



TRANSCRIPT

Episode 103
Internet Archive Digital Library
Final Transcript
2010.06.15

(ANIMUSIC.COM THEME SONG PLAYING)

Grycz>> THE MODEST AMBITION OF THE INTERNET ARCHIVE IS "UNIVERSAL ACCESS TO HUMAN KNOWLEDGE". THEY'RE PUTTING EVERYTHING IMAGINABLE ON THE INTERNET AND MAKING IT AVAILABLE TO EVERYONE... FOR FREE. HOW DO YOU DO THAT? CAN IT BE DONE? "THE PRACTICAL IDEALIST OF THE INTERNET ARCHIVE;" COMING UP ON "GREAT LIBRARIES OF THE WORLD."

(MUSIC PLAYING)

Garnett>> HI, I'M TOM GARNETT. I'M FROM THE SMITHSONIAN INSTITUTION LIBRARIES AND I'M WORKING WITH NINE OTHER NATURAL HISTORY LIBRARIES ON THE PROJECT CALLED "THE BIODIVERSITY HERITAGE LIBRARY." WE HAVE A CRISIS WITH EXTINCTION OF SPECIES ON EARTH. AND WE HAVE A CRISIS IN OUR KNOWLEDGE OF LIFE ON EARTH. THE BIODIVERSITY HERITAGE LIBRARY INSTITUTIONS WANT TO DO SOMETHING REAL AND TANGIBLE IN THE NEXT FIVE YEARS TO CRACK THIS KNOWLEDGE-JAM IN IDENTIFYING THE FULL SCOPE OF BIODIVERSITY ON OUR PLANET.

Simpson>> I CAN SPEAK FROM THE PERSPECTIVE OF A LIBRARIAN WORKING IN A UNIVERSITY TO THE INTERNET ARCHIVE AND THE OPEN CONTENT ALLIANCE PROJECT, IN GENERAL. I THINK, OBVIOUSLY, IT'S (AS MANY OTHER PEOPLE, PROBABLY, HAVE SAID HERE TODAY) A TREMENDOUSLY IMPORTANT PROJECT.

Penniman>> I THINK A LIBRARY IS A MECHANISM BY WHICH WE CAN CREATE A GENERATION OF INDEPENDENT PROBLEM SOLVERS. PEOPLE, WHO KNOW HOW TO USE INFORMATION; KNOW HOW TO GET AT IT; KNOW HOW TO USE IT IN ORDER TO RESOLVE ISSUES THAT ARE CRUCIAL TO THEM. SO LIBRARIES PLAY A VARIETY OF ROLES: THEY ARE A BRIDGE TO THE PAST; BUT THEY ARE ALSO A BRIDGE TO THE FUTURE.

Garnett>> MUCH OF WHAT IS KNOWN ABOUT LIFE ON EARTH EXISTS IN PUBLICATIONS IN A FEW INSTITUTIONS. THIS IS DERIVED FROM OVER 250 YEARS OF SCIENTIFIC STUDY OF LIFE. THIS MATERIAL WE WANT TO DIGITIZE AND MAKE FREELY AVAILABLE TO THE WORLD AND FOR RESEARCHERS EVERYWHERE. RIGHT NOW, TO ESTABLISH EVEN THE SIMPLEST SORT OF SCIENTIFIC FACT

ABOUT A BIOLOGICAL ORGANISM, REQUIRES NOT ONLY RESEARCH IN THE FIELD, BUT SLOGGING THROUGH MATERIALS IN MULTIPLE LIBRARIES. IN FACT THERE'S EVEN A TERM FOR IT CALLED THE "TAXONOMIC IMPEDIMENT."

Grycz>> THE "OPEN CONTENT ALLIANCE", FORMED IN 2005, IS A CONSORTIUM OF LIBRARIES, UNIVERSITIES, AND MUSEUMS DEDICATED TO CREATING A PERMANENT DIGITAL ARCHIVE OF THEIR HOLDINGS. THE ALLIANCE WAS MADE POSSIBLE BY THE INTERNET ARCHIVE, WHICH SCANS AND HOSTS THE ALLIANCE'S MATERIAL.

Christenson>> THE UNIVERSITY OF CALIFORNIA HAS ONE OF THE LARGEST LIBRARIES IN THE WORLD. I'M THRILLED TO BE HERE, THRILLED TO BE PART OF THE OPEN CONTENT ALLIANCE.

>> WE'RE EXCITED, FROM MY OWN TEAM ON CREATING INNOVATION, TO MAKE ALL THIS CONTENT, (NOT JUST DIGITIZED BOOKS OUT THERE), BUT HOW WE REALLY CREATE INNOVATION ON TOP OF THAT.

>> IF YOU WERE TOO VOCAL IN THE LIBRARY, (WHICH I WOULD OCCASIONALLY BE) HE WOULD SORTA LEAN OVER, AND HE WOULD SAY, "WOULD YOU LIKE TO LEAVE THE LIBRARY?"

(LAUGHING)

AND IT WASN'T A QUESTION.

(LAUGHING)

>>SO, I WAS TALKING TO BREWSTER ABOUT THE OPEN CONTENT ALLIANCE, THIS IDEA OF CREATING A DIGITAL LIBRARY IS SO FUN FOR US.

