT.

AND AWE THEN THYS CITY VERY SIMILAR TO THE AUTHENTICITY ISSUE THAT WE TALKED ABOUT WITH PHYSICAL MEDIA.

SO THE CHALLENGERS ARE RECOGNIZING AND PROCESSING ALL THESE MANY DIFFERENT TYPES OF FILES IS VERY DIFFICULT AND COMPLICATED.

AND THAT A FORMATELL IT IS COMPUTER ESSENTIALLY WHAT TO DO IN ORDER TO ACCESS THAT OBJECT BUT IT'S VERY DIFFICULT TO PRESERVE THAT INFORMATION.

SO I'LL SKIP THROUGH THESE AND I GUESS THE KEY POINT TO MAKE HERE IS JUST TO TRY TO REMAIN COGNIZANT THAT THERE ARE OPEN FILE FORMATS LIKE TIFF AND PDF-A IS A PRESERVATION THAT'S PDF. SO SAVING THINGS IN OPEN FILE FORMATS THAT MEANS THAT THE COMMUNITY KNOWS WHAT THE PROGRAMMING IS THAT ALLOWS IT TO BE OPENED AND IF THE COMMUNITY KNOWS IT THEN IT WILL BE OBVIOUS TO PEOPLE THROUGHOUT TIME.

WHEREAS, YOU KNOW, SPECIFIC PIECES OF VENDOR COMMERCIAL SOFTWARE WE DON'T ALWAYS KNOW WHAT THE CODE IS BECAUSE THAT'S WHERE THEY MAKE THEIR MONEY. SO THE ACTION IS JUST TO REMAIN COGNIZANT OF OPEN FORMATS.

AND SO THEN FINALLY AND THIS WILL BE THE LAST PART OF THIS SECTION IS THE CONDITIONAL OBJECT.

I TALKED ABOUT THIS A LITTLE AT THE BEGINNING AND THAT'S JUST REMEMBERING THAT FOR EVERY INSTITUTION THERE'S DIFFERENT LEVELS OF RESOURCES AND KNOWLEDGE AND INSTITUTIONAL KNOWLEDGE AND INSTITUTIONAL PRACTICES THAT CAN BE DEDICATED TOWARDS DIGITAL PRESERVATION. AND SO SORT OF OUR JOB AS PRESERVATIONISTS OR CURATORS ARE YOU KNOW, NONPROFIT WORKERS WORKING WITH CULTURE HERITAGE MATERIAL OR OTHER DIGITAL INFORMATION IS JUST TO DO WHAT WE CAN WITH OUR INSTITUTIONAL RESOURCES IN ACCORDANCE TO COMMUNITY-ACCEPTED STANDARDS AND BEST PRACTICES.

AND SO THAT'S THE CONDITIONAL OBJECT.

AND HERE ARE SOME FRAME WORKS OF A DIGITAL PRESERVATION SYSTEMS. NOW, THESE ARE VERY COMPLEX AND HIGH LEVEL BUT WE'VE SORT OF DONE SOME EXERCISES AND SOME OF THE RESOURCES SO THAT YOU CAN AT LEAST LOOK THROUGH THEM TO TRY TO GET A SENSE OF WHAT PEOPLE -- THE LARGE FRAMEWORKS THAT BIG INSTITUTIONS ARE USING HAVE A LOT OF VALUE AS FAR AS TELLING YOU WHAT THEY'RE DOING IN A VERY SIMPLE WAY AND I KNOW THESE DIAGRAMS DON'T LOOK SIMPLE AT ALL!

BUT IF YOU ACTUALLY READ THROUGH SOME OF THEIR REQUIREMENTS WHEN YOU LOOK AT THE DOCUMENTATION FOR THESE FRAMEWORKS YOU'LL PULL OUT LITTLE PIECES THAT CAN HELP YOU.

IMPLEMENT POLICIES WITHIN YOUR INSTITUTION REGARDLESS OF WHETHER YOU'RE ACTUALLY GOING TO MEET ALL THE REQUIREMENTS IN ONE OF THESE FRAMEWORKS.

AND I CALLED THIS A SERIES OF
NOWS BECAUSE I THINK THAT'S A
NICE PHRASE FOR THINKING ABOUT
DIGITAL PRESERVATION AND BACKUPS
AND STORAGE BECAUSE AS WE TALKED
ABOUT WITH SOME OF THE -- IN THE
PREVIOUS SLIDES, THIS IS
SOMETHING THAT HAPPENED THROUGH
TIME, RIGHT?

SO THINGS WILL ALWAYS COME
OBSOLESCE OR WILL CHANGE
FORMATS AND MEDIAS AND WHAT
WE'RE TRYING TO DO-- AND THIS
TIES INTO THE INSTITUTIONAL
PIECE-- IS TO TRY TO DO THE BEST
WE CAN RIGHT NOW AND THAT SORT
OF NOW WILL ALWAYS BE CHANGING
AND IT WILL ALWAYS BE NEVER
ENDING ESSENTIALLY.

BUT DIGITAL PRESERVATION IS ALL
ABOUT ACCESS INTO THE FUTURE AND
SO NOBODY CAN TELL THE FUTURE SO
WHAT WE TRY TO DO IS WHAT WE CAN
RIGHT NOW.

SO THAT FINISHES PART ONE.

SO DANIELLE IF YOU WANT TO TRY
TO ORGANIZE SOME QUESTIONS WE
CAN MAYBE TALK ABOUT IT.

WE'LL GET A LITTLE MORE SPECIFIC
ON BACKUPS AND STORAGE NEXT SO
YOU CAN SAVE THOSE FOR THAT
SECTION.

>> YEAH, SO I THINK PEOPLE WERE
TALKING AND THERE WAS SOME
INFORMATION GOING BACK AND FORTH
IN CHAT, SOME GOOD ADVICE THAT
CAME OUT OF THIS.

BUT IN GENERAL PEOPLE WERE
ASKING ABOUT PROCEDURES FOR
FIXITY AND SOMEONE PHRASED IT AS
SO IF I HAVE TWO COPIES OF IT
AND ONE COPY GOES BAD, THE
FIXITY CHANGES, THEN I GET RID
OF THE OLD ONE OR THE BAD ONE
AND I REPLACE IT WITH THE GOOD
ONE.

IS THAT KIND OF HOW YOU WOULD

SUMMARIZE IT OR DO YOU HAVE
OTHER THOUGHTS ABOUT HOW --
>> YES.

>> -- FIXITY WORKS IN A
PRACTICAL SENSE.

>> YEAH.

SO WE'LL TALK ABOUT THOSE IN
BACKUP AND STORAGE BUT OBVIOUSLY
MULTIPLE COPIES IS KEY COMPONENT
OF DIGITAL PRESERVATION, RIGHT?
I MEAN, MULTIPLE COPIES,
MULTIPLE PLACES.

SO, YES, WHEN YOU'RE RUNNING
FIXITY CHECK YOU CAN RUN THEM --
YOU'LL RUN THEM SEPARATELY ON
DIFFERENT SETS OF ITEMS BUT AS
FAR AS A WORK FLOW IT WILL BE AT
THE FRONT OF THE WORK FLOW AND
THEN WHEN CORRUPTION IS DETECTED
YOU WOULD REPLACE IT WITH THE
AUTHENTIC COPY.

>> AND CATHERINE FOLLOWED UP
WITH "IS THERE ANY WAY TO
REVERSE BIT ROT IN A CORRUPTED
FILE?"

>> THERE IS NOT.

>> OKAY.

>> YEAH, BASICALLY NO.

IF YOU HAVE AN ORIGINAL COPY YOU
COULD BUT IF YOU HAVE AN
ORIGINAL COPY YOU WOULD JUST BE
REPLACING IT, YOU WOULDN'T
REPAIR THE OTHER ONE.

THE GREAT THING ABOUT DIGITAL
MATERIAL IS THEY'RE EASY TO COPY
AND REPLICATE.

SO ONCE YOU DETECT CORRUPTION
YOU WOULD GENERALLY DELETE THAT
FILE AND REPLACE IT.

>> SO THIS IS ONE OF THE REASONS
WE TALK ABOUT DIGITAL OBJECTS
BEING FRAGILE IS THAT WE CAN'T
DO REPAIR FOR THE MOST PART.

>> RIGHT.

YES.

AND THAT'S WHY FIXITY IS SORT OF
AT THE FRONT OF A WORK FLOW OR

INGEST OR ACCESSIONING, HOWEVER YOU WANT TO SAY IS SO IMPORTANT, RIGHT?

BECAUSE AS SOON AS YOU GET THEM YOU WILL BE WANTING, IF YOU CAN, TO GENERATE FIXITY INFORMATION SO THAT YOU CAN THEN TRACK IT AND AUTHENTICATE IT WHILE THE OBJECT IS IN STORAGE OR IF IT GETS PASSED AROUND TO BE DESCRIBED OR IF IT GOES INTO DIFFERENT ACCESS SYSTEMS OR COLLECTION MANAGEMENT SYSTEMS. SO, YEAH, INCORPORATING FIXITY AND, AGAIN, IT'S NOT TOO OVERWHELMING A PROCESS OR PROCEDURE AND IF YOU HAVE AN I.T. DEPARTMENT THEY WILL BE FAMILIAR WITH THIS CONCEPT AND WITH SOME OF THESE TOOLS SO CERTAINLY WORTH TALKING TO THEM.

>> OKAY, WELL, I THINK THAT -- THEN WE HAD SOME QUESTIONS ABOUT SOME SPECIFIC TOOLS AND A QUESTION ABOUT CHUCK SUMS ON MACINTOSHS, PEOPLE WERE REFERRING TO SPECIFIC TOOLS THAT THEY'VE USED OR HAVE INFORMATION ABOUT.

YOU WANT TO TALK ABOUT SOME OF THE ONES YOU BRIEFLY PUT IN THE RESOURCES?

>> SURE, I PUT ONE IN THE RESOURCES BOTH FOR WINDOWS AND FOR Mac.

