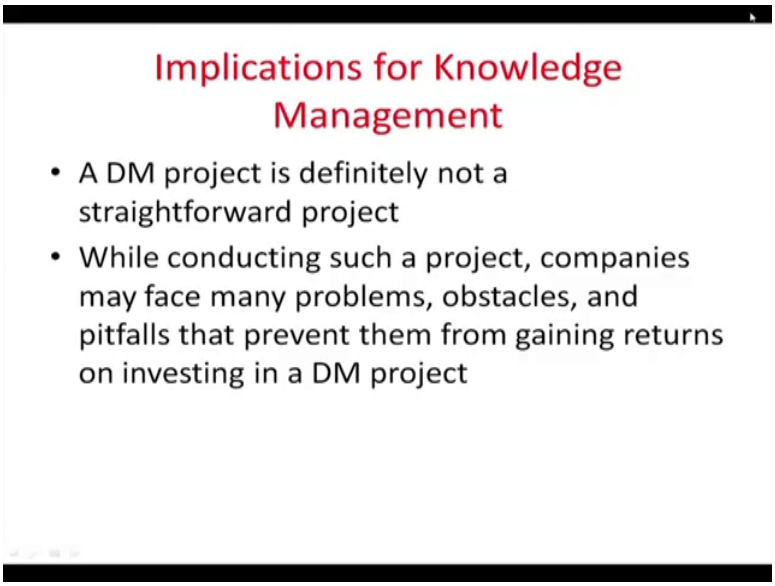
Knowledge Management
Prof. K B L Srivastava
Department of Humanities and Social Science
Indian Institute of Technology - Kharagpur

Lecture – 30
Knowledge Portals

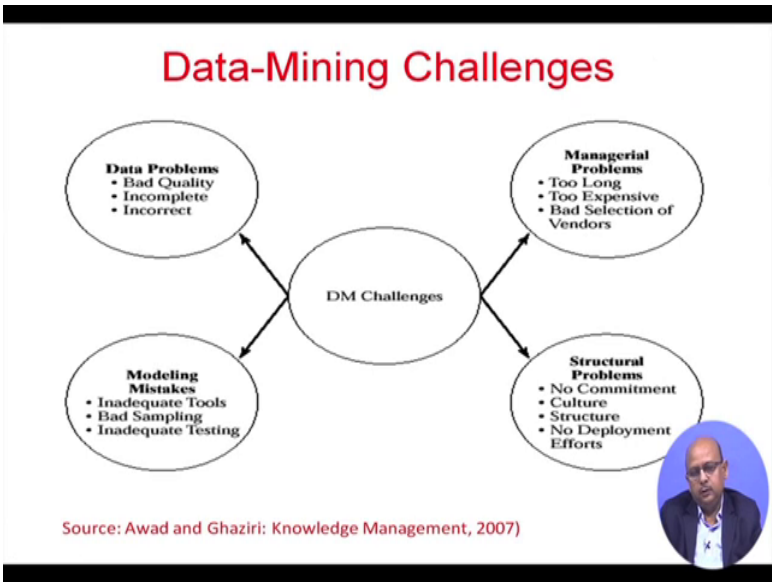
So now we have been discussing about the data mining and its implications for knowledge Management.

(Refer Slide Time: 00:28)



But will also need to ensure that when you are going for data mining what kind of data we are using a data mining right.

(Refer Slide Time: 00:34)



And then you have to see that certain problems or challenges which I associate to data mining are removed because these are acting as a constraint then what will happen your data would not be correct whatever mining efforts is made, through statistical analysis or machine learning and then whatever outcomes are coming out of this is not correct right. So there could be different kind of problems like they could be data related problems right.

If the quality is good, not good or the data is not completed or the missing data of these incorrect data right. So then it will not get good results right. Similarly there are problems related to the management right. If the data is too long too expensive and you are going to select bad vendors and not this data mining activities may be hampered. Similarly how the data is structure structural problem may be there right.

A structural problem is related to basically the people and the culture part right. Whether people are really interested to go for data mining are not right. Whether the commitment and commitment of the analyst knowledge developers, knowledge users for this kind of things or not. What kind of cultured structure you have in the organization okay. And whether you have a norm in the organization to go for these kinds of data mining activities regularly extract information and knowledge from the data or not okay.

What kind of structure you have a structure means the structure of the organization. Whether this can structure across levels facilitate this kind of process or not. And whether there is coordination among the data mining miners and other managers vertically and horizontally to see that how this data mining activities integrated across processes to ensure good information coming out of it and is being used by the manager at cross levels.

