Transcript
ERIC Update: 2015 Accomplishments and Future Enhancements

Slide 1

Good afternoon everyone. I am Erin Pollard and I am the Project Officer for ERIC in the U.S. Department of Education. I'd like to welcome you to our ERIC update webinar.

Slide 2

Today we are going to talk about ERIC's accomplishments in 2015, and give a preview of what to expect in the coming year. 2015 was a great year for ERIC! The education community and general public used ERIC heavily. Overall, we had twelve million people from 242 countries, resulting in more than 42 million website page views! Here are some highlights about our collection, our sources, and key products and services.

Slide 3

ERIC is known for its extensive collection of education-related resources. We've been providing access to the literature in education since 1966, and have indexed nearly 1.6 million journal articles, reports and other resources. In 2015, we added more than 48,000 new records to the collection, including over 36,600 peer-reviewed materials. Of those peer-reviewed materials, more than 11,600 of them have full text available for download. This is an increase of 238 percent since last year, and a huge success for ERIC! We'll be talking more about the latest ERIC content a bit later in the presentation.

Slide 4

We'll also talk more about the sources that provide content to ERIC, including our journal and grey literature publishers and individual authors. ERIC currently indexes materials from 1,821 active sources, and added 87 new sources in 2015. These new sources include 59 new journals and 28 new publishers of reports, conference papers, and other documents. In addition, last year IES grantees contributed 80 peer-reviewed full-text reports as a result of the IES Public Access Policy.

Slide 5

Here's a snapshot of the products and services ERIC provides. Overall, in 2015:

- ERIC users downloaded more than 6 million full-text documents from the collection.

- Users made more than 3,800 inquiries to the ERIC Help Desk. If you haven't used the Help Desk before, simply go to the Contact Us area of the ERIC website and either call or send us an email.

- We updated the ERIC Thesaurus last year to add new descriptors and update existing terms. In addition,

- We published two new videos. One video is for all users and the other is specifically for IES grantees and contractors. You can find these on our multimedia page.

- We also gave presentations at 2 major library conferences (the American Library Association's Annual Conference and the annual conference of the Special Libraries Association), and held 2 webinars, including a webinar on the ERIC Thesaurus and another on our proposed changes to the ERIC selection policy. The webinars and presentations can also be found on our multimedia page.

- We also had a major update to the ERIC selection policy that was released at the very end of 2015. Some of you may have attended the Town Hall webinar we held on this topic last year and weighed in on our proposed changes to the policy. The revised policy is now live on the ERIC website and we'll give you some more information about this shortly.

- We're also excited about a new infographic we just released that helps explain how articles are added to ERIC. You'll see this infographic, and hear how research gets added to the collection, a bit later in the presentation.

Slide 6

So to begin – who uses ERIC? While ERIC is designed to be a U.S. based resource, about half of our users are based outside of the U.S. In a typical month we see users from nearly every country in the world. This presents an interesting challenge because we want to design our resources to fit the needs of American users, but also be responsive to global needs.

Slide 7

When we think about our users, we think of four distinct groups with four different sets of needs.

The first audience is academics. These are defined as librarians, students, researchers, and faculty that are based in a college or university. This user group often needs high quality research and has access to subscription databases where they can get full-text articles if ERIC can't provide the full text.

The second audience group is educators. This audience typically has little to no access to subscription databases, but still needs access to high quality research that can inform practice.

The third group is the general public. These are parents and community members that are often looking for more general, easy to understand resources and rarely have access to full-text databases.

Finally, we serve policymakers, such as boards of education or legislators. These individuals need information quickly and are often looking for free full text.

We think of these user groups equally as we form our collection priorities and develop our services.

Slide 8

So that's who uses ERIC, but what is in ERIC? ERIC gets content from journal publishers, grey literature providers, and individual authors. Grey literature providers contribute a variety of reports, conference papers and other materials to ERIC. They are made up of education associations, universities, research centers, and other organizations in the field of education. We also get content from federal and state government agencies, including IES. The smallest category of providers are

individual authors, including IES grantees and contractors, who submit their work to ERIC via our online submission system.