YOU KNOW, THE ONE THING ABOUT FIXITY I'LL NOT IS THERE ARE NOT A LOT OF TOOLS WITH GRAPHICAL INTERFACES.

IT'S USUALLY SOMETHING THAT'S BUILT INTO EVERY OPERATING SYSTEM AS FAR AS A COMMAND LINE FUNCTION.

AND THAT MEANS WHERE YOU'RE ACTUALLY ENTERING THE COMMANDS INTO A TERMINAL WINDOW IS WHAT IT'S CALLED ON MAC.

IT'S THE COMMAND PROMPT ON
WINDOWS MACHINES.
BUT I DID PUT TWO PIECES OF
SOFTWARE ON THAT THERE THAT ARE
BOTH FREE AND VERY EASY TO USE
AND UNDERSTAND.
MD-5 IS THE ONE FOR MAC AND MD-5
SUMMER IS THE ONE FOR WINDOWS SO
THOSE CAN BE USED BY PEOPLE ON
THEIR DESKTOPS TO GENERATE AND
AUDIT.
THEY BOTH HAVE AUDIT FUNCTIONS
THAT YOU CAN RUN.
THEY'LL CREATE A FILE ONCE THEY
RUN A FIXITY CHECK.
SO THOSE ARE GOOD.
AND THEN THERE ARE OTHER ONES
OUT THERE AS FAR AS COMMAND LINE
USE.
>> AND THEN NICOLE ASKS THIS
QUESTION AND HONESTLY I'VE NEVER
TRIED TO DO THIS ON A LARGE DATA
SET OF ANY KIND.
SO ANY RECOMMENDATIONS ON
SOFTWARE TO CREATE FIXITY ON
LARGE AMOUNTS OF DATA UP TO 1
TERABYTE?
HAVE YOU TRIED THAT?
>> YEAH, THE COMMAND LINE TOOLS
WOULD WORK BETTER FOR THAT.
I MEAN, JUST LIKE ANY PIECE OF
CODE OR SOFTWARE, IT WILL TAKE
TIME BUT, YES, I'VE RUN FIXITY
CHECKS ON A TERABYTE WORTH OF
MATERIAL.
IT WILL TAKE A WHILE, BUT NOT,
YOU KNOW, HOURS.
YOU KNOW, IT -- I DON'T KNOW
WHAT RATE IT CREATES A CHECKS ON
THAT BUT IT'S A LOT.
>> (LAUGHS)
>> I MEAN, IT'S PROBABLY 100
FILES A SECOND, MAYBE.
>> DIFFERENT CHECK SUMS I THINK
WORK AT DIFFERENT SPEEDS, BUT
YEAH.
>> YEAH, THAT'S TRUE.

>> SOMEONE ELSE -- BRYCE ASKS
WILL DIFFERENT CHECK SUM
GENERATING SOFTWARE ALWAYS
CALCULATE IDENTICAL CHECK SUMS?
>> YES, I TALKED ABOUT THE
ALGORITHM WE USED.
CHECK ALGORITHM WILL ALWAYS
PRODUCE THE SAME CHECK SUM.
SO THERE'S MD-5 IS ONE POPULAR
ONE.
THAT'S WHEN YOU ENTER A PASSWORD
INTO FACEBOOK OR TWITTER OR
SOMETHING THE WAY THAT IT
ENCRYPTS YOUR PASSWORD IS
THROUGH MD-5.
SO WHEN WE HEAR NEWS STORIES
ABOUT PEOPLE'S PASSWORDS GETTING
STOLEN ESSENTIAL LIVE WHAT
PEOPLE HAVE DONE IS STEAL A
DATABASE OF CHECKSUMS AND THEN
BECAUSE MD-5 HAS BEEN AROUND FOR
A LONG TIME SUPREME ACTUALLY
MANAGED TO FIGURE OUT WHAT SOME
OF THE ALGORITHMS ARE.
SO WHEN YOU HEAR ABOUT THESE
PASSWORD HACKS IT'S ACTUALLY
PEOPLE JUST FIGURING OUT WHAT AN
MD-5 CHECKSUM IS.
SO I GUESS THE RECOMMENDATION
THERE IS TO USE SHA-256 BUT, YOU
KNOW, OBJECTS WITHIN OUR
CULTURAL HERITAGE INSTITUTIONS
ARE NOT THINGS THAT PEOPLE ARE
REALLY TRYING TO GAIN A LOT OF
ACCESS TO A CHECKSUM FOR.
>> (LAUGHS)
>> I WOULDN'T WORRY ABOUT M.D.-5
AS FAR AS ITS SECURITY FOR THE
KIND OF INFORMATION WE'RE
TALKING ABOUT.
>> RIGHT.
THAT MIGHT BE A FUTURE PROBLEM
BUT NOT RIGHT NOW, ANYWAY.
>> YEAH.
>> OKAY, SO I THINK SOME OF THE
OTHER QUESTIONS WHICH GET INTO
THINGS LIKE HOW CAN WE MINIMIZE

THE LIKELIHOOD OF BIT ROT, ET CETERA, YOU'LL ADDRESSED A DRESS AS WE GO INTO BACKUPS.

>> SO WE WERE GOING TO START WITH A POLL TO SEE HOW PEOPLE'S PRACTICES ARE.

THIS IS HOW MANY COPIES OF DIGITAL CONTENT DOES YOUR INSTITUTION KEEP AND IN HOW MANY DIFFERENT GEOGRAPHIC LOCATIONS DOES IT KEEP IT IN?

SO WE HAVE A COUPLE OF OPTIONS THERE.

>>

WELL, THIS IS PRETTY EXCITING. TWO COPIES; TWO PLACES IS THE MINIMUM STANDARD BUT IT'S GREAT TO SEE SO MANY PLACES DOING IT BECAUSE NOT MANY I HAVE THAT ENCOUNTERED ARE ACTUALLY DOING TWO PLACES.

>>

SO THAT'S GOOD.

ALMOST HAVE THE RESPONDENTS WITH TWO COPIES, TWO PLACES I'M HAPPY TO SEE PEOPLE HAVING MULTIPLE COPIES WHICH IS IMPORTANT AND TWO PLACES IS REALLY IMPORTANT BECAUSE I'M NOT GOING TO TALK ABOUT IT TOO MUCH TODAY BUT OBVIOUSLY THE ENVIRONMENTAL HAZARD AND THE INSTITUTIONAL HAZARD AND WE HAD HURRICANE SANDY IN NEW YORK AND WE HAVE HAD DISASTERS IN THE PAST SO TWO PLACES ARE VERY IMPORTANT AND THAT SHOULD BE TWO PLACES OUTSIDE OF YOUR BUILDING, HOPEFULLY.

SO I DIDN'T DEFINE THAT TOO SPECIFICALLY BUT WE CALL THAT COLLOCATION SO THEY SHOULD NOT BE LOCATED IN THE SAME SPOT. THAT'S KEY TO TWO PLACES. AND SO IN BACKUP AND STORAGE WE'LL TALK A LITTLE BIT MORE LOW LEVEL AND KEEP IT SIMPLE WITH

BEST PRACTICES.

AND SO I SORT OF HAVE THIS
TRIPLE DEUCES RULES WHICH THREE
2s.

SO TWO COPIES IN TWO PLACES,
WHICH IT WAS GREAT TO SEE
EVERYBODY DOING AND TWO MEDIA
TYPES IS IMPORTANT, TOO, RIGHT?
WE TALKED EARLIER ABOUT THE
LONGEVITY AND DEGRADATION LIFE
CYCLE OF SPECIFIC FORMS OF MEDIA
SO IF YOU PUT TWO THINGS ON THE
SAME PIECES OF MEDIA THAT ARE
HYPOTHETICALLY DEGRADING AT THE
SAME TIME OR BECOMING OBSOLETE
AT THE SAME TIME THEN YOU SORT
OF -- YOU KNOW, THERE'S A
CERTAIN HAZARD TO THAT.

SO MOST PEOPLE WILL HAVE
SOMETHING ON, LIKE, SPINNING
DISK WHICH IS JUST A HARD DISK
EITHER EXTERNAL, INTERNAL OR ON
A SERVER AS WELL AS, SAY,
OPTICAL DISKS.

SO THREE DEUCES IS SORT OF THE
MINIMAL STANDARD FOR BACKING UP
DIGITAL INFORMATION.

TWO COPIES, TWO PLACES AND TWO
TYPES OF MEDIA.

THEN THREE COPIES OBVIOUSLY IS A
SLIGHTLY BETTER STANDARD FOR THE
OBVIOUS REASONS THEN ALSO TWO
PLACES AND TWO MEDIA TYPES.

AND SO THE OTHER IMPORTANT THING
-- WE TALKED ABOUT THIS EARLIER
WHEN WE WERE TALKING ABOUT
PHYSICAL LOGICAL OBJECTS IS
KNOWING WHAT YOU HAVE AND, YOU
KNOW, HAVING WORKED WITH A LOT
OF CULTURAL HERITAGE
INSTITUTIONS, ESPECIALLY SMALLER
ONES AND NONPROFITS AS WELL THEY
MIGHT BE PRESERVING THINGS ON
MULTIPLE COPIES AND MULTIPLE
PLACES BUT THEY DON'T -- YOU
KNOW, IF YOU PULL OUT A C.D.-R,
IT'S OFTEN VERY DIFFICULT FOR

SOMEONE TO TELL YOU WHAT IS ON THERE BOTH LOOKING AT IT AND WHEN YOU PUT HIT IN THE COMPUTER SO INVENDORRYING WHAT IS WHERE BOTH ON THE DISK, ON THE PIECE OF MEDIA AS WELL AS WHERE IT IS IN THE BUILDING OR OUTSIDE THE BUILDING.

THOSE ARE ESSENTIAL CONCEPTS. THEN FIXITY WE TALKED ABOUT IS SOMETHING THAT PEOPLE SHOULD BE DRIVING TO GENERATE WHEN THEY'RE PRESERVING DIGITAL INFORMATION. AND DETAILS OF THESE ARE DEPENDENT ON YOUR INSTITUTIONAL RESOURCES AND INSTITUTIONAL REQUIREMENTS.