>> WHEN YOU CLICK ON IT IT TAKES YOU TO A BOOK, AND THIS IS A BOOK READER.

Grycz>> THE FIRST OCA CONFERENCE WAS ENERGIZED BY THE ENTHUSIASM AND VISION OF INTERNET ARCHIVE FOUNDER, BREWSTER KAHLE'S KEYNOTE ADDRESS.

Kahle>> "UNIVERSAL ACCESS TO ALL HUMAN KNOWLEDGE". IT WILL BE REMEMBERED AS ONE OF THE GREAT THINGS THAT HUMANS HAVE EVER DONE. UP THERE WITH THE GUTENBERG PRESS AND THE MAN-ON-THE-MOON. THANK YOU VERY MUCH.

(CLAPPING)

Kruger>> IT'S ALWAYS BEEN A BIG COMPONENT TO SHARE OUR COLLECTIONS AND NOT JUST, UM, SIT ON THEM FOR OURSELVES AND OUR OWN LITTLE COMMUNITY OF USERS IN CHAMPAGNE, ILLINOIS, (WHICH ISN'T ACTUALLY "A LITTLE COMMUNITY".) WHAT DIGITIZING OUR PUBLIC DOMAIN HOLDINGS AND GETTING THEM ONLINE PROVIDES US WITH THE OPPORTUNITY TO DO, IS SHARE THOSE COLLECTIONS BEYOND THE STATE OF ILLINOIS AND BEYOND THE NUMBER

OF USERS WE CAN LEND OUR CONTENT TO IN ANALOG FORM.
SO, IT REALLY GIVES AN OPPORTUNITY TO SHARE THE HOLDINGS OF THE UNIVERSITY
OF ILLINOIS LIBRARY, WITH THE WORLD.

Kammerer>> I LEARNED ABOUT THESE MEETINGS BECAUSE WE ARE PARTNERING WITH
THE INTERNET ARCHIVE TO DO THE SCANNING OF ABOUT 50,000 STATE
PUBLICATIONS. WE ALSO HAVE AN ARCHIVE-IT SUBSCRIPTION WHEREBY WE'RE
ARCHIVING STATE AGENCY WEB SITES THAT WE, AGAIN, ESTIMATE TO CHANGE ABOUT
EVERY 45 DAYS. FOR THOSE TWO REASONS, WE HAVE THIS CONNECTION TO THE
INTERNET ARCHIVE.

Garnett>> AN ENTOMOLOGIST FROM PERU FELT OBLIGATED TO TRAVEL TO WASHINGTON
D.C. TO USE THE SMITHSONIAN, TO RESEARCH HIS OWN "BIOTA". WHEN HE
DISCOVERED THAT THE WORKS HE WAS LOOKING FOR HAD BEEN DIGITIZED AS PART OF
THE BIODIVERSITY HERITAGE LIBRARY, HE WAS SO FULL OF PRAISE AND THANKS
THAT THIS WAS AVAILABLE, NOT ONLY TO HIM, BUT FOR PEOPLE ALL OVER PERU WHO
COULD NOW LEARN ABOUT THEIR OWN "FLORA AND FAUNA" WITHOUT HAVING TO TRAVEL
SEVERAL THOUSAND MILES.

Norton>> WE DECIDED THAT WE WANTED TO PUT TOGETHER AN OPEN-ACCESS LIBRARY
FOR BIODIVERSITY HERITAGE. AND THAT'S GOING TO HAPPEN. BUT I ALSO AM A
MEMBER OF THE "BOSTON LIBRARY CONSORTIUM" WHICH IS ALL OF THESE PEOPLE,
HERE-- [AND] 19 OTHER LIBRARIES. AND WE BELIEVE IN THE OPEN-CONTENT
ALLIANCE.

Preece>> MAURA AND I MET BREWSTER AT CNI [COALITION FOR NETWORKED
INFORMATION] LAST YEAR AND WE "CAUGHT THE FEVER".

Christenson>> WE'RE MAKING AS MUCH AS POSSIBLE AVAILABLE IN DIGITAL FORM.
WE HAVE, I THINK, OVER 34 MILLION BOOKS IN THE UNIVERSITY OF CALIFORNIA
LIBRARY SO WE'RE JUST NOW SEEING, I THINK, A CRITICAL MASS OF WHERE WE CAN
REALLY START TO PROVIDE SERVICES TO THE SCHOLARS ON TOP OF THOSE BOOKS FOR
THINGS LIKE TEXT MINING; FOR THINGS LIKE SERVICES FOR PEOPLE WHO ARE
VISUALLY IMPAIRED. ALL THESE WONDERFUL SERVICES ARE NOW POSSIBLE AND
BECAUSE WE HAVE THIS CORPUS OF DIGITIZED BOOKS. THE POSSIBILITIES FOR
CONNECTING IDEAS AND CONCEPTS AND HUMAN THOUGHT OVER TIME. ...AND OVER
SPACE TOO; MATERIALS COME FROM ALL OVER THE WORLD.
IT'S REALLY AN EXCITING PROSPECT FOR OUR SCHOLARS.