Slide 9

Last year, ERIC selected 59 new journals and 28 grey literature sources from a large pool of candidates for indexing. We gave priority to peer-reviewed materials that are releasable in full text, either immediately or after an embargo. As you've heard we've been prioritizing these resources.

With the addition of these publishers, ERIC now has 1,017 journal publishers and 804 grey literature sources.

If there are new publishers you would like to see added to ERIC, please contact the ERIC Help Desk via the Contact Us link in the footer of the ERIC website. We will add them to our list for future review.

Slide 10

All of the agreements lead to individual pieces of content, or ERIC records. Overall, ERIC added 48,097 new records in 2015, that's slightly more than 4,000 per month. Our content mix is largely made up of journal articles, reports, conference papers, other documents, and books.

As you can see on the chart, the largest group of records, at 52 percent, is the peer-reviewed content where we cannot display the full text. The majority of these records are peer-reviewed journal articles that we don't have permission from the publishers to display in full text. Next, at 24 percent, is peer-reviewed, full-text content. This category includes peer-reviewed journal articles we have permission to provide in full text and IES reports. This has been a key area of the collection that we have been trying to build, so we are really pleased with this number.

The two smallest categories are the documents that aren't peer-reviewed where we cannot display the full text at nearly 16 percent, and the documents that are not peer-reviewed, but we can display full text and not peer reviewed materials (at 8 percent). These are mainly reports from non-profits, conference papers, and the like. Now let's compare this update with a look at the content added to ERIC in 2011.

Slide 11

Here are two pie charts that show a comparison of the mix of work we added to ERIC in 2015 vs the mix for 2011. The percentage of full text content that is not peer reviewed remained fairly constant (at 7 and 8 percent), but there were large shifts in all of the other content areas. We added less of the "no full text and peer reviewed" content in 2015, but note the very large change in the percentage of full-text, peer-reviewed content added in 2015 vs. the amount added in 2011 – more than 260 percent! We have worked to build this part of the collection to serve the needs of our non-academic users – the ones without access to subscription databases, but still need high-quality research.

Speaking of full text as a huge part of the value ERIC brings for the user community, we are very excited to let you know that the College Board has asked ERIC to become the official repository for all of their full-text content. We are now in the process of acquiring and indexing all of the College

Board content that we don't already have available in the collection, and they will be linking to the PDFs in ERIC instead of on their site.

Slide 12

So in 2015 peer-reviewed, full text was a quarter of the new records, but when you look at how this compares to the whole collection, only 2 percent is peer-reviewed, full text. Because ERIC has been in existence for over 50 years and has nearly 1.6 million records, it takes a good bit of effort to shift the collection.

When you look at the breakdown of the rest of the collection, you will see that we are indexing about the same amount of peer-reviewed work where we cannot display the full text, and less non-peer-reviewed work where we cannot display the full text. We have also decreased the amount of non-peer-reviewed full text—we are trying to replace some of that content with peer-reviewed content instead.

Slide 13

So that is our collection as it stands today. However, it raises questions about how content gets into ERIC. We've heard questions from users, like "does ERIC take everything that publishers send them?" and "what happens if I submit my paper to ERIC – will ERIC perform a peer review of my work?" To respond, we created an infographic to help clarify our processes from the beginning to the end of the content work flow.

Now I'm going to turn the presentation over to Dave Brady, who directs the ERIC contract at AEM Corporation. He's going to talk about the processes for how records get added to ERIC.

Slide 14

Thanks, Erin.

I'm going to use the new infographic to give you a behind-the-scenes view of the inputs and processes involved with adding records to the ERIC collection.

We have split the infographic into two pieces to make it easier to see the text associated with each of the processes. This slide shows the top half of the infographic which depicts our source identification and selection processes. The next slide contains the bottom half of the infographic which depicts our content acquisition, record processing, and online collection updating steps. Let's start with the yellow box on the top right of your screen.

This box shows our largest source of potential new content for ERIC - recommendations for journals and other new sources to be added to the collection. These recommendations come from publishers, ERIC users, and members of the ERIC team and we typically get 25 to 50 recommendations each month. Most of these recommendations come from publishers and, in recent years, we've seen an increase in the number of international publishers asking to have their journals indexed in ERIC. So after we receive a recommendation for a new source, then what?