BUT CERTAINLY THE TRIPLE DEUCES IS WHERE TO START AND THEN THREE COPIES IF YOU CAN MANAGE IT.

AND SO WHAT ARE THE TYPES OF STORAGE THAT PEOPLE USE?

THERE ARE BASICALLY THREE TYPES: ONLINE, NEAR-LINE AND OFFLINE. ONLINE IS ESSENTIALLY NETWORK ATTACHED STORAGE SO THIS MEANS YOU HAVE IMMEDIATE ACCESS TO IT.

A LOT OF US HAVE THE NETWORK DRIVES IN OUR OFFICES AND SO THAT IS ONLINE STORAGE BOTH WITHIN THE OFFICE ENVIRONMENT. SOMETHING, YOU KNOW, IF YOU HAVE WEB ACCESS YOU OBVIOUSLY HAVE IMMEDIATE ONLINE RESOURCE TO SOMETHING YOU MIGHT BE STORING ON AN EXTERNAL SERVER OR IN THE CLOUD OR EVEN, YOU KNOW, SAY IN GOOGLE DOCS IF YOU'RE DOING SOMETHING LIKE THAT.

SO ONLINE IS BASICALLY MEANS IMMEDIATE ACCESS.

AND THEN OFFLINE OBVIOUSLY MEANS YOU DON'T IMMEDIATE ACCESS TO IT.

AND THAT'S OFTEN CALLED A DARK ARCHIVE, IF PEOPLE HAVE HEARD THAT TERM WHERE YOU CAN STORE

THINGS THAT YOU DON'T EXPECT TO BE AVAILABLE TO YOU NOT JUST IMMEDIATELY BUT REALLY WITHIN ANY REASONABLE RESPONSE TIME. THEN NEAR-LINE IS SORT OF IN BETWEEN THOSE THINGS SO IT'S NOT ATTACHED TO THE NETWORK SO IT DOES NOT HAVE DIRECT AVAILABILITY BUT YOU CAN GET IT. A LOT OF UNIVERSITIES HAVE NEAR-LINE STORAGE FOR PRESERVATION IN THE SPACE. IT MEANS YOU CAN GET IT WITHIN LIKE A DAY OR TWO IF YOU NEED A LARGE SET OF IMAGES OR PUBLICATIONS OR SCHOLARLY PUBLICATIONS SO THOSE ARE OUR THREE TYPES OF STORAGE AND QUESTIONS TO ASK IN YOUR INSTITUTIONS WHEN YOU'RE MAKING DECISIONS AROUND THIS, OBVIOUSLY A LOT OF US OPERATE MOSTLY IN THE ONLINE OR NEAR-LINE STAGE IS HOW OFTEN DO YOU NEED TO ACCESS THIS MATERIAL, ARE PRESERVATION COPIES SEPARATE FROM YOUR ACCESS COPYS?

A LOT OF PEOPLE HAVE THEIR IMAGES ONLINE, THOSE ARE NOT NECESSARILY THEIR PRESERVATION COPIES OF THEIR DIGITAL OBJECTS. HOW ARE YOUR PRESERVATION ACCESS COPIES CREATED AND/OR MANAGED? A LOT OF PEOPLE KEEP THEIR PRESERVATION COPYS IN AN OFFLINE STORAGE ENVIRONMENT AND ACCESS COPIES WILL BE IN THEIR OPAK OR UP ON FLICKR OR SOMETHING LIKE THAT.

AND HOW DO SYSTEMS AND WORK FLOWS PLAY IN WITH THESE OTHER SYSTEMS?

SO OFFLINE IS GOING TO BE SOMETHING YOU'LL MANAGE WITH SPECIFIC PIECES OF SOFTWARE AND ONLINE IS SOMETHING YOU MIGHT WANT TO HAVE TIED INTO YOUR

CONTENT MANAGEMENT SYSTEM OR
YOUR OPAK.

SO THOSE ARE THINGS TO THINK
THROUGH WHEN TALKING ABOUT
STORAGE.

AND SO WE WANTED TO PUT THIS
SLIDE IN HERE, THE TECH STACK
JUST TO SORT OF GIVE PEOPLE AT
LEAST EXPOSURE TO SOME OF THE
VERBIAGE WHEN YOU'RE WORKING
WITH I.T. I KNOW A LOT OF PEOPLE
ON HERE PROBABLY ARE THEIR OWN
I.T. DEPARTMENT AND SO SOME OF
THIS MIGHT NOT BE SUPER
APPLICABLE BUT WHATEVER YOU'RE
WORKING WITH I.T. FOLKS, BE IT
INTERNAL OR CONTRACTORS IT'S
JUST GOOD TO AT LEAST HAVE SOME
EXPOSURE TO SOME OF THE KEY
TERMINOLOGY BECAUSE IT REALLY
DICTATES A LOT OF WHAT YOU CAN
DO WHEN YOU'RE DOING DIGITAL
PRESERVATION AS WELL AS WHEN
YOU'RE BUILDING ONLINE
COLLECTIONS AND OPAKS AND
CATALOGS AND SO A COUPLE --
WE'LL GO THROUGH THE TERMINOLOGY
QUICKLY.

A LAMP STACK IS SOMETHING YOU
OFTEN HEAR AND THAT'S WHAT THIS
IMAGE IS.

THAT'S BASICALLY A -- IT'S SORT
OF AN ENVIRONMENT FOR HAVING A
SERVER.

AND A SERVER WILL HAVE LINUX,
WHICH IS A COMPUTING CODE THAT
TELLS THE SERVER WHAT TO DO,
APACHE IS BASICALLY SORT OF THE
TYPE OF OPERATING SYSTEM THAT'S
ON THERE, MYSQL IS A DATABASE SO
I'LL PUT THE INFORMATION INTO
THE SERVER AND STRUCTURE IT AND
THEN USUALLY IT'S PHP.

THE P_s ARE PYTHON.

THESE ARE CODE BASICALLY SORT OF
MIDDLE WARE DHED WILL ALLOW WHAT
YOU'RE DOING ON THE INTERNET OR

YOUR DESKTOP TO INTERACT WITH THAT SERVER.

SO THINGS LIKE OMECCA IF PEOPLE ARE FAMILIAR WITH THAT, THEY CAN ONLY BE BUILT IN THIS LAMP STACK ENVIRONMENT BECAUSE THEY NEED THE SPECIFIC PIECES IN ORDER TO OPERATE.

SO IF YOU HAVE, SAY, A WINDOWS SERVER YOU MIGHT NOT BE ABLE TO INSTALL CERTAIN PIECES OF SOFTWARE.

SO THAT'S GOOD TO TALK TO I.T. ABOUT THAT.

RAID IS ESSENTIALLY A DEVICE THAT HAS MANY DIFFERENT HARD DRIVES IN IT AND IT WILL STORE INFORMATION ACROSS ALL THOSE HARD DRIVES.

SO IT'S NOT NECESSARILY MULTIPLE COPIES BECAUSE THERE'S USUALLY JUST ONE COPY.

INSTEAD OF IT BEING STRUCTURED ON ONE HARD DRIVE IT'S SPLIT UP ACROSS MANY AND SO NETWORK ATTACHED STORAGE, WE SORT OF TALKED ABOUT THAT WITH NEAR-LINE AND ONLINE.

THAT MEANS YOUR STORAGE SYSTEM IS TIED INTO A NETWORK SO YOU CAN ACCESS IT FROM YOUR COMPUTER.

ISAN, STORAGE AREA NETWORK IS A NETWORK OF NETWORKS SO YOU WOULD HAVE MULTIPLE DEVICES CONNECTED TO A NETWORK AND THEN A SAN WOULD BE THAT NETWORK.

THEN CMS IS CONTENT MANAGEMENT SYSTEM AND DAMS IS DIGITAL ASSET MANAGEMENT SYSTEM.

IT'S HARD FOR US TO GIVE YOU RECOMMENDATIONS ON WHAT TO DO OR WHAT TO USE OR SPECIFIC APPLICATIONS BUT CERTAINLY KNOWING THE TERMINOLOGY HELPS YOU MAKE THE DECISIONS THAT ARE REQUIRED FOR THE TECHNOLOGY.

WE SO WE'LL PROBABLY GET A GOOD NUMBER OF QUESTIONS ABOUT THAT AND WE'LL TRY TO ADDRESS THEM. WHEN YOU'RE THINKING ABOUT DIGITAL PRESERVATION SYSTEMS FOR YOUR STORAGE AND BACKUP THERE ARE A COUPLE DIFFERENT TYPES. YOU CAN WORK WITH A VENDOR. YOU CAN GET A TURNKEY SOLUTION IS OFTEN THE TERM PHRASE FOR A PIECE OF SOFTWARE THAT IS INSTALLED AND DOES EVERYTHING. IT WILL BOTH HAVE SORT OF A FRONT-END INTERFACE THAT YOU CAN USE AS WELL AS MANAGE YOUR STORAGE SYSTEM AND YOUR LAMP STACK AND OBVIOUSLY A VENDOR WOULD MOST LIKELY SET THAT UP UNLESS YOU HAVE THE I.T. DEPARTMENT TO HELP YOU DO THAT. THEN OPEN SOURCE SOFTWARE I JUST WANT TO SORT OF MENTION BECAUSE IT GETS A LOT OF PRESS IN OUR COMMUNITY AND IT IS VERY VALUABLE AND VERY COMMUNITY DRIVEN BUT IT'S ALSO WORTH REMEMBERING THAT IT DOES REQUIRE ITS OWN SORT OF MAINTENANCE AND INVOLVEMENT AS FAR AS USE SO IT IS TECHNICALLY FREE TO DOWNLOAD AND INSTALL BUT AS FAR AS SUPPORT AND MAINTENANCE IT DOES REQUIRE THINGS AS FAR AS STAFF TIME OR EXPERTISE. THAT'S IMPORTANT TO REMEMBER. SO IF YOU'RE SORT OF THINKING ABOUT BACKING UP STORAGE SYSTEMS YOU NEED TO THINK ABOUT YOUR RESOURCES AND EXPERTISE THAT ARE IN HOUSE. WHAT YOU CAN PAY FOR THEM IF YOU PLAN ON WORKING WITH THE VENDOR, REQUIREMENTS AND NEEDS, WHAT TYPE OF MATERIAL ARE YOU WORKING WITH, WHAT ARE YOUR REQUIREMENTS AS FAR AS NEAR-LINE OR OFFLINE STORAGE.