Marx>> THE BOSTON PUBLIC LIBRARY HAS HAD THE WORDS "FREE TO ALL"
EMBLAZONED OVER ITS DOORWAY SINCE 1894, SO WE HAVE LONG PRACTICED THESE
PRINCIPLES. AND WE SUPPORT THESE PRINCIPLES, STILL.

Christenson>> ONE OF THE MOST WONDERFUL THINGS ABOUT INTERNET ARCHIVE AND
ITS COLLECTIONS IS THAT INDIVIDUALS CAN POST UP CONTENT TO THE INTERNET
ARCHIVE. LIBRARIES ARE CONTRIBUTING. LARGE ORGANIZATIONS ARE CONTRIBUTING.
BUT, ALSO, INDIVIDUALS... TO REALLY BRING THIS ALL TOGETHER,

MY FAVORITE BOOK ON THE INTERNET ARCHIVE COMES FROM MY FIVE-YEAR-OLD SON, WHO POSTED UP HIS BOOK "ALL ABOUT BOOKS." SO IT WARMS MY HEART AS A LIBRARIAN.

Grycz>> SO THIS A BOOK THAT HE WROTE?

Christenson>> ABOUT BOOKS.

Kumar>> WE KINDA MIMIC THE EXPERIENCE OF ACTUALLY PHYSICALLY HOLDING A BOOK IN YOUR HAND. WE LAUNCHED IT ABOUT THREE YEARS AGO, FIRST GENERATION. WE HAVE SINCE REVISED IT TO BE ABLE TO WORK BETTER ON MODERN COMPUTERS WITH BIG SCREENS.

Grycz>> THE INTERNET ARCHIVE HAS PUT OVER ONE MILLION BOOKS ONLINE, WHICH ARE SEARCHABLE AND DOWNLOADABLE TO ANYONE.

Bromley>> LAST SUNDAY, (WHILE I WAS IN AIRPORTS, YOU KNOW, FROM BUFFALO TO CHICAGO TO SAN FRANCISCO) AND OUT OF TOUCH; NOT KNOWING WHAT IT WAS DOING; WONDERING IF IT WAS STILL RUNNING. I FINALLY GOT HERE ON SUNDAY, AND HEADED FROM THE AIRPORT INTO THE ARCHIVE OFFICE. I PULLED UP MY LAPTOP AND RECONNECTED, AND SAW THAT THERE'S MY JOB! AND IT'S STILL RUNNING! AND WE'RE NOW WITHIN ABOUT 20,000 OR 30,000 OF OUR ONE MILLION GOAL. AND DURING THE MEETING, WHILE WE WERE THERE, WE KEPT REFRESHING THE PAGE THAT SHOWS THE NUMBER. AND FINALLY IT FLOPPED OVER A MILLION!

Swartz>> I STARTED OUT BASICALLY STARTING FROM SCRATCH. WE HAD TO DESIGN THE SITE; WHAT IT WOULD LOOK LIKE; WHAT IT WOULD DO; WHAT ALL THE PIECES WERE. AND THEN I HAD TO FIND THESE PEOPLE. THERE'S THE "IMPORT PART", WHICH EDWARD HANDLES. WHICH IS TAKING ALL THESE CRAZY RECORDS WE GET FROM BOOKSTORES AND LIBRARIES AND PUBLISHERS AND NEWSPAPERS-- (AND THEY'RE ALL IN THEIR OWN DIFFERANT FORMAT; THEY'RE ALL WEIRD AND STRANGE AND BROKEN IN SOME WAYS) AND EDWARD GOES THROUGH AND CLEANS THEM ALL UP, MAKES THEM ALL SHINE AND MERGES THEM ALL TOGETHER INTO ONE BIG COLLECTION THAT WE CALL "THE OPEN LIBRARY".

Betts>> WE'VE BEEN GIVEN AN ALGORITHM TO TRY AND DECIDE WHETHER TWO DIFFERENT RECORDS REFER TO THE SAME EDITION OF A BOOK.

Swartz>> COLLECTING THESE BOOKS, FINDING THEM, FILTERING THEM, ALL THAT STUFF HE TAKES CARE OF. AND THEN, ANAND, HE RUNS KIND OF THE WEB FRONT END.

Chitipothu>> I INTERACT ONLY WITH THE DESIGNER.

Swartz>> MOVING ACROSS THINGS, INTERACTING, LOGGING IN, ALL THAT KIND OF STUFF IS KINDA HIS AREA THAT HE TAKES CARE OF.

Bromley>> IT'S JUST A WONDERFUL THING THAT BREWSTER'S BEEN DOING

WITH MAKING INFORMATION OF ALL KINDS AVAILABLE TO PEOPLE, FREELY. ...AND BYPASSING THE WHOLE LOCKDOWN THAT WE'VE SEEN INCREASINGLY OVER THE LAST DECADE OR SO.

Grycz>> THE INTERNET ARCHIVE. IT DOESN'T LOOK LIKE A TRADITIONAL LIBRARY.