The first step is to conduct a high level review to see if the source is within ERIC's scope. We do this initial review because many of the recommendations that we receive are for sources that are not

education-related. For example, about 50 percent of the suggestions we receive are for journals in other disciplines – such as psychology, medicine, or engineering. Another part of this initial review is determining if the sources are published in English. If a source is education-related and published in English, it moves to the next stage for further consideration.

Now, let's move over to the yellow box on the top left of your screen. These are inputs we receive from individual authors through the online submission system. These materials tend to be conference papers and reports, and also include funded research submitted by IES grantees and contractors. On occasion, we receive journal articles from sources that are not currently indexed in ERIC. These submissions sometimes help us become aware of new journals that we may want to consider as potential sources in the future. These top two yellow boxes represent the key inputs to the next stage in the process which is known as Source Selection.

During the source selection process, we screen materials to determine if they are appropriate for ERIC, and we do this screening using the ERIC selection policy. Broadly speaking, the review against the selection policy ensures that each selected source is relevant to one or more of the topic areas in the IES authorizing legislation, and is education research. There are other criteria that we also consider and, if you would like more information, please see the selection policy on the ERIC website at http://eric.ed.gov/?selection. Online submissions pass through the selection process as soon as they are uploaded, and ERIC notifies authors if their materials are not accepted. Online submissions that are accepted are represented by the yellow box on the left side of the screen and go onto the processing step, which I'll describe on the next slide.

Recommendations for new sources are reviewed twice per year. Because the list of recommendations is lengthy, we aren't able to review all of the suggestions in a single review cycle. We sometimes prioritize our reviews of sources based on our collection priorities at the time.

Publishers that are selected for ERIC sign an agreement to have their content indexed in the system which is represented by the middle yellow box. This agreement specifies if ERIC is allowed to display the full text of the article, either immediately or after an embargo, or whether the document can be accessed on a publisher website.

The final yellow box on the bottom right notes that sources that are not consistent with the selection policy may be held for a future review. Now let's move to the next set of steps in the process.

Slide 15

Now that we have covered the source identification and selection processes, let's move to the bottom half of the infographic which shows our content acquisition, record processing, and online collection updating steps.

The publisher agreement we mentioned on the previous slide also specifies how ERIC will acquire the articles and other content for indexing. This providing of content is depicted in the top yellow box on the slide. Some publishers upload their content to an ERIC FTP site, others email it, and a number of publishers make the content available on their website for ERIC staff to download.

Coming in via the yellow arrow on the top left of your screen are the approved online submissions from the prior slide. These online submissions are directly uploaded from the submission system located on the ERIC website. The brown building in the middle of your screen contains the high-level steps associated with taking the publisher content and online submissions and processing them.

The ERIC team takes all of this content and creates searchable bibliographic records. Processing ERIC records requires skilled hands and many steps. Information specialists add the title, author, and other source information about the material to the record into specified fields following established guidelines. As needed, they also will write an abstract for the article if no abstract is provided by the author or publisher. These abstracts accurately capture the content of the material and approximately 15 percent of each month's content requires creation of an abstract.

Next, indexers on the ERIC team review the subject matter of the content to apply descriptors from the ERIC Thesaurus to the record. These descriptors increase discoverability and allow searchers to find the content that most closely matches their interest. As appropriate, indexers will also add identifiers to augment search results. We will talk more about identifiers later in the webinar.

After all of the metadata elements have been added, our quality control personnel review the ERIC records to ensure that the fields are properly populated and the descriptors and identifiers are relevant to the subject matter. These processing activities produce the 4,000 records each month that are forwarded to the Department of Education for updating in the ERIC collection as depicted at the bottom of your screen.

Now, I will turn it over the Erin to tell you more about changes to the ERIC Selection Policy.

Slide 16

Thanks, Dave. If you would like to download that infographic, it is on our FAQ page.

Now that we have covered the high-level overview of how a record gets into ERIC, we are going to go into a little more detail on a few updates we have made to certain parts of the process.