HOW IT TIES INTO WHAT SYSTEMS YOU'RE CURRENTLY USING, EVEN SOMETHING AS SIMPLE AS MICROSOFT OFFICE.

THAT PLAYS NICE WITH CERTAIN PIECES OF TECHNOLOGY AND NOT SO NICE WITH OTHERS.

AND THEN ALSO, YOU KNOW, FOR THOSE OF US IN CULTURAL HERITAGE WE OBVIOUSLY ARE PUTTING A LOT OF IMPORTANT ASSETS AND METADATA AND SPENDING TIME CREATING METADATA AND WE'RE PUTTING THAT INTO THESE SYSTEMS AND SO IT'S ALWAYS IMPORTANT TO KEEP IN MIND THAT SOME TIME IN THE FUTURE BECAUSE TECHNOLOGY ALWAYS CHANGES YOU'LL BE NEEDING TO GET YOUR DATA OUT OF THAT SYSTEM.

SO DATA IN AND DATA OUT IS SOMETHING TO REMEMBER WHATEVER TALKING TO VENDORS OR I.T. ABOUT BACKUP SYSTEMS AND STORAGE AND NO SOLUTION IS PERMANENT SO OUR WHOLE SERIES OF.

SO THESE ARE A COUPLE AND I JUST PUT UP THE LOGOS FOR PEOPLE TO RESEARCH AND THEY'RE GENERALLY IN THE RESOURCES AND ARE PRETTY EASILY FINDABLE ONLINE.

ARCHIVEMATIKA IS A FRONT LINE SYSTEM THAT DOESN'T MANAGE STORAGE BUT IT DOES HELP YOU -- IT HAS A FIXITY GENERATION TOOL IN IT AND A NUMBER OF OTHER TOOLS TO ADDRESS, SOME OF THE CHALLENGES WE TALKED ABOUT IN THE EARLIER SLIDES.

ARCHIVE SPACE IS WHAT ARCHIVES TOOL KIT IS GOING TO BE TURNING INTO SO IT'S A NEW ARCHIVES COLLECTION MANAGEMENT SYSTEM.

ISLANDORA IS A PRESERVATION SYSTEM.

DURACLOUD IS AN ONLINE STORAGE SYSTEM GEARED TOWARDS CULTURAL HERITAGE AND NONPROFITS AND THEN

META-ARCHIVES AND LOCKSS, LOCKSS STANDS FOR LOTS OF COPIES KEEPS STUFF SAFE AND THESE ARE CHRA B AREIVE THE INITIATIVES AND WE'RE SURE LIZ AND TOM WILL TALK ABOUT THEM IN THEIR PRESERVATION NEXT WEEK FOR PEOPLE TO SHARE INFRASTRUCTURE RESOURCES AS FAR AS PRESERVING DIGITAL OBJECTS SO YOU WOULD SAY HAVE SOME SORT OF COMPUTER HARDWARE ON YOUR SITE AND YOU WOULD BE HOSTING INFORMATION FS FROM OTHER INSTITUTIONS IN YOUR CONSORTIUM AND THEY WOULD BE HOSTING YOURS SO THERE'S LOTS OF COPIES WOULD BE SPREAD AROUND THE NETWORK THAT THE CONSORTIUM HAS AND THERE'S THE FIXITY CHECKING ELEMENT INVOLVED.

THEN FEDORA IS JUST A DIGITAL RESOURCE.

SO JUST SOME NAMES TO THINK ABOUT.

AND THE KEY CONCEPTS JUST TO SUM UP THIS WHOLE PART OF THE PEREZ STATION, MULTIPLE COPIES; MULTIPLE PLACES, SEEMS LIKE PEOPLE ARE DOING A PRETTY GOOD JOB ABOUT THAT.

MULTIPLE MEDIA TYPES AS WELL. UNIQUE UNIVERSAL IDENTIFIERS, WE TALKED ABOUT THAT EARLIER.

JUST WHEN A DIGITAL OBJECT HAS AN IDENTIFIER IT NEEDS TO STAY CONSISTENT THROUGHOUT TIME REGARDLESS OF WHERE IT'S LOCATED AND THAT REALLY -- IT JUST MAKES EVERYTHING EASIER AND YOU WOULD BE SURPRISED HOW MANY PLACES WILL CHANGE THE NAME OF A FILE OR A DIGITAL OBJECT WHEN THEY PUT IT ON A DIFFERENT PIECE OF MEDIA OR WHEN IT SOMEHOW GOES TO A VENDOR SOMETHING LIKE THAT.

SO UUID IS SORT OF THE ACRONYM FOR THAT.

INVENTORY AND IDENTIFY, WE
TALKED ABOUT THAT AT THE
BEGINNING.

KNOWING WHAT IT IS IN YOUR
DIGITAL ENVIRONMENT BUT ALSO IN
YOUR PHYSICAL BUILDING AS WELL
AS IF YOU'RE KEEPING THINGS IN
MULTIPLE PLACES KNOW WHERE THEY
ARE.

RECORD AND MONITOR FIXITY
INFORMATION.

SO HOPEFULLY SOME OF THE HOME
WORK WILL HELP PEOPLE UNDERSTAND
THAT BETTER AND THERE'S YOU
HAVE?

THE RESOURCES THAT PEOPLE CAN
REFER TO TO SORT OF TRY TO WRAP
THEIR HEAD AROUND PUTTING THAT
PIECE INTO THEIR WORK FLOWS.
WORK WITH I.T., I PROBABLY DON'T
NEED TO TELL ANYBODY THAT.
BE ADAPTIVE.

WE TALKED ABOUT THE SERIES OF
NOWS AND THAT SOFTWARE WILL
CHANGE, FORMATS WILL CHANGE,
MEDIA WILL CHANGE.

SO, YOU KNOW, DIGITAL
PRESERVATION IS ALWAYS TRYING TO
REMAIN FLEXIBLE AND THAT TIES IN
WITH THE DATA IN; DATA OUT
PRINCIPLE AND THAT LINKS INTO
OUR LAST BULL IT WILL WHICH IS
SYSTEMS CHANGES BUT DATA
SHOULDN'T.

SO THAT'S THE KEY PART OF
DIGITAL PRESERVATION.

SO THAT'S BACKUP AND STORAGE AND
IT LOOKS LIKE WE HAVE SOME
QUESTIONS SO DO YOU WANT TO TALK
ABOUT THEM A BIT, DANIELLE, AND
I'LL TRY ANSWER?

>> SURE.

BOY DO WE HAVE QUESTIONS!

(LAUGHS)

JUST A QUICK FIRST NOTE, WE
POSTED THE LINK TO THE HANDOUTS
AND UNFORTUNATELY WHEN ADOBE

CONNECT TOOK HIS SLIDES IT DID MAKE THE BACKGROUND VERY, VERY PROMINENT SO IF YOU'RE HAVING A HARD TIME READING BE SURE TO LOOK AT THE ACTUAL HANDOUTS YOU CAN DOWNLOAD AS A P.D.F.

THEY SEEM TO BE PRETTY READABLE.

>> I DON'T KNOW WHY IT'S THAT GRAY BUT THAT'S ALL RIGHT.

>> ADOBE CONNECT DOES FUNNY THINGS SOMETIMES.

SO WE'VE HAD A LOT OF QUESTIONS AND I DON'T KNOW THAT WE'LL BE ABLE TO GET TO ALL OF THEM BUT I'M GOING TO TRY TO DIGEST.

THE FIRST ONE WAS CAN YOU EXPLAIN RAID A LITTLE BIT MORE AND THEIR DIFFERENT RATE OPTIONS AND NICOLE ASK CAN FILES BECOME CORRUPT ON A RAIDS SYSTEM?

HOW DOES THAT WORK

>> WELL, AN EASY WAY TO THINK OF IT IS LIKE A NETWORK OF HARD DRIVES, PERHAPS, SO -- BUT THEY'RE ALL WITHIN ONE SYSTEM.

SO IT'S A RANDOM ARRAY OF INDEPENDENT DISK IS WHAT THE ACRONYM STANDS FOR AND THAT JUST MEANS THAT THERE IS A PIECE OF SOFTWARE WRITING INFORMATION TO ALL THOSE DISKS BUT IT'S NOT WRITING MULTIPLE COPIES OF THE INFORMATION.

SO, YOU KNOW, WHAT THAT SAVES YOU IS THAT A DISK FAILURE IN A RAID ARRAY WILL NOT AFFECT ALL OF THE INFORMATION STORED IN THERE, IT WILL ONLY AFFECT THE ONE DISK THAT FAILS.

AND SO, YOU KNOW, HARDWARE, THE EXTERNAL DRIVES AND INTERNAL DRIVES OFTEN FAIL VERY EASILY BECAUSE OF THE WAY THEY OPERATE IS VERY DEPENDENT ON BEING HERMETICALLY SEALED AND, YOU KNOW, THERE ARE JUST A TINY -- IT'S SPINNING VERY QUICKLY AND A

HARD DRIVE CRASH IS A TERM YOU OFTEN HEAR SO A RAID ARRAY SORT OF ADDRESSES THAT BY ISOLATING THE CATASTROPHIC DAMAGE THAT MIGHT OCCUR WITH A HEAD CRASH.

>>

IF THAT SORT OF HELPS. SO IT'S JUST MULTIPLE HARD DRIVES SO IT'S NOT WRITING INFORMATION MULTIPLE TIMES, IT'S WRITING ACROSS THEM INSTEAD OF JUST ON ONE DRIVE.

>> OKAY.