Kahle>> WE'D STARTED A COMPANY CALLED "WAIS." ACTUALLY IT WAS A PROJECT AT FIRST, ALL OPEN-SOURCE SO PEOPLE COULD TURN TO THEIR COMPUTERS TO ANSWER QUESTIONS. SOME SAY, "GOSH, THAT'S KIND OF OBVIOUS", BUT IT WASN'T OBVIOUS IN 1988 AND 1989! AND SO WE HAD TO HAVE THE RIGHT INFORMATION, A USER-FLEXIBLE WAY THAT YOU COULD JUST TYPE A QUESTION AND TRY TO GET ANSWERS. AND THAT BECAME THE "WAIS SYSTEM" [WIDE AREA INFORMATION SYSTEM]. IT CAME BEFORE "GOPHER" AND BEFORE THE "WEB". IT WAS THE FIRST INTERNET PUBLISHING SYSTEM. WE WORKED WITH NEWSPAPERS TO PUT THEIR INFORMATION ON THIS NEW THING CALLED "THE INTERNET". AND THE IDEA WAS TO TRY TO MAKE A WORLDWIDE SERVICE OF LIBRARIES, TO BUILD "THE BIG LIBRARY" OUT OF LOTS AND LOTS AND LOTS OF LITTLE PIECES... OF PUBLISHERS, PEOPLE'S PERSONAL INFORMATION, CORPORATE INFORMATION, GROUP'S INFORMATION... ALL INTER-OPERATING ON THIS NEW THING CALLED "THE INTERNET".

Grycz>> YEAH.

Kahle>> AND IT WORKED. THE SYSTEM THAT REALLY ENDED UP TAKING THE WORLD BY STORM WAS "THE WEB". THE LATE 80'S AND EARLY 90'S WAS TRYING TO GET IT SO THIS "INTERNET THING" WOULD "GO OPEN" AND GET PUBLISHING ONLINE SO THAT WE COULD THEN BUILD THE LIBRARY. AND BY '94,'95 THINGS WERE GOING WELL-ENOUGH IN TERMS OF HOW THE INTERNET WAS GROWING; AND HOW PUBLISHERS AND PEOPLE WERE PUTTING THEIR INFORMATION ONLINE; THAT I COULD TURN MY ATTENTION TO "BUILDING THE LIBRARY".

TONIGHT, WE'RE GOING TO TALK ABOUT "THE FUTURE OF BOOKS". THE IDEA IS TO HAVE A DISTRIBUTED SYSTEM OF READERS (ON THEIR DIFFERENT READING DEVICES) TALKING TO DIFFERENT LIBRARIES, DIFFERENT BOOK SELLERS, DIFFERENT PUBLISHERS, ANYWHERE AROUND THE WORLD. IF YOU'VE GOT SOMETHING TO SAY; AND IF SOMEBODY'S WILLING TO PAY FOR IT; THERE'S A SMOOTH AND EASY WAY TO MAKE THAT AVAILABLE. TONIGHT, I'M GOING TO TRY, LIVE, TO DEMONSTRATE THIS SERVICE WORKING: CONNECTING READERS ...TO SEARCH ENGINES ...TO LIBRARIES ...TO PUBLISHERS AND ...BOOK SELLERS.

Grycz>> THE INTERNET ARCHIVE EXISTS TO EXERCISE WHAT IT CALLS "OUR RIGHT TO REMEMBER."

Mohr>> THE SOFTWARE THAT WE CALL "A CRAWLER" OR "A HARVESTER" IS PRETENDING TO BE A PERSON BROWSING THE WEB. EXCEPT, IT'S A VERY COMPULSIVE PERSON WHO VISITS A PAGE, AND THEN VISITS EVERY LINK ON THE PAGE, IN ORDER. SO, THE AIM IS REALLY TO GET SOME SORT OF COMPREHENSIVE COVERAGE OF THE MATERIAL ON A CERTAIN WEB SITE; WITHIN A CERTAIN RANGE OF WEB SITES; OR WITHIN THE WHOLE WEB. WHEN YOU TAKE THAT PROCESS

AND SCALE IT UP SO THAT IT'S HAPPENING AT A RATE MUCH FASTER THAN A PERSON COULD BROWSE (OR ON MANY DIFFERENT MACHINES), THEN YOU ACHIEVE THE LEVEL OF COVERAGE THAT WE'VE MANAGED TO ACHIEVE. THERE ARE A LOT OF CHOICES THAT NEED TO BE MADE IN HOW YOU COLLECT THINGS. THERE ARE PORTIONS OF THE WEB THAT HAVE ALMOST INFINITE AMOUNTS OF CONTENT OF NEGLIGIBLE VALUE (AUTO-GENERATED CONTENT). YOU WANT TO, SOMEHOW, TRAIN THE CRAWLERS TO NOT SPEND SO MUCH TIME COLLECTING THAT DATA. ON THE OTHER HAND, THERE ARE AREAS THAT ARE ESPECIALLY HARD TO COLLECT BECAUSE THEY'RE USING CERTAIN ADVANCED WEB MEDIA, ADVANCED WEB DESIGN OR HIGHLY INTERACTIVE PAGES. SOMETIMES THOSE REQUIRE THE ATTENTION OF A HUMAN OPERATOR TO CONTROL THE CRAWLER AND SAY, "ON THIS SITE, MAKE SURE YOU GET ALL LINKS OF THIS FORM". SO, IT'S A MIXTURE OF ESTABLISHING STANDARD RECIPES THAT GET MOST OF THE DATA; AND, THEN, CATCHING THE EXCEPTIONS. THE WAYS THE WEB CHANGES... EVERY MONTH SOMETHING NEW COMES OUT; SOME NEW BIT OF WEB DESIGN COMES INTO STYLE. ADJUSTING THE CRAWLER SO THAT IT STILL GETS AN ACCEPTABLE SAMPLE OF WHAT'S HAPPENING ON THE WEB.