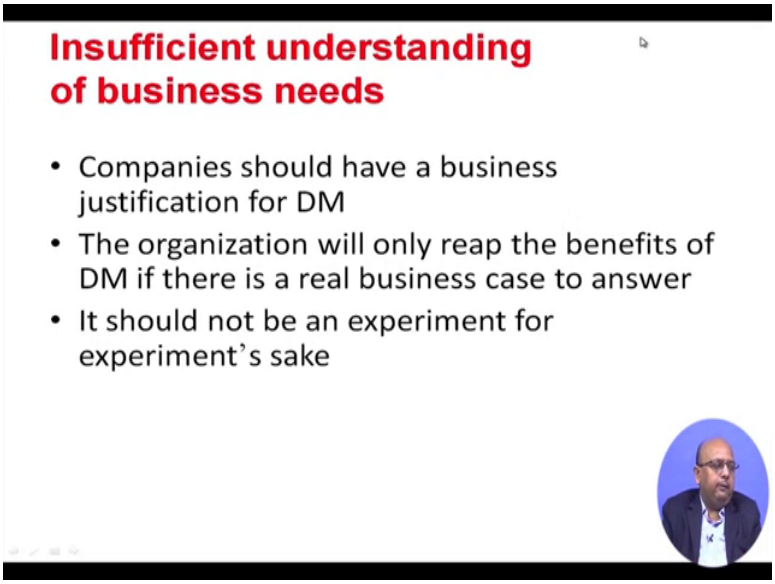
And then the other issues like modeling stakes with the suppose you have developed certain model but the model does not fit into the data then what will you do? It means either your tools are not correct or your assuming are not correct right. Or the kind of sample that you have taken is not correct. So when you are going to use certain sampling techniques make sure that it is more randomized in the process right.

Because sometimes convenient and none data collected through convenient sampling upper bases sampling may not give good results. So it is always advised to go for a more apprentice sampling and so that the data fits into the normal distribution curb and then you get better

results right and then you test it properly. Use appropriate statistical tools to get certain inferences from the data.

And that is why you need to see that these barriers or obstacles to data is not there whether it is related to data itself with the organization, or the management argument when you are going to the analysis of the data especially the mining part okay. Those data cannot be mine unless it is good quality it is perfect data and then we have developed better models, good models right, for testing the data right.

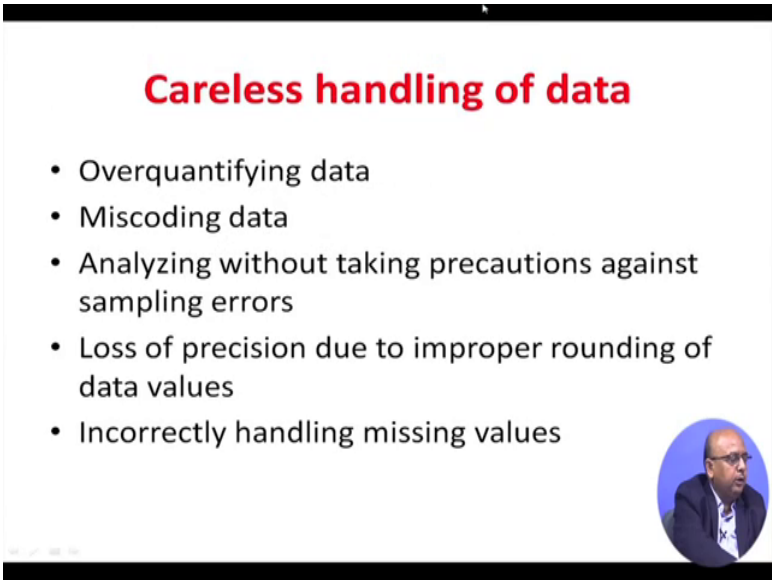
(Refer Slide Time: 03:38)



So it is very, very important that you have to understand the business requirement before you go for it okay. Because unless you have a business requirement or business justification you should not go for data mining activities right. Otherwise you will not be able to get any benefit out of it. So you make sure that what is the objective of that data mining at the first stage.

And whether you have a real business case to answer for this data mining or not, then only you go for it otherwise not right. you do not take an experiment or do not go for experiment to go for data mining and then go for data mining and then because there will be even if there is certain outcomes you will not be able to make use of it and then you try to find out a context where to use it what to do with this information right.