AND THEN THE OTHER QUESTION WE GOT A LOT OF VARIANCE OF CONCERNS CLOUD STORAGE AND PEOPLE WONDERING WHAT EXACTLY IS MEANT BY CLOUD STORAGE AND I GAVE THEM MY FLIP RESPONSE WHICH IS THAT IT'S A HARD DRIVE YOU DON'T CONTROL.

BUT, YOU KNOW, SOME OF THE CLOUD STORAGE OPTIONS AND PEOPLE WONDERING ABOUT SAFETY AND PERHAPS SOME OF THE TOOLS THAT YOU MIGHT BE ABLE TO ACCESS.

>> YEAH, I WILL -- YOU KNOW, CLOUD STORAGE IS CERTAINLY AN EMERGING THING SO WE'RE FAMILIAR WITH IT AS FAR AS ALL OF OUR GMAIL AND SOCIAL MEDIA THAT WE USE BUT AS FAR AS A PRESERVATION OPTION I GUESS THE FIRST THING TO SAY IS THAT IT SHOULD NEVER BE THE ONLY PLACE YOU'RE STORING INFORMATION.

PARTIALLY FOR THE REASONS DANIELLE MENTIONED.

I MEAN, YOU ARE DEPENDENT ON A CORPORATION, YOU'RE DEPENDENT ON HARD DRIVES AND SERVER SYSTEMS THAT YOU HAVE NO CONTROL OVER AND YOU DON'T KNOW WHERE THEY ARE AND A LOT OF PEOPLE PROBABLY HEAR ABOUT AMAZON WEB SERVICES, AWS, GOING OUT A LOT ON THE EAST COAST AT LEAST AND THEN YOU

CAN'T ACCESS GMAIL OR WHATEVER
CLOUD-BASED DEVICE YOU MIGHT BE
USING.

SO -- AND THERE ARE OPTIONS FOR
IT AS A BACKUP AND DURACLOUD IS
A GREAT ONE AND AMAZON CAME OUT
WITH AMAZON GLACIER RELATIVELY
RECENTLY AND THIS IS BASICALLY
OFFLINE ACCESS SO THEY GIVE YOU
VERY CHEAP STORAGE BUT THEY --
VERY HIGHLY LIMIT YOUR ACCESS TO
IT AS FAR AS THE AMOUNT THAT YOU
CAN PULL OUT OF IT ONCE YOU PUT
IT IN THERE.

SO IT'S INTENDED FOR SORT OF
LONG-TERM STORAGE.

SO, YOU KNOW, THE CLOUD IS A --
AN EMERGING AND CHEAP WAY TO
STORE THINGS ON LINE BUT IT
SHOULD NEVER BE THE ENTIRETY OF
YOUR PRESERVATION SOLUTIONS.

SO YOU SHOULD ALWAYS HAVE
MULTIPLE COPIES BUT I THINK AS A
PLACE TO STORE ARCHIVAL DARK
ARCHIVES, OFFLINE STORAGE IT HAS
A LOT OF POTENTIAL.

>> JUST A COUPLE MORE THINGS
THEN WE'LL MOVE ON ALTHOUGH I
THINK WE'RE DOING QUITE WELL IN
TERMS OF TIME.

BUT YOU KNOW YOU'RE -- YOUR
TRIPLE DEUCES ANALOGY, SARA
ANDREWS WANTED TO KNOW IF THAT
CARRIES OVER TO HIGHER STORAGE
AND MEMORY NEED ITEMS SUCH AS
VIDEO AND AUDIO COLLECTIONS HOW
DO YOU HANDLE THINGS THAT ARE
REALLY LARGE?

FOR TWO TYPES OF MEDIA.

>> I'VE TRIED TO TALK ABOUT
INSTITUTIONAL DEPENDENTSYS WITH
THE RESOURCES AVAILABLE WITHIN
YOUR INSTITUTION AND AV ALWAYS
COMES UP.

I HAD A BIG MEETING ABOUT THIS
YESTERDAY WITH A MUSEUM IN NEW
YORK VIDEOTAPING EVERY PUBLIC

EVENT THAT HAPPENS AND THEY HAVE THIS MASSIVE -- IT'S NOT EXACTLY AN ARCHIVE IT'S BASICALLY JUST STORAGE ON SITE AND THEY'RE NOT DOING THIS MULTIPLE COPIES IN MULTIPLE PLACES BECAUSE IT IS COST PROHIBITIVE AND THEY WANT TO AND THEY'RE LOOKING FOR SOLUTIONS TO ADDRESS THIS. IT'S A PRINCIPLE AND YES IT DOES APPLY TO A.V. AND A.V. CAN BE GIGABYTES IN SIZE FOR SINGLE FILES SO IT IS APPLICABLE BUT I AM CERTAINLY COGNIZANT OF ITS DIFFICULTY TO BE APPLIED TO COLLECTIONS THAT GROW QUITE LARGE IN SIZE.

SO THAT'S WHERE CLOUD STORAGE CAN COME INTO PLAY OR THINGS LIKE LOCKS AND OTHER COLLABORATIVE APPROACHES TO INFRASTRUCTURE.

BUT, YEAH, TWO COPIES IS APPLICABLE TO A.V. AND OTHER LARGE FILES.

>> AND THINGS THAT DON'T FIT ON A D.V.D. ARE CHALLENGING WHEN YOU WANT TO GET DIFFERENT MEDIA TYPES IN THERE.

>> YEAH.

>> IT LOOKS LIKE THEY'VE SWITCHED OUT THE POWERPOINTS TO ONE THAT DOESN'T HAVE A BACKGROUND SO FOR THOSE OF YOU WHO ARE HAVING PROBLEMS I HOPE THIS WILL BE BETTER.

ONE LAST QUESTION AND THIS IS MORE OF A PHILOSOPHICAL QUESTION ABOUT DIGITAL OBJECTS I THINK WHEN YOU MAKE A BACKUP OR YOU -- AND IF YOU DON'T CAPTURE THE FIXITY, PERHAPS, DO YOU LOSE AUTHENTICITY WITH THE DIGITAL OBJECT?

WHAT ARE YOUR THOUGHTS ON THIS?

>> (LAUGHS)

>> IT IS A PHILOSOPHICAL

QUESTION.

>> YEAH.

>> HOW DO YOU PRESERVE THE ORIGINAL WHEN WE'RE MOVING IT AROUND?

>> RIGHT.

WELL, I -- THIS IS A GREAT QUESTION AND ONE THAT COMES UP A LOT AND IT MAKES GREAT ARTICLES AND TALK OVER BEERS THE GREAT THINGS ABOUT DIGITAL OBJECTS THEY SEEM FRAGILE AND DIFFICULT TO MANAGE AT TIMES BUT ONE OF THE BEAUTIES OF DIGITAL INFORMATION IS THAT IT'S EASY TO COPY SO I THINK AUTHENTICITY IN A DIGITAL ENVIRONMENT BECOMES MUCH MORE DEPENDENT ON METADATA AND MANAGEMENT AND THIS IT DOES ON INFORMATION CONTENT.

AND, YOU KNOW, SOME PEOPLE WOULD DISAGREE WITH SOME OF THIS BUT IF YOU CAN MAKE A BIT BY BIT COPY OF A DIGITAL OBJECT THAT HAS THE SAME CHECKSUM, IT'S SUCCESSFUL BY THE SAME SOFTWARE, HAS THE SAME UUID, EVERYTHING ABOUT IT IS THE SAME THEN IT'S JUST AS AUTHENTIC AS THE THING THAT MAY HAVE BEEN GIVEN TO YOU. BUT WHAT HAS CHANGED IS THAT YOU'VE MADE THAT COPY, RIGHT? AND SO I THINK THIS IS SORT OF PREMISE -- THE PREMISE METADATA SCHEMA TRIES TO CAPTURE A LOT OF THIS.

IT'S A DIGITAL PRESERVATION SCHEMA.

I THINK DANIELLE YOU PROBABLY REFERENCED IT A BIT YESTERDAY AND IT'S BUILT TO CAPTURE HOW OBJECTS CHANGE, WHAT CHANGES AROUND THEM EVEN IF THEY THEMSELVES DON'T CHANGE.

SO, YOU KNOW, WE WILL NEVER BE ABLE TO KEEP AD INFINITUM INTO

THE FUTURE A FLOPPY DRIVE, JUST LIKE A PIECE OF PAPER. THINGS DEGRADE OVER TIME AND WILL ROT AND FALL AWAY AND THAT'S TRUE FOR PHYSICAL MEDIA AND DIGITAL INFORMATION IS DEPENDENT ON PHYSICAL MEDIA BUT THE ACTUAL OBJECT AS LONG AS IT'S THE SAME AND IS DUPLICATED IT IS -- IT'S JUST AS AUTHENTIC IN MY BOOK.

SO AUTHENTICITY BECOMES VERY DIFFICULT BECAUSE IT BECOMES MORE ABOUT HOW IT'S MANAGED IN THE INFORMATION ABOUT HOW IT WAS ACQUIRED AND MANAGED AND ITS CONTEXT THROUGH TIME OR SO THAN I THINK IN SORT OF TRADITIONAL PRESERVATION AND ARCHIVES.

>> EXCELLENT AND I WOULD JUST ADD THAT SOMETHING MIGHT BE AUTHENTIC BUT IF YOU AREN'T CAPTURING THAT FIXITY INFORMATION AND CHECKING ON IT YOU JUST DON'T KNOW.

>> RIGHT.

SO A LOT OF THE DIGITAL FORENSICS THINGS THAT -- WHAT WE PUT IN THE RESOURCES AS WELL AS THE ONLINE COMMUNITY, THAT COMES OUT OF THE LEGAL WORLD AND THE CRIMINAL WORLD SO A LOT OF THESE TOOLS COME FROM AUTHENTICITY IN A LEGAL CONTEXT AND THAT'S NOT ONE WE OPERATE WITH MUCH WITH AND NONPROFITS AND CULTURAL HERITAGE BUT MUCH OF THE DIGITAL FORENSIC TOOLS AS WELL AS THE IDEA OF HOW WE INTERACT FORENSICALLY WITH INFORMATION COMES OUT OF THIS AUTHENTICITY CONTEXT DRIVEN BY LEGAL RIGHTS AND THINGS LIKE THAT.