Grycz>> THE DIGITAL LIBRARY OF TOMORROW, LARGER THAN 80% OF THE PUBLIC LIBRARIES IN THE UNITED STATES.

Negulescu>> THE THING THAT WE HAVE BEEN WORKING ON MOST DILIGENTLY, IN THE PAST TWO TO THREE YEARS, IS THE IDEA THAT WE'VE COME A LONG WAY IN FIGURING OUT HOW TO COLLECT A LOT OF THIS DATA. AND WE HAVE A LOT OF THE SUPPORT STRUCTURE AND COLLABORATION WORKING GLOBALLY ON WAYS THAT WE CAN CONTINUE TO MAKE SURE WE CAN CAPTURE WHAT'S MAKING ITS WAY TO BEING PUBLISHED ON THE WEB. WE'RE JUST IN THE EARLY STAGES OF FIGURING OUT HOW TO MAKE IT ACCESSIBLE IN A COMPELLING AND USEFUL WAY TO A BROAD AUDIENCE; NOT JUST A RESEARCHER THAT IS AN EXPERT SEARCHER WHO KNOWS THEIR WAY INSIDE AND OUT OF ADVANCED SYNTAX; BUT ALSO, THE GENERAL PUBLIC: SOMEBODY WHO'S LOOKING FOR "WOW, THERE WAS THIS TREMENDOUS ITALIAN RECIPE THAT USED TO PUBLISHED ON MY FAVORITE RECIPE SITE. IT'S NO LONGER THERE, SO I'M GOING TO GO IN THE ARCHIVE AND TRY TO LOCATE IT." PROVIDING SERVICES WHERE YOU CAN ACTUALLY SEARCH USING KEYWORDS... NAVIGATE "BY DATE" AND "BY KEYWORD" IN A WAY THAT MIGHT BE VISUAL... AS WELL AS, USING TERMS THAT YOU MIGHT RECOGNIZE; WORDS THAT ARE FAMILIAR; WORDS THAT SPARK SOME MEMORY OF A RESOURCE THAT YOU'RE INTERESTED IN. THOSE ARE THE THINGS WE STRUGGLE WITH, MOST, ON A DAY-TO-DAY BASIS. BECAUSE WE'RE DEALING WITH SUCH A LARGE VOLUME OF DATA. WE HAVE OVER 150 BILLION WEB CAPTURES NOW. 1.5 PETABYTES (COMPRESSED)! SO JUST THE SHEER SCALE AND SCOPE OF THAT EFFORT IS DAUNTING.

Jaquith>> I WOULD SAY THAT IT'S A REALLY INTERESTING PROJECT. IT'S A LITTLE CHAOTIC. IT'S A LITTLE ANARCHIC. BUT IT'S WONDERFUL IN THAT SENSE. YOU'RE NEVER GOING TO GET PIGEONHOLED AND STUCK IN SOME CERTAIN AREA FOR A WHILE BECAUSE WE'RE CONSTANTLY CHANGING. WE'RE ADDING NEW MATERIAL. WE'RE ADDING NEW COLLECTIONS. WE'RE ADDING DIFFERENT MEDIA TYPES. AND THE WEB KEEPS CHANGING. SO WE TRY TO STAY ON TOP OF THE WAY THE WEB'S CHANGING, AS WELL AS THE WAY MATERIALS ARE CHANGING. AND WE TRY TO MERGE THE TWO.

THE BIG REASON I CAME BACK AFTER GOING TO THE "FOR PROFIT" WORLD, FOR A BIT; AND WORKING WITH A FINANCIAL START-UP (AND JUST NOT LIKING IT BECAUSE I FELT I'M CONTRIBUTING TO THE "WALL STREET" "FOR PROFIT" "RICH" PART OF AMERICA) AND I WANTED TO DO WHAT I FELT WAS MORE IMPORTANT AND MORE MEANINGFUL FOR THE REST OF THE WORLD, AND FOR SOCIETY. AND TO WORK ON A DIGITAL ONLINE LIBRARY THAT'S TRYING TO BE CUTTING-EDGE AND NEW AND DIFFERENT! I CAN'T IMAGINE A MORE EXCITING PROJECT.