Earlier this month we posted a major revision to the ERIC selection policy. When we were considering possible changes to the policy, we reached out to ERIC users and our advisory group to gather their input on our proposed changes. We posted a call for feedback on the ERIC website, and held a Town Hall meeting with interested ERIC users. Our advisory group provided their input in a conference call in October. After considering all feedback, we revised the policy. It is now posted on the ERIC website at the selection policy link in the footer.

Slide 17

We've talked about this in previous webinars but, for those of you who were not in attendance, by far the biggest change to our selection policy relates to how we flag peer review on ERIC records.

The new peer-review criterion adds a definition of the review process ERIC will accept in order to apply the peer-review designation to a record. Acceptable review processes include blind peer review and expert peer review. The policy also extends the peer-review flag to qualified grey literature, or non-journal content, like conference papers and reports. This is by far the most significant change of all.

The policy also defines how we apply the status to approved sources and online submissions. For journals, we apply the peer-review designator at the journal level to all records created for this source. For non-journal sources, we may assign the designator to a specific peer-reviewed series or type of publication (like conference papers) where we have proof of a peer-review process. Work submitted by Institute of Education Sciences, or IES, grantees and contractors is all peer reviewed, so the status is automatically assigned. Other authors submitting peer-reviewed work through the online submission system must provide proof of that their work went through an acceptable peer-review process.

Slide 18

Why make this change? Before the revised policy went into effect, ERIC flagged peer-reviewed journal content and also extended the flag to some types of grey literature records, such as IES publications and other funded work, and not other grey literature. This created user confusion, and we wanted to provide the most complete and accurate information possible. We also wanted to apply the flag in a fair and equitable way to all types of qualified content. This meant we had to develop a clearer policy.

Also, when we look for new sources to add to our collection, we give preference to sources that publish peer-reviewed content because we believe that users find great value in these materials. By extending our definition of peer review, we are now able to expand this priority to the highest quality grey literature material our users most want to use.

What will be the impact of these changes? We forecast that we may be able to add the peer-review indicator on 100 approximately new peer-reviewed sources a year. The majority of these will be conference papers and government sponsored reports. However, we expect the new policy has a broader implication. ERIC will continue to build upon its tradition of high-quality grey literature and elevate the status of grey literature in the field. During a review cycle, if we find that a grey literature source has a peer-review process, they will receive the same collection priority as peer-reviewed journals.

If you want to learn more about the background surrounding the change to ERIC's peer-review policy, see the two archived webinars in the ERIC multimedia area: Changes to the ERIC Selection Policy and ERIC Town Hall Meeting.

Slide 19

We made a couple of other changes to the selection policy that are worth mentioning here. For example, we incorporated our working policy related to multi-language journals to add clarity for international publishers. Our policy is that a journal must have at least 80 percent of articles written in English in order to be indexed in ERIC. Additionally, the publisher must provide us with the full text so our indexers can verify the language of the articles.

We also added a preservation policy, which responds to the questions and requests we receive related to removing records or content from ERIC. It is our policy to index records in perpetuity. We do not remove content from the collection except in extreme circumstances, such as claims of plagiarism or social security numbers in a full-text document.

Slide 20

In addition to a major selection policy update, we also made a major Thesaurus update. The Thesaurus is the backbone of ERIC. It's how we index our content. We updated it this past October to capture changes in the terminology, outdated words, and new concepts. We want to make sure that valuable resource remains useful, current, and relevant for ERIC users.

Slide 21

When we update the Thesaurus, we work from a pool of candidate terms that have been suggested by ERIC users and indexers. As we review these terms, we ask:

- Is there literary warrant? Literary warrant means the term appears in ERIC content. Terms with a high level of literary warrant – 1,000 hits or more – are considered to be priorities.

- Is there a need? If we've received multiple requests for the same addition or change, we consider those requests to be a priority. Although we can't select every request for further development, we never get rid of any requested terms. All requests are kept and examined when a Thesaurus update is scheduled.

We also take into account if the change promotes indexing consistency, reflects the language of our users, and brings our Thesaurus into compliance with NISO standards.

Slide 22

Given the care and consideration that goes into an update, it took us nearly a year to develop the final changes. In this update we added 19 new descriptors – here are some examples, including Common Core State Standards and Low Income Students. We also added 28 new synonyms, and made 17 changes to existing terms, including the descriptor change from "mental retardation" to "intellectual disability."