SO GOOD TO REMEMBER.

>>

>> YEAH.

I THINK WE'LL MOVE ON.

YOU COULD ADDRESS REALLY QUICKLY
ACTUALLY SOMEONE POPPED UP WITH
WHAT IS NEAR-LINE STORAGE IF YOU
WANTED TO GIVE A QUICK
DEFINITION OF THAT ONE AGAIN.

>> SURE.

SO WE TALKED ABOUT ONLINE IS
VERY IMMEDIATE ACCESS THAT YOU
CAN GO RIGHT TO IT, PULL
SOMETHING IN, PULL SOMETHING OUT
AND OFFLINE IS A DARK ARCHIVE
WHICH IS NOT MEANT TO BE
ACCESSED MUCH AT ALL IF EVER.
SO NEAR LINE IS FUZZY BUT IT'S
BETWEEN THOSE TWO SO NEAR-LINE
STORAGE WOULD BE MAYBE YOUR I.T.
DEPARTMENT HAS A SERVEER IF
YOU'RE BIG INSTITUTIONS AND YOU
CAN REQUEST OF THEM CONTENT THAT
IS HELD IN THEIR SYSTEMS, IN
THEIR STORAGE SYSTEMS.

AND THEY WILL DELIVER IT TO YOU,
SAY, IN TWO OR THREE DAYS.

SO THIS WILL BE THINGS THAT ARE
NOT, OBVIOUSLY -- DIFFERENT
COPIES MIGHT BE ONLINE FOR
DISPLAY OR ACCESS BUT PERHAPS
THE PRESERVATION COPY ISN'T
ACCESSIBLE BY PEOPLE BECAUSE
IT'S SORT OF ENDANGERED BECAUSE
YOU CAN HAVE LOTS OF USERS GOING
IN AND PEOPLE HAVE ACCESS TO
ONLINE STORAGE.

NEAR-LINE IS SORT OF MANAGED
BECAUSE IT'S NOT DIRECTLY
ACCESSIBLE BY AS MANY USERS.

>> LET'S MOVE ON TO LEVELS OF
PRESERVATION.

WE HAVE HAD A NUMBER OF PEOPLE
WHO ARE FROM SMALL INSTITUTIONS
SAYING HOW DO I DEAL WITH THIS
AND MAYBE THIS PART WILL MAKE IT
CLEAR TO YOU.

>> YEAH.

SO, YOU KNOW, AS I MENTIONED IN
THE BEGINNING OF THE PEREZ
STATION THIS IS SORT OF A

PROJECT A WHICH WHICH -- BY
INDSA, THE NATIONAL STEWARDSHIP
ALLIANCE WHICH CAME OUT OF THE
LIBRARY OF CONGRESS AND SORT OF
-- T BACKGROUND TO THIS IS
BASICALLY TO ANSWER WHAT
DANIELLE WAS TALKING ABOUT WHICH
THESE PRINCIPALS ARE ACCESSIBLE
WITH VERY LITTLE JARGON IN HERE,
IT'S NOT ONE OF THOSE BIG
FRAMEWORK DIAGRAMS THAT I SHOWED
EARLIER THAT IS JUST DIFFICULT
TO LOOK AT AND CERTAINLY -- BUT
IT COVERS KEY CONCEPTS
THROUGHOUT TODAY AND PROCESS
AREAS AND INDSA IS OVER 150
INSTITUTIONS OR MEMBERS.

I CERTAINLY ENCOURAGE PEOPLE TO
JOIN FPBD IF THEY'RE INTERESTED.
IT'S A COLLABORATIVE COMMUNITY
GROUP ON SOME LARGE RESEARCH
UNIVERSITIES TO TINY HISTORICAL
SOCIETIES SO IT'S A VERY DIVERSE
MEMBERSHIP SO THIS WAS DRIVEN BY
PEOPLE AT ALL DIFFERENT SKILL
LEVELS AT ALL DIFFERENT TYPES OF
INSTITUTIONS COMING TOGETHER TO
DEFINE WHAT ARE ACCEPTED BEST
PRACTICES THROUGH COMMUNITY
PARTICIPATION AND THE CREATION
OF THIS AND IT TRIES TO BUILD ON
YOUR BASELINE CONSIDERATIONS AND
CAN BE USED AS A SORT OF
SELF-ASSESSMENT TOOL AND WE'LL
WALK THROUGH SOME OF HIT IN THE
NEXT SLIDE.

AND AS WE SAID ALL ALONG, WHAT
IT DOESN'T COVER IS
INSTITUTIONAL CONTEXT.

IT IS MEANT TO BE ACCESS TO
BELIEVE EVERYONE BUT IT CAN'T
ANSWER TO EVERYONE'S SPECIFIC
UNIQUE NEEDS AND LEVELS AND
EXPERTISE.

IT DOESN'T MAKE SPECIFIC
TECHNOLOGY RECOMMENDATIONS AS
FAR AS PIECES OF SOFTWARE OR

TYPES OF SYSTEMS.

IT DOESN'T TELL YOU HOW TO DO SOMETHING SPECIFICALLY AND POLICY IS SOMETHING THAT EVERYONE WILL HAVE TO DEVELOP ON THEIR OWN.

SO ANOTHER GOAL WAS TO KEEP IT ON ONE SHEET OF PAPER SO THE MOST ACCESSIBLE THING TO ANYONE IS HAVING ONE THING THAT YOU CAN HOLD IN YOUR HANDS SO I'LL TRY TO WALK THROUGH THE BOXES AND TALK ABOUT HOW THEY APPLY TO WHAT WE'VE TALKED ABOUT AND ANSWER ANY QUESTIONS.

SO YOU HAVE LEVEL ONE, LEVEL TWO LEVEL THREE, LEVEL FOUR.

SO THE DOCUMENT IS SCALABLE THEN YOU HAVE THESE SORT OF AREAS OF PRACTICE, STORAGE, GEOGRAPHIC LOCATION, FILE FIXITY, INFORMATION SECURITY WE DIDN'T TOUCH ON TOO MUCH, METADATA AND FILE FORMATS.

SO YOU CAN USE THIS TO SEE WHERE YOU ARE AS FAR AS DOING DIGITAL PRESERVATION BUT ALSO USE IT TO PLAN FOR WHERE YOU WANT TO GET TO AND, OF COURSE, THE LEVELS AND THE AREAS OF PRACTICE ARE INTERDEPENDENT SO YOU COULD BE AT DIFFERENT LEVELS FOR DIFFERENT ACTIVITIES AND WITHIN EACH BOX THERE'S GOING TO BE A COUPLE OF THINGS THAT ARE CONSIDERED COMMUNITY ACCEPTED BEST PRACTICES FOR DOING DIGITAL PRESERVATION IN THIS AREA AND AT THIS LEVEL.

SO WE'LL TALK MOSTLY ABOUT LEVEL ONE AND TWO AND SO TWO COMPLETE COPYS THAT ARE NOT CO-LOCATED, THAT JUST MEANS TWO COPIES IN TWO PLACES WHICH WE HAD A GREAT RESPONSE AND THAT'S GREAT AND FOR DATA ON HETEROGENEOUS MEDIA THAT MEANS BASICALLY WHAT WE'RE

TALKING ABOUT WHEN WE WERE
TALKING ABOUT PHYSICAL MEDIA
EARLIER.

GET THE CONTENT OFF THAT MEDIUM
AND INTO YOUR STORAGE SYSTEM SO
THERE IS AN EXPECTATION THAT
PEOPLE WILL BE MIGRATING OFF OF
THINGS THAT CAN BE -- BECOME
OBSOLESCEMENT OR CAN DEGRADE.
SO FLOPPYS, WE HAD A LOT OF
THINGS ABOUT FLOPPYS BECAUSE
PEOPLE ARE PRESERVING THEM BUT
ALSO CAN BE DANGEROUS BECAUSE
THEY ARE WELL PAST THEIR
LONGEVITY AT THIS POINT.

SO THOSE TWO, FILE FIXITY AND
DATA INTEGRITY WE TALKED ABOUT
CHECK FILE FIXITY AND IF IT
HASN'T BEEN CREATED, CREATE IT.
WE TALKED ABOUT HOW TO AT LEAST
THINK ABOUT THAT AND WE'LL HAVE
YOU DO IN THE YOUR HOME WORK AND
WE PROVIDED TOOLS IN CASE PEOPLE
WANT TO IMPLEMENT IT IN THEIR
OWN INSTITUTIONS.

INFORMATION SECURITY WOULDN'T
TOUCH ON TOO MUCH BUT I TALKED
ABOUT IT A LITTLE IN TALKING
ABOUT ONLINE AND NEAR-LINE
STORAGE SO IDENTIFYING WHO HAS
THE AUTHORITY TO READ, WRITE,
AND MOVE AND DELETE INDIVIDUAL
FILES, DIGITAL OBJECTS, RIGHT?
SO NEAR LINE STORAGE IS TO
ADDRESS THAT, YOU GIVE FEWER
PEOPLE ABILITY TO ACCESS THINGS
AND BY MANAGING ACCESS THAT
HELPS PRESERVE CONTENT PEOPLE
ARE NOT NECESSARILY THERE TO TRY
TO ALTER THINGS INTENTIONALLY
BUT, YOU KNOW, WE ALL OBVIOUSLY
HAVE LOST FILES PROBABLY AND
OVERWRITTEN THINGS WE HAVE DONE,
ACCIDENTALLY DELETED THINGS SO
THERE IS THAT FRAGILITY THAT IT
CAN BE MISMANAGED ACCIDENTALLY
AND INFORMATION SECURITY TRIES

TO ADDRESS SOME OF THAT.
SO KNOWING WHO HAS THOSE
AUTHORIZATIONS AND LISTING IT IS
LEVEL ONE INFORMATION SECURITY.
METADATA IS HAVING AN INVENTORY
WHICH WE TALKED ABOUT AND WHERE
IT'S LOCATED I TALKED ABOUT THAT
THEN ENSURING ITS BACKUP AND NOT
LOCATED OF THE INVENTORY AS WELL
AS THE DATA.