Kahle>> THE INTERNET ARCHIVE HAS GOT NEWS REPORTS, LECTURES, MUSIC, PERFORMANCES, ALL SORTS OF DIFFERENT KINDS OF CREATIVITY. ...SOME SCHLOCKY; AND SOME VERY ERUDITE. BUT PEOPLE ARE ENJOYING BEING ABLE TO SEE IT AND USE IT. I'M NOT SURE IT'S AS DETRIMENTAL TO "INSTITUTIONAL ROLES AND RESPONSIBILITIES" THAT THERE'S A WIDER SHARING OF THESE MATERIALS. WE'RE FINDING THAT PEOPLE ARE CURIOUS ABOUT THINGS... THE GROWTH OF WIKIPEDIA... WHERE PEOPLE ARE NOT ONLY JUST READING VORACIOUSLY ABOUT VERY SMALL SUBJECTS, THEY'RE CORRECTING AND ADDING AND BUILDING IT THEMSELVES. WE NOW HAVE AN OPPORTUNITY TO TURN EVERYONE INTO A PUBLISHER; AND MAKE EVERYONE INTO THEIR OWN ARCHIVIST OF THEIR OWN MATERIALS, IF WE HAVE SOME OF THE RIGHT PUBLIC INSTITUTIONS AROUND TO PLAY THE RIGHT ROLES. AND THAT'S WHERE THE INTERNET ARCHIVE COMES IN.

Miller>> I THINK THE "INTERNET ARCHIVE DIGITAL LIBRARY" CREATES SEVERAL DIFFERENT TYPES OF COMMUNITIES. THERE'S THE COMMUNITY OF PEOPLE WHO ARE COMMITTED TO COLLECTING, ADDING TO, AND PRESENTING THE LIBRARY. AND THEN THERE'S THE FACELESS, NAMELESS, PEOPLE WHO USE THE LIBRARY. WE HAVE OVER 650,000 VIRTUAL LIBRARY CARDS THAT HAVE BEEN GIVEN OUT FOR THE INTERNET ARCHIVE, WHICH IS PRETTY AMAZING. IT'S PROBABLY MUCH MORE ACTIVITY THAN MOST COMMUNITY LIBRARIES HAVE, OR UNIVERSITY LIBRARIES HAVE. SO, IN ONE WAY, WE HAVE A WORLD LIBRARY. WE HAVE A LOT OF FOOTSTEPS COMING IN. WE MEASURE THOSE FOOTSTEPS BY DOWNLOADS. WE MEASURE THOSE FOOTSTEPS IN ACTIVITY; BY FEEDBACK THAT PEOPLE LEAVE AS THEY USE OUR MATERIALS. WE HAVE A "WEB 2.0" INTERFACE WHERE PEOPLE CAN ADD COMMENTS TO THE MATERIALS ONLINE. THE ONLY COMMENTS I EVER SAW ON A LIBRARY OF CONVENTIONAL MEANS ARE THE FOOTNOTES AND THE MARGINALIA; SOME BY DESIGN, SOME BY HAPPENSTANCE, AND SOME FROM ABUSE. SO I THINK, IN SOME WAYS, WE'RE REDEFINING WHAT "COMMUNITY" AND "LIBRARY" MEANS. I THINK THERE'S ALSO A TREMENDOUS OPPORTUNITY IN AN INTERNET-ARCHIVE-TYPE LIBRARY TO INTERMIX, MINGLE AND SAMPLE KNOWLEDGE THAT TYPICALLY WOULD BE DIFFICULT TO DO IN A REGULAR LIBRARY.

I THINK THAT'S A BIG VICTORY.

Prelinger>> I FIRST HEARD ABOUT THE INTERNET ARCHIVE IN 1996, SOON AFTER IT WAS FOUNDED. I GOT BREWSTER'S ARCHIVE MAILERS. AND I THOUGHT "THIS IS A REALLY, REALLY INTERESTING IDEA." AND I SAVED THEM. AND THEN THEY JUST WENT INTO "E-MAIL LIMBO". BUT SOON AFTER I MOVED TO CALIFORNIA IN 1999, WE WERE IN A PROCESS OF TRYING TO FIND A LONG-TERM HOME FOR OUR MOVING-IMAGE

ARCHIVES. WE DIDN'T WANT IT TO GO INTO THE COMMERCIAL SECTOR (WHERE IT MIGHT GET CUT UP; IT MIGHT GET JUNKED; THERE WAS NO GUARANTEE OF LONGEVITY). WE WANTED IT TO BENEFIT THE PUBLIC. BUT IN ORDER FOR THAT TO HAPPEN, WE NEEDED SOME HELP. A NUMBER OF PEOPLE ADVISED ME TO TALK TO BREWSTER. AND I CALLED BREWSTER AND LEFT A MESSAGE. AND HE CALLED ME BACK, A FEW DAYS LATER, AND SAID, "I DON'T KNOW WHO YOU ARE, BUT WHAT CAN I DO FOR YOU?" AND I SAID, "WELL, YOU KNOW, I HAVE THIS FILM ARCHIVE..." AND HE INTERRUPTED ME IMMEDIATELY AND--ON ABOUT A TWENTY SECOND ACQUAINTANCE--HE SAID, "SO, LAST NIGHT, WE WERE SITTING AROUND AT DINNER WONDERING WHERE WE COULD FIND A FILM ARCHIVE TO PUT ONLINE FOR FREE." "HOW WOULD YOU LIKE TO PUT YOUR FILM ARCHIVE ONLINE FOR FREE?" AND I THOUGHT THIS WAS PRETTY WEIRD. I THOUGHT INFORMATION WAS SUPPOSED TO BE EXPENSIVE. (AT THAT POINT, I WAS NEW IN CALIFORNIA AND HADN'T BEEN INCULCATED IN THE "GOSPEL OF OPEN ACCESS": THAT INFORMATION SHOULD BE OPEN. AND SO I THINK I JUST KIND OF STUTTERED A LITTLE BIT LIKE RALPH CRAMDEN. BUT WE CAME TO KIND OF A MEETING-OF-THE-MINDS. AND SINCE I'M A CONTRARIAN AND IT SEEMED TO ME THAT IT MADE SENSE TO TRY TO SEE IF WE COULD MAKE DIGITAL VIDEO COPIES AVAILABLE ONLINE. AND THAT WAS REALLY THE MOMENT WHEN MY THINKING STARTED TO EVOLVE FROM "LET'S MAINTAIN COLLECTIONS OF SCARCE PHYSICAL OBJECTS" TO "LET'S FIGURE WAYS TO COPY THESE PHYSICAL OBJECTS AND MAKE THEM AVAILABLE TO WIDE GROUPS OF PEOPLE AT THE MINIMUM NECESSARY COSTS". AND SO WE BUILT THIS "ONLINE FILM ARCHIVE" WHICH GREW, VERY QUICKLY, TO A LITTLE OVER 2,000 ITEMS. [IT IS] AVAILABLE FOR PEOPLE TO DOWNLOAD, TO REUSE, TO REMIX, TO PUT INTO THEIR OWN PRODUCTIONS. [THERE IS] NO CHARGE. (CREATIVE COMMONS, PUBLIC DOMAIN, LICENSE).