Slide 23

The full list of updated terms is available by browsing the Thesaurus on the ERIC website. To learn more about the Thesaurus, see the video of the webinar in the Multimedia area. The video provides background information and a description of the update process. If you would like to download the updated Thesaurus file, go to the ERIC download area or browse the Thesaurus.

Slide 24

So that is what we accomplished in the past year. As you can tell, we have been really busy! In addition to our ongoing work, we've started some new projects for 2016. These include our project to restore unreadable PDFs, improvements we're making to the ERIC identifier field, and hyperlinks we're adding to provide added value for our users. Dave is going to provide more information on each of these projects.

Slide 25

Thanks, Erin.

We have started a new project to begin restoring PDFs that were taken offline a number of years ago because they were not readable. These are the documents marked as "PDF Pending Restoration" in the collection.

We are converting these documents to make them searchable PDFs, and we have begun returning them online. However, this is not a quick nor straightforward activity given the quality of some of the documents that we are seeking to restore. To understand some of these challenges it is helpful to remember that when ERIC was created 50 years ago, it was primarily a microfiche-based system. When ERIC went online, the microfiche was converted to PDFs to make it easy for users to access documents. However, the technology for PDFs 10+ years ago was different than it is today and the PDFs were made as images. This means that they were not machine readable, which enables searching and is ideal for individuals with disabilities.

Slide 26

To make the documents searchable, ERIC used Optical Character Recognition (called OCR) software to scan the PDFs. This process takes the existing image and then assigns letters, numbers, or punctuation to it to create a layer of readable text that a user can search, copy, and paste. The text created by the OCR process is known as Alternate Text.

The OCR process works really well for a crisp document, like one that you would type in Microsoft Word and then print on a laser printer, but many historic PDFs were not crisp. Let's look at some examples of these less crisp documents.

Slide 27

It is important to remember that a large number of ERIC's historic documents were often typed on a typewriter, converted to microfiche, and then converted to PDF. As many long-time ERIC users may have seen, many of these documents were pretty hard to read with a human eye. Machines (via the OCR process) could not do much better.

Some documents are so faint that they are not readable to the human eye nor via the OCR process, and therefore are not a focus of the current effort to restore unreadable PDFs. Other documents that are in better shape and can be processed may have words misspelled, truncated, or missing letters.

In this example, which is one of the clearer documents being processed, you can see that there are some issues such as stray markings on the image which make it difficult to read, and impact the accuracy of the OCR process. The OCR process created the alternate text shown, but this alternate text is not entirely accurate and therefore requires manual editing to make it readable. Manual editing involves correcting text and punctuation to match the image.

Making these manual edits can be challenging. We don't want to infer what the author is saying and inadvertently change the alternate text associated with the document.

In the example you see here, it's fairly easy to edit some of the alternate text by comparing it with the image. The words "legislation," "distinction," and "between" are clear in the image, so we can accurately edit the alternate text. Where we run into issues is with words that are partially or wholly

obscured, or have missing letters. We could infer a word, but the risk is that it may not be the right word. We are sensitive to the fact that many of these documents are copyrighted materials and that we need to preserve the true meaning of the text.

The benefits of making edits when it's possible to interpret the text is that we obtain a readable layer that will enable users to search the full text. That way users are more likely to find the information they are looking for.

Now let's look at an example which is a bit more difficult to address.

Slide 28

In this example, the document contains text that is very difficult for the OCR process given the stray markings and marks in the margins. In this case you can see that basically the alternate text bears some semblance to the text in the image, but a very significant portion would have to be corrected or retyped in order to properly reflect the text and punctuation in the document. Given the challenges that exist in these records, restoring the PDFs to be readable is a time consuming process.

We are combining the use of manual edits and reviews and automated tools to make the restoring of these previously unreadable documents as efficient and timely as possible. These efforts are part of an incremental effort to be performed over a number of years to restore these unreadable historical documents as we believe that there is value in bringing as many of these full text documents as we can back into the collection.

Now, let's shift from the PDF restoration activities to the ERIC identifiers.