IT'S NOT JUST MULTIPLE COPIES
MULTIPLE PLACES OF THE ACTUAL
DIGITAL OBJECTS, IT'S MULTIPLE
COPIES, MULTIPLE PLACES OF THE
DOCUMENTATION OR METADATA THAT
TOLLS YOU WHAT THEY ARE AND
WHERE THEY ARE.

FILE FORMATS.

I MENTIONED IT'S IMPORTANT FOR
PEOPLE TO KEEP IN MIND TRYING TO
USE AND ENCOURAGE DONORS OR
PEOPLE THAT YOU WORK WITH TO USE
OPEN FORMATS.

SO THEY ARE A SIZABLE THROUGH
TIME.

SO THAT'S LEVEL ONE.

LEVEL TWO, YOU KNOW, THE IDEA OF
THE LEVELS IS THEY GET MORE
RIGOROUS BUT THEY SORT OF TOUCH
ON THE SAME FEATURES SO WE MOVE
UP TO THREE COPIES, ONE IS IN A
FULLY DIFFERENT GEOGRAPHIC
LOCATION CREATING DOCUMENTATION
ABOUT YOUR STORAGE SYSTEM YOU'RE
ACTUALLY CHECKING FIXITY AT
LEVEL TWO.

USING WRITE BLOCKERS WHICH WE
TALKED ABOUT.

VIRUS CHECKING WHICH WE TALKED
ABOUT, INFORMATION SECURITY IS
DOCUMENTING THOSE RESTRICTIONS
IN LEVEL ONE AND THEN METADATA
IS GETS MORE ROBUST AS YOU GO UP
THE LEVELS.

THEN HAVING AN INVENTORY OF FILE
FORMATS AND USE.

SO I WILL STOP THERE BECAUSE I

THINK THOSE TWO LEVELS ARE MORE APPLICABLE TO EVERYONE ON CALL AND WE'LL JUST -- WE GOT, WHAT, 15, 20 MINUTES LEFT?

SO I THINK IT WOULD BE GOOD TO TAKE QUESTIONS ABOUT THE LEVELS OF PRESERVATION.

WE HAVE A POLL IN THE NEXT SLIDE SO HAVING EXPLAINED IT AND GONE THROUGH IT WE'LL DO A QUICK POLL TO SEE WHERE PEOPLE THINK THEY'RE AT.

THEN WE CAN TALK ABOUT SOME OF -- WE'LL GO BACK TO THE PREVIOUS SLIDE AFTER THAT.

TALK ABOUT THE BOXES AND SUGGESTIONS THAT ARE INCLUDED THERE

>> THAT SOUNDS LIKE A GREAT IDEA
>> SO THIS IS KIND OF WHAT WE HAD PREDICTED, I THINK.

>> YEAH, I MEAN, IT LOOKS LIKE A PRETTY -- IT'S A GOOD MIX OF -- AND I THINK IT WILL BE INTERESTING TO GET SOME QUESTIONS FROM PEOPLE NOT YET AT LEVEL ONE CERTAINLY AS TO WHAT THEY SEE AS THE BIGGEST HURDLE TO GETTING THERE.

AS I WAS TALKING TO DANIELLE BEFORE THERE'S ALWAYS BEEN THE TALK ABOUT HAVING A LEVEL 0 THAT'S KIND OF AN ONGOING DISCUSSION BUT IT'S VERY DIFFICULT TO DEFINE WHAT THAT WOULD LOOK LIKE AND TELLING PEOPLE THEY'RE AT LEVEL ZERO IS NOT THE THING TO DO.

>> I'D LIKE TO SAY THINGS LIKE "NOT YET AT LEVEL 1."

>> (LAUGHS)

YEAH.

A LOT OF THE DISKIGS WAS WHAT IS LEVEL 1 AND, YOU KNOW, THOSE ARE BASICALLY THE BEST PRACTICES.

WHAT HAVE THE R THE BARE MINIMUM BEST PRACTICES FOR PRESERVING

DIGITAL INFORMATION?

THERE WAS OBVIOUSLY A LOT OF DEBATE PUT ON LEVEL 1 SO WE CAN CLOSE THE POLL AND I'LL GET BACK TO --

>> I THINK AT THIS POINT WE WERE -- FINAL QUESTION.

SO IF PEOPLE HAVE COMMENTS ABOUT -- OR QUESTIONS SPECIFICALLY ABOUT HOW THEY MIGHT HELP THEIR INSTITUTIONS GET TO LEVEL 1 FEEL FREE TO SHARE THAT BUT WE DO DEFINITELY HAVE A FEW THAT WE HADN'T GOTTEN TO EARLIER.

AND SO I WANTED TO GET BACK TO SOME OF THEM ONE OF THE THINGS THAT TIED OVER A COUPLE SECTIONS OF YOUR PRESERVATION WAS WE TALKED ABOUT PRESERVING PHYSICAL EQUIPMENT BUT NICOLE ASKED THE QUESTION WHAT ABOUT PRESERVING OR THE LONGEVITY AND/OR ACCESS TO THE CHECKSUMS USED TO CREATE FIXITY.

THE ALGORITHMS FOR THOSE.

IS THERE DISCUSSION ABOUT THAT

>> THAT'S A GOOD QUESTION WITH THE PROBLEMS WE MENTIONED ABOUT M.D.-5 SUGGESTS PEOPLE WILL BE USING IT LESS IN THE FUTURE BUT I THINK IT COMES BACK TO OUR SERIES OF NOWS IDEA AS LONG AS THAT ALGORITHM CAN DO WHAT IT NEEDS TO DO AT THAT TIME WHICH IS AUTHENTICATE A CHECKSUM, YOU CAN THEN SIMULTANEOUSLY GENERATE ONE USING A NEW ALGORITHM.

SO AT THE SAME TIME YOU AUTHENTICATE THIS DIGITAL OBJECT IS THE SAME YOU CAN THEN USE A MORE UPDATED FIXITY TOOL THAT YOU HAVE THAT WILL GENERATE MULTIPLE ONES AT ONE TIME SO YOU'RE NEVER NECESSARILY RELIANT ON ONE ALGORITHM.

AND THE TOOLS -- ALL THE TOOLS ARE DOING, IMPLEMENTING THE

ALGORITHM ITSELF SO YOU'LL NEVER BE DEPENDENT ON A PIECE OF SOFTWARE TO DO IT.

>> ANOTHER RING TO REMEMBER IS THIS KIND OF HARKS BACK TO THE PRESENTATION ON DIGITIZATION WHERE FOCUSING ON OPEN FORMATS IS A VERY GOOD THING.

OPEN AND POPULAR SO IF YOU CHOOSE SOME PROPRIETARY ALGORITHM THAT ONLY ONE PROVIDER WILL LET YOU CALCULATE THE NUMBER THEN YOU MIGHT HAVE MORE PROBLEMS IN THE FUTURE.

>> YEAH, M.D.-5 AND SHA-256 WILL PROBABLY ALL YOU'LL NEED.

>> AND THEY'RE BOTH OPEN, WE CAN RECREATE NEW SOFTWARE TO CALCULATE THEM THE FUTURE.

THIS GOES BACK TO A MEDIA QUESTION IF YOU HAVE A MASTER OR PRESERVATION COPY OF YOUR IMAGE, DOCUMENT, WHATEVER, DOES THE MEDIA WEAR OUT OVER TIME? DO YOU HAVE TO WORRY THAT EVENTUALLY YOUR C.D. WILL DIE AND YOU CAN'T COPY THINGS OFF OF IT?

>> YES.

THAT'S A BIG PROBLEM WITH OPTICAL DISK IS THAT ONE SCRATCH MEANS THE ENTIRE DISK IS INACCESSIBLE SO THAT'S VERY DIFFERENT THAN THE ANALOG WHICH IS WHAT'S EASY TO CHEAPLY PRODUCE.

THIS IS ONE REASON THEY ARE SO CHEAP TO PURCHASE.

SO THIS REALLY JUST GETS BACK TO THE MULTIPLE COPIES IDEA.

PHYSICAL MEDIA ALWAYS DEGRADES AND AND LOT OF THEM HAVE SHORT LIFE SPANS.

>> SO IT'S NOT MUCH THAT THEY WARE OUT THE WAY YOU WOULD THINK OF A TAPE WEARING OUT BECAUSE IT'S BEEN PLAYED A LOT BUT IT'D

JUST BEEN DAMAGED OVER TIME.

>> I MEAN, YES, IT DOES WEAR OUT.

ANYTHING PHYSICAL WEARS OUT THIS IS A SAD FACT OF LIFE.

BUT, YEAH, AND ACTUALLY DIGITAL OBJECTS ALSO -- THE MORE THEY ARE ACCESSED THE MORE THEY ARE LIKELY TO GET BITROT SO -- BIT CORRUPTION GENERALLY WILL HAPPEN MORE FREQUENTLY BUT T MORE FREQUENTLY SOMETHING IS ACCESSED.

THIS IS WHY THE WHOLE REASON PEOPLE CREATE ACCESS AND PRESERVATION COPIES BECAUSE IF YOU WERE ACCESSING ONE SPECIFIC THING THE WHOLE TIME IT'S MUCH MORE LIKELY TO DEGRADE.

SO PHYSICAL MEDIA AND A LOT OF THROWN TOGETHER THINGS FROM THE EARLY COMPUTER AREA THAT ACTUALLY DEGRADES OVER TIME AND MAKES THEM MORE DIFFICULT TO READ AND ACCESS.

>> SOMEONE WAS ASKING ABOUT MULTIPLE COPIES AT THE POINT OF DIGITIZATION AND I THINK HE HAD A CLARIFICATION OF THAT SOMEWHERE AND I CAN'T FIND IT SOMEWHERE RIGHT NOW BUT JUST IN TERMS OF WORK FLOW ARE THERE RECOMMENDATIONS FOR HOW TO MAKE MULTIPLE COPIES, MULTIPLE MEDIA ET CETERA AT THE TIME YOU'RE DIGITIZING MATERIALS?