>> LIVE MUSIC ARCHIVE. THERE WAS AN EXISTING COMMUNITY THAT RECORDED ROCK CONCERTS (AS LONG AS THE BANDS AGREED) AND TRADED THEM WITH EACH OTHER FOR NO MONEY. IT WAS A TRADITION STARTED BY "THE GRATEFUL DEAD". THEY WERE TRADING AROUND CASSETTE TAPES, (ACTUALLY, REEL-TO-REEL CASSETTE TAPES), AND THEN--WHEN IT MOVED DIGITAL--THEY STARTED MOVING AROUND ON THE NET. BUT, IT WAS TOO EXPENSIVE FOR THEM TO HOST. THEY PUT THEM UP FOR A LITTLE WHILE. THEY'D TELL PEOPLE ABOUT IT. THEY'D DOWNLOAD IT UNTIL IT GOT SHUT DOWN. SO WE--AS A LIBRARY--CAME TO THIS EXISTING COMMUNITY AND SAID, "WHY DON'T WE TRY TO HELP YOU? WE'RE IN THE BUSINESS, IF YOU WILL, OF MAKING THINGS AVAILABLE". RECORD LABELS WENT OUT OF BUSINESS IN A HUGE WAY IN THE LATE 90'S AND EARLY 2000'S AND A LOT OF THEM, THE ONES WHO WERE STILL PASSIONATE ABOUT WHAT THEY WERE DOING, WENT ONLINE TO GO AND DISTRIBUTE THIS MATERIAL, ANYWAY, FOR FREE. [IT WAS] CALLED "NET LABELS." AND THERE'S A WHOLE GENRE OF RECORD LABELS THAT GOES INTO FILTERING; AND TALKS TO ARTISTS; AND MAKES RELEASES; AND PUTS THEM OUT, BUT THERE'S NO MONETARY TRANSACTION. AND BECAUSE THE INTERNET ARCHIVE HOSTS THESE MATERIALS AND DISTRIBUTES THEM, FOR FREE, WE WORK WITH THESE GUYS SO THAT THEY DON'T HAVE COSTS FOR GIVING SOMETHING AWAY. ITS VERY IMPORTANT: IT SHOULDN'T, SHOULDN'T COST YOU SOMETHING TO GIVE SOMETHING AWAY.

Prelinger>> ONE OF THE THINGS WE'RE DOING AT OUR LITTLE LIBRARY, HERE, THE PRELINGER LIBRARY, IS TO PUT LISTS OF BOOKS (LITTLE SUBJECT BIBLIOGRAPHIES) ONLINE. SO YOU CAN SAY, "HEY, HERE'RE ALL

OUR BOOKS FROM THE WPA". "HERE'RE ALL OUR BOOKS ABOUT EARLY TELEVISION". BUILDING THESE CREATIVE COLLECTIONS. THE INTERNET ARCHIVE IS FOCUSING A LOT OF EFFORT, NOW, IN TRYING TO TIGHTEN UP ITS COLLECTIONS, SO THAT IT JUST ISN'T A QUERY, IT JUST ISN'T A DATABASE. (DATABASES!... THEY MAY BE ONE OF THE GREAT DUMBING-DOWN OF THE WORLD.)

Grycz>> THE INTERNET ARCHIVE SUPPORTS OTHER CREATIVE ENDEAVORS TO KEEP BOOKS ALIVE. THERE IS PRINT-ON-DEMAND FOR EXAMPLE; BUT NOT ONLY PRINT-ON-DEMAND, BUT "PRINT-ON-DEMAND ON WHEELS!"