Slide 29

Another project that is underway is some exciting work we are doing with a less known treasure trove of information within ERIC– the identifiers field. In order to talk about this work, let's start with some background information, beginning with defining what identifiers are.

Identifiers are proper nouns that are assigned to ERIC records to provide greater indexing specificity. Identifiers are search aids that help users find information about geographic locations—such as a major cities, states or countries, laws—such as the Individuals with Disabilities Education Act, and measures—such as tests and assessments.

Slide 30

Currently, all of this information is in one field that we made searchable in 2015.

However, this year we are going to break apart this field into three separate fields to make it more searchable. Specifically, having three separate fields will:

- allow more precise searching;

- make it possible to refine search results using identifiers the same way you can with descriptors, sources, and other ERIC fields today; and lastly

- this effort will help add rigor and consistency to the identifiers as a controlled vocabulary.

We think these searchable fields will be a huge benefit for researchers.

Slide 31

Here you can see some examples of how identifiers can be helpful in searches.

Identifiers can be combined with search terms as is shown in the first example. This search would allow you to find work that's been done in Tennessee on reading using the geographic location identifier "Tennessee."

The second search example would return all documents that have "Florida Comprehensive Assessment Test" as a test and measure identifier.

The last example shows how you can combine the use of multiple identifiers to find specific documents. In this search example, you would find records containing the test and measure identifier associated with "Woodcock Johnson" and the geographic location identifier "Indiana."

Slide 32

One key aspect of improving the identifier fields is improving the data associated with them. A key aspect of this effort is to standardize the values associated with identifiers.

We have to keep in mind that ERIC is a 51 year old system. There have been changes over the years in naming conventions and the way identifier terms have been assigned to records. As a result, we now sometimes have a field with different terms for the same entity, making it difficult to find all records. We are taking steps to standardize the language so it's possible to locate everything in ERIC related to a particular term, such as the Higher Education Act, in a single search.

In this example, you see three variations appearing in the ERIC identifier field relating to the Higher Education Act: Higher Education Act, Higher Education Act of 1965, and Higher Education Act 1965. Today, in order to find all records about this Act you would have to search using all three terms. To address these types of situations, we will be standardizing identifier values by finding identifier values that have similar names and then ultimately replacing these terms with standardized terms. We will use software tools to help us develop lists of all identifiers used in ERIC records that either contain or closely resemble similar terms such as those associated with "Higher Education Act." We will then manually review each of these similar terms and provide recommendations to the Department of Education regarding which of these existing identifiers should be mapped to a list of standardized terms. After approval of the mappings of these terms, we will use software to update the appropriate ERIC records to populate the new identifier fields with standardized terms.

The result of these efforts will be a more standardized and less redundant list of identifiers and improved ERIC records.

Slide 33

We are also taking some additional steps to specifically improve the usefulness of the values contained in the Geographic Location identifier. One such action is adding consistency to how U.S. cities are identified.

In the past, multiple standards have been applied to how U.S. cities were tracked in identifiers, ranging from including the city name as a standalone value, not indicating a city name, and using a construct of indicating the name of the state of the city followed in parentheses by the city name itself. Our approach will be to use the format of State name followed by City name in parentheses, as indicated on the slide. We will therefore find all examples of Philadelphia in the identifiers listings of ERIC records, and put them into this new format to provide greater standardization and improved searching.

A second improvement is to eliminate regional references in the identifiers used for states. For example, we currently have records that include identifiers that identify specific regions within states, such as Alabama (Northeast) and Arizona (South). Given that there can be multiple interpretations regarding what comprises northeast Alabama, we will be replacing such identifiers with the identifier for the relevant state. This will provide greater consistency and a broader range of records that searchers can view to obtain the information they are seeking.

A third exciting project has to do with adding hyperlinks to ERIC records, and now I will turn it back to Erin to tell you more.

Slide 34

Thanks, Dave.

As we have worked to fully make ERIC an Institute of Education Sciences investment, we have tried to find ways to add information to records that is useful to our users, keeping in mind the different audiences that I mentioned at the beginning of the presentation. We have identified three new fields that we think will be especially useful. They involve linking to sections of the IES website that are really useful.