>> YEAH, IT DEPENDS ON HOW YOU'RE DOING DIGITIZATION AND I THINK JAKE PROBABLY TALKED ABOUT THIS A GOOD BIT IN HIS EARLIER PEREZ STATION BUT YOU'RE CREATING PRESERVATION COPIES AT THE POINT OF DIGITIZATION AND GENERATING AN ACCESS COPYRIGHT AFTER THAT OR AT SOME POINT DOWN THE LINE.

SO SCANNING AT PRETTY HIGH

QUALITY, 300 OR 600 DPI OR SOMETHING.

THAT'S USUALLY A PRESERVATION COPY.

SO YOU CAN CREATE ANOTHER PRESERVATION COPYRIGHT THERE, YOU CAN CREATE AN ACCESS COPY AND NEVER USE THE PRESERVATION COPY AGAIN.

I THINK IT MAKES SENSE TO MAKE IT PART OF THE DIGITIZATION WORK FLOW AS FAR AS MAKING ANOTHER COPY AND PUTTING IT SOMEWHERE AND ON PAST DIGITIZATION PROJECTS I'VE WORKED ON, THIS IS WHAT WE DID.

YOU WILL SCAN AT A HIGH RESOLUTION ALL DAY, HAVE A BIG FILE, MAKING A COPY OF THE PRESERVATION MASTER AND WORK WITH THAT BUT THE ORIGINAL ONE WILL GO ON TO THE SERVER, ON THE A SPINNING DISK INTO YOUR STORAGE SYSTEM.

AND THE CLARIFICATION THAT I FOUND AND LOOKS LIKE KRISTEN ALSO FOUND WAS HE WAS TALKING ABOUT A 200-PAGE DOCUMENT THAT YOU'RE SCANNING AND IT FAILS AT PAGE 113.

SO PRESUMABLY THE LAST HOWEVER MANY PAGES ARE NOT IMAGED OR THERE'S SOMETHING THAT CAUSES THE WHOLE FILE TO FAIL.

THIS JUST -- AND IT'S THE SAME FOR -- ESPECIALLY FOR AUDIO AND VIDEO.

THIS IS WHY SOMEONE, USUALLY A HUMAN BEING, NEEDS TO LOOK AT THE FILE OR LISTEN TO THE FILE AFTER YOU DIGITIZE IT TO VERIFY THAT YOUR DIGITIZATION WORKED.

>> YEAH, TEXT FILES ARE COMPLICATED BECAUSE IT'S BASICALLY A WHOLE BUNCH OF IMAGES THAT NEED TO BE STITCHED TOGETHER THROUGH METADATA BUT

THEY ARE -- EVERY SCAN IS A DISCRETE DIGITAL OBJECT SO THE PREVIOUS ONES WOULD NOT BE PROBLEMATIC AND HOPEFULLY YOU CAN JUST REDO WHATEVER HAS FAILED

>> SO HOPEFULLY BEFORE YOU DESIGNATE SOMETHING AS YOUR PRESERVATION COPY YOU'VE VERIFIED THAT THERE IS THE INFORMATION IN IT THAT YOU WANT TO BE PRESERVING.

>> YEAH.

>> SO THIS IS I THINK GOING TO BE OUR LAST QUESTION AND IT'S ANOTHER SORT OF PHILOSOPHICAL QUESTION CHERYL McCLELLAN SAYS IF YOU HAD AN EXTERNAL DRIVE WITH A COUPLE THOUSAND JPEGs CREATED OVER A SEVEN-YEAR PERIOD OBVIOUSLY YOUR COLLECTION IS AT RISK.

EXTERNAL HARD DRIVES GO BAD AFTER A WHILE.

WHAT NEXT?

SO WHAT IS THE NEXT TYPE OF STORAGE MEDIA THAT WE THINK WILL BE THE SALVATION OR AT LEAST NOW OUR NEXT THING THAT WE PUT IT ON?

>> RIGHT.

SPINNING DISK IS WHAT WE'RE GOING TO USE INTO THE FUTURE SO THAT CAN BE THOUGHT OF A COUPLE DIFFERENT WAYS AN EXTERNAL DRIVE IS USUALLY A SPINNING DISK BUT IT'S GOING TO BE A LARGE SERVER SYSTEM WITH A RAID ARRAY.

WHAT NEXT I WOULD SAY AT LEAST A COPY ON ANOTHER HARD DRIVE AS WELL AS A COPY ON ANOTHER TYPE OF MEDIA.

IF IT'S LIKE SPECULATIVE WHAT'S THE NEXT BIG MEDIA THING IT'S GOING TO BE SPINNING DISK INTO THE FUTURE BECAUSE OF THE COSTS AND OBVIOUSLY A HARD DRIVE HAS

GOTTEN MUCH LARGER BUT ALSO FOR I.T. DEPARTMENTS AND SERVER WRACKS SO THE COST HAS GONE DOWN ENOUGH THAT I THINK SPINNING DISKS WILL REMAIN THE STORAGE OF CHOICE BUT IT MIGHT BE CAPE TO BELIEVE TRANSITION FROM EXTERNAL AND INTERNAL DRIVES TO SERVER SYSTEM, RAID ARRAYS AND CLOUD STORAGE.

SO IT'S PROBABLY NOT A WHOLE LOT AND IT WOULD BE EASY AND CHEAP TO STORE ON LINE AS A BACKUP. IN CASES LIKE THAT I THINK THE CLOUD DOES OFFER SOME ADVANTAGES AS FAR AS COST.

AND YOU ARE DEPENDENT ON THOSE THINGS.

THE CORPORATE INFRASTRUCTURE AND THINGS LIKE THAT.

SO YOU WILL HAVE ANOTHER COPY EVEN IF THAT FAILS.

>> OKAY SO WITH THAT THANK YOU AGAIN FOR ALL OF YOUR HELP AND YOUR WONDERFUL PEREZ STATION TODAY.

I WONDER IF YOU COULD JUST SAY A QUICK WORD ABOUT THE HOME WORK WHICH INVOLVES CALCULATING SOME CHECKSUMS.

>> RIGHT, SURE.

SO WE TALKED ABOUT --

ESSENTIALLY WE'RE GOING TO HAVE PEOPLE GRAB A VERY SIMPLE TEXT FILE AND USE AN ONLINE TOOL TO GENERATE A CHECKSUM SO THIS IS BASICALLY JUST UPLOADING A DOCUMENT TO THE SITE AND IT WILL GIVE YOU AN M.D.-5 CHECKSUM AND THEN WHAT YOU'RE GOING TO DO IS BASICALLY MAKE A VERY SMALL ALTERATION TO THE FILE AND REUP LOAD IT AND COMPARE THE TWO CHECKSUMS.

SO IT'S IMPORTANT TO THINK ABOUT WHAT'S GREAT ABOUT FIXITY AND ABOUT THESE ALGORITHMS AND

CHECKSUMS IS THAT THE TINIEST,
TYNE NEST CHANGE, LIKE 1 OR 0
THROUGHOUT MILLIONS OF THEM THAT
MIGHT COMPRISE YOUR FILE, IF ONE
TINY THING CHANGES THEN A
CHECKSUM WILL BE DRAMATICALLY
DIFFERENT SO THEY WON'T LOOK
ANYTHING LIKE AY LIKE SO WE'RE
GOING TO HAVE A SIMPLE EXERCISE
TO GENERATE A CHECKSUM AND THEN
CHANGE A FILE AND THEN GENERATE
ANOTHER CHECKSUM AND MATCH THEM
UP.

BASICALLY WE'RE DOING A FIXITY
AUDIT.

>> AND THERE'S SOME ADDITIONAL
EXERCISES IN A HANDOUT SO IF YOU
WANT TO TRY THIS AT A MORE
ADVANCED LEVEL THERE'S
INSTRUCTIONS ON HOW TO DO THAT
AS WELL.

>> SO WE LISTED SOFTWARE TOOLS
AND YOU CAN DO THE SAME PRACTICE
BUT ON A WHOLE DIRECTORY SO YOU
CAN DO TEN OR 20 FILES AND DO
THE SAME EXERCISE AND WE HAVE A
COUPLE OTHER EXERCISES SO WORK
WITH LEVELS OF DIGITAL
PRESERVATION DOCUMENT AND
THERE'S A WEB SITE IF PEOPLE
WANT TO COMMENT OR GET MORE
FEEDBACK AND OBVIOUSLY MY E-MAIL
HERE WILL BE AVAILABLE AND I'D
LOVE FOR PEOPLE TO GET MORE IN
TOUCH WITH THAT DOCUMENT AS WELL
AS THE PEREZ STATION.

SO TO USE THE LEVEL SORT OF THE
ASSESS WHERE YOU'RE AT AND WHERE
YOU THINK YOU CAN BE AND HOW YOU
PRACTICE THAT AND WE HAVE A
COUPLE OF OTHER EXERCISES JUST
SORT OF BASICALLY ANSWERING
QUESTIONS TO DETERMINE YOUR
INSTITUTIONAL RESOURCES AND
EXPERTISE AND HOW YOU CAN BUILD
THOSE AND WHAT QUESTIONS YOU
NEED TO BE ASKING ABOUT BUILDING

A STORAGE SYSTEM OR MAKING
BACKUPS.

>> WELL, GREAT, I THINK THAT IS
ALL WE HAVE, KRISTEN.
ANYTHING ELSE?

>> THANK YOU, DANIELLE AND
JEFFERSON.

THANKS, EVERYONE, FOR HELPING
EVERYONE IN THE CHAT AND WE WILL
GET ANSWERS TO SOME SPECIFIC
QUESTIONS MAYBE OUT TO YOU
PRIVATELY.

AGAIN, JOIN US ON MONDAY AT 2:00
EASTERN TIME FOR OUR LAST CLASS
IN THIS COURSE.

>> THANKS, EVERYONE.