Kahle>> WE'VE BEEN FINDING THE LAST MILE OF GETTING INFORMATION TO PEOPLE SOMETIMES IS THE HARDEST, ESPECIALLY WHEN YOU'RE DEALING WITH PEOPLE IN RURAL AREAS OR IN DEVELOPING COUNTRIES. SO WE BUILT A PRINT-ON-DEMAND BOOKMOBILE. [IT'S] LIKE AN OLD BOOKMOBILE EXCEPT IT DOESN'T CARRY PRINTED BOOKS. IT CARRIES COMPUTER EQUIPMENT AND A SATELLITE DISH THAT CONNECTS TO THE INTERNET. WE DOWNLOAD A BOOK; PRINT IT; BIND IT. AND KIDS WALK AWAY WITH A BOOK.

WHAT HAPPENS WITH THE LIBRARIES IS "THEY'RE BURNED." AND THEY TEND TO BE BURNED BY GOVERNMENTS. HISTORICALLY. (THIS IS NOT A POLITICAL STATEMENT. IT'S JUST IF YOU LOOK AT THE HISTORY OF LIBRARIES, THEY TEND TO BE BURNED BY "THE NEW GUYS" THAT COME IN. THEY DON'T LIKE THE OLD STUFF AROUND. THEY'RE SORRY ABOUT IT LATER. BUT IT'S TOO LATE.) SO IF WE'RE GOING TO BUILD "THE LIBRARY OF ALEXANDRIA, VERSION 2" LET'S SEE WHAT WE CAN DO ABOUT THAT ONE BASIC PROBLEM THAT, WELL, MAKES IT SO THAT WE DON'T HAVE "THE LIBRARY OF ALEXANDRIA, VERSION 1", WHICH WAS DESTROYED THROUGH A WHOLE SEQUENCE OF A COUPLE HUNDRED YEARS. WHAT CAN WE DO? WELL, THERE'S, ONE APPROACH. IT'S TO GO AND MAKE A COPY AND BURY IT THE DESERT SOME PLACE. THE EGYPTIANS FOUND THAT YOU COULD ACTUALLY MAKE THINGS LAST, PRETTY WELL, THAT WAY. ANOTHER, IS TO MAKE COPIES. AT ALL TIMES--SINCE THE SUMERIANS--THERE'S BEEN AN UP-AND-RUNNING PART OF THE WORLD THAT--IF WE WERE BETTER AT COPYING WHAT IT IS WE HAVE--AND PUTTING IT IN OTHER PLACES AND OTHER PEOPLE'S HANDS--WE WOULD HAVE THE "OTHER WORKS OF ARISTOTLE". SO WE'VE TRIED TO START CALLING IT "INTERNATIONAL LIBRARY SYSTEM." ORGANIZATIONS THAT WANT TO WORK TOGETHER TO GO AND BUILD COPIES OF EACH OTHER. WHERE WE ARE NOT DEPENDENT ON EACH OTHER, SUCH THAT IF ONE GETS BURNED (OR DESTROYED, OR AN IRON CURTAIN ISOLATES IT FOR 50 OR 100 YEARS) THAT WHEN THOSE THINGS ARE CORRECTED THE OTHERS WILL GO AND RESTORE THE COLLECTIONS THAT WERE THERE.

Miller>> I THINK ONE OF THE WONDERFUL THINGS ABOUT THE INTERNET ARCHIVE IS WE'RE AGNOSTIC ABOUT THE MATERIAL THAT CAN BE PLACED ONLINE. WE ENCOURAGE ALL PEOPLE TO BRING MATERIAL REGARDLESS OF WHAT SHAPE OR CONDITION IT'S IN. IN ESSENCE IT'S THE STATUE OF LIBERTY INSCRIPTION "BRING US YOUR TIRED, YOUR DOWN-TRODDEN." "BRING US YOUR DOG-EARED, YOUR WORN PAGES. WE'LL HOST IT ALL." SOME LOOKS BETTER THAN OTHERS. BUT THE CONTENT, ITSELF, IS THE REAL GOLD. THE CONTENT IS THE REAL JEWEL OF THE COLLECTIONS THAT PEOPLE ARE ENTRUSTING FOR US TO PUT UP AND SHARE WITH OTHERS.

Grycz>> THE "WAYBACK" MACHINE AT THE INTERNET ARCHIVE RESPONDS TO 500 QUERIES A SECOND, GIVING ACCESS TO OVER 150 BILLION WEB PAGES CAPTURED AND STORED SINCE 1996. THE INTERNET ARCHIVE AND SUN MICROCOMPUTERS HAVE FORMED A PARTNERSHIP. THANKS TO ADVANCED TECHNOLOGY, MULTIPLE COPIES OF THE ENTIRE INTERNET FIT IN ONE, EIGHT-FOOT BY EIGHT-FOOT BY TWENTY-FOOT SHIPPING CONTAINER, THAT CAN BE MOVED ANYWHERE. "THE LIBRARY OF ALEXANDRIA ON WHEELS." HERE IS A CONCRETE EXAMPLE OF THE CREATIVE ENTREPRENEURIAL SPIRIT THAT THRIVES AT THE INTERNET ARCHIVE.

(MUSIC PLAYING)

(CLAPPING, CHEERING)

I'M CHET GRYZC, JOIN US NEXT TIME WHEN WE VISIT ANOTHER LIBRARY ON "GREAT LIBRARIES OF THE WORLD".

(ANIMUSIC.COM THEME SONG PLAYING)