Slide 35

The first example is adding a direct link to the IES website to view the publication page for each record. This will have companion reports and supplementary materials. For example, this is a toolkit created by REL Southwest. ERIC just indexes the toolkit, but with the link users will get 31 handouts and 23 videos. Four of these videos summarize recommendations for evidence-based best practices for teaching, and the remaining videos show teachers putting the recommendations into practice in actual classrooms at three different grade levels. This is a fantastic resource that we think ERIC users will appreciate a link to.

Slide 36

The second example of providing a link is to a What Works Clearinghouse practice guide page. Here you will hear from the authors and download different versions of the report. Again, these materials are valuable, and we think giving users the link will be beneficial.

Slide 37

The second type of link goes to a summary of the grant funding. This page gives information about what IES grant funded the study, and includes information like the year and amount of the grant,

the purpose, and an abstract. This is really helpful for policymakers and for the general public to see what their tax dollars are funding.

Slide 38

The last set of hyperlinks we are adding are to a new page that summarizes a What Works Clearinghouse review of a study. The What Works Clearinghouse reviews studies of the research on the different programs, products, practices, and policies in education, and provides ratings on whether the study provides credible evidence based on the design of the study and whether it meets standards that the WWC has established.

 During this process they produce study review guides with information about the population that the study focused on, the design of the study, and if there are any methodological concerns with the study. There is a ton of information in these guides that will be a tremendous value to researchers and students, and it will all be directly linked from ERIC. This will enable novice users to know if the evidence presented in the study is rigorous enough to base decisions off of and will enable more advanced users to have all of the tools needed to get a summary of the impact findings of the study.

 We are really excited about this resource and think it will be a tremendous help to users.

Slide 39

So that's what we've worked on in the past year and are working on going forward. Now we are going to take any questions that you have. If you think of a question after this presentation, or think your question would be best answered offline, please email us at ERICRequests@ed.gov. We will answer your questions that way as well.

Questions and Answers

Please use the chat box to send this to all panelists and we will answer any questions that we see.

Question 1: Where can we find the infographic that was shown at the beginning?

There are two infographics that we have done. The first one you saw, which is what we have accomplished in 2015, is currently on our Notes page. And the one on how a record gets into ERIC is on our FAQs page.

Question 2: How does ERIC define grey literature resources?

We define grey literature as not a journal. So this would be a conference paper, a report (whether it's by a government agency, a think tank or a non-profit), it could be a wide variety of things, but generally just not a journal.

Question 3: What is the current proportion of journal articles to other documents? At one point it was 60/40 – is that still valid?

I would give you the information for 2015 but we don't have this here. I would say that in general that around 60/40 is what we are aiming for. We don't think about it as journals/non-journals, and

that's one of the reasons why we don't have that number off-hand. We think of it more as peer-reviewed versus not peer-reviewed content.

Question 4: Are all articles indexed in the Thesaurus terms by real people – humans?

Yes, they're humans with names and faces. We find that they index better than machines, so we've continued to rely on their expertise and have been very pleased with the fact that they can outperform machines regularly.

Question 5: What is the status of the documents that were removed due to personal information of the author?

In 2013, we temporarily disabled all of our documents due to concerns about Personal Identifiable Information, particularly SS numbers appearing in a large, but not overwhelming, proportion of our documents. While the issue was occasionally about the author, it was not exclusively about the author. Those documents are not available online and there is no plan to return them online. We cannot return them online. If you look at the multimedia section of our website, you can see a webinar that we gave on the PDF restoration process, and some information about what the problems are – why we can't black line it, why we can't put up those documents in other ways, why I can't give you a list of those documents, because we don't want to risk giving out people's personally identifiable Information.

Question 6: Will this webinar be available for review later? Yes. Everyone who registered will be getting a link to this webinar when it is online. We hope to have it online soon, and we will let you know when it is available.

Wrap Up

If you think of questions after this presentation, or if you think your question would be best answered offline, please e-mail us at the ERICRequests@eric.gov link that is on the slide. This is also the same e-mail for the Contact Us desk at the bottom of any ERIC screen and we can answer your questions that way.

Thank you so much for attending this webinar. Please look for the slides and archive of this webinar to be posted in a few weeks at eric.ed.gov/?multimedia.