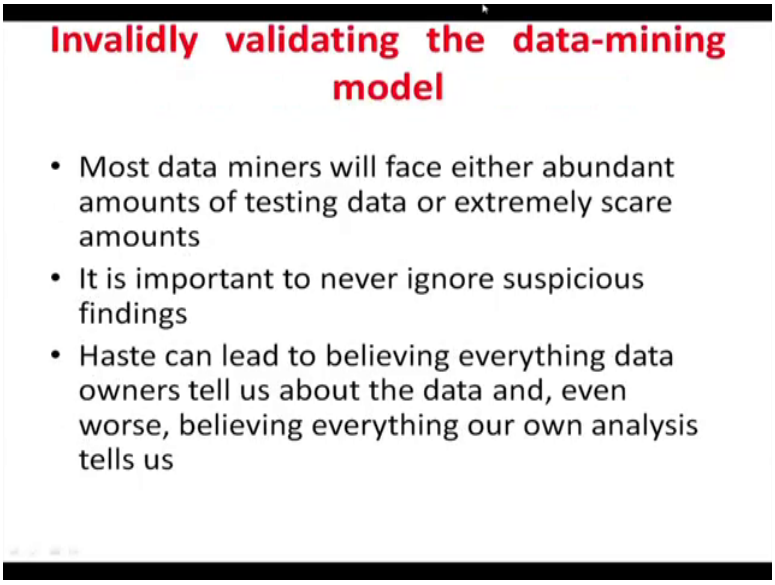
(Refer Slide Time: 04:24)



So first you need to identify and then do not carelessly handle the data right. There is lot of issues like over do not go for over quantifying the data, do not go for miscoding the data. So these are basically things relate to the preparation of the data okay. Identify what are the sampling errors. That is very important okay, and then you identify the kind of precision you are looking forward to okay.

At what level of significance you want to conduct it, and then if there are missing values whether you are going to include and exclude. If you are going to include them what could be the rules for putting values there or if you want to exclude, why you want to exclude this? Because this may reduce your dataset and at the same time this may not give you correct information also right. So these are some issues related data handling which must be looked in to.

(Refer Slide Time: 05:13)

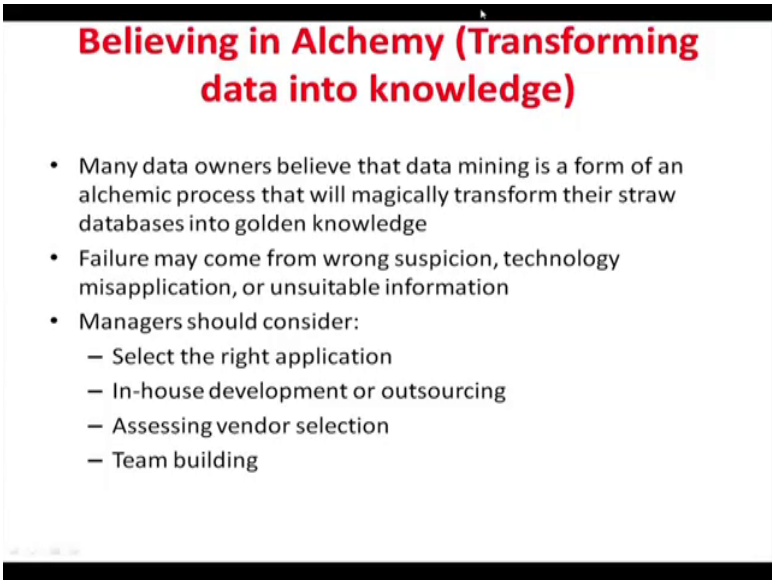


Another important thing is that sometimes you go from validate in certain models and you force your data to fit into the model right because of having a lot of data then you will think that it will okay, it will fit into the data it might fit into the data but sometimes it is not given giving good results. And there are certain results which maybe or may not be correct okay.

So you need to identify that when you are going to validate certain models using certain data mining techniques with whether it is statistical or using artificial intelligence techniques either further classification or prediction or just simply associating certain variables right you need to ensure that model better fits to the data. There are no outliers right, it fits into the normal distribution car right.

Otherwise what will happen even if you have the data and it is not proper data and you are going to validate it? It may not be correct right .so make sure that the data that is there at your hand is correct it is proper, So that you can make better use of this data.

(Refer Slide Time: 06:22)



Then most important thing is that why you want to go for data mining the basic objective is transforming data into knowledge. So you move from data to information right. You will organize them and ultimately you are going to analyze the data to get the information right. So data mining is not an alchemy process see it is not that okay automatically data would be converted into knowledge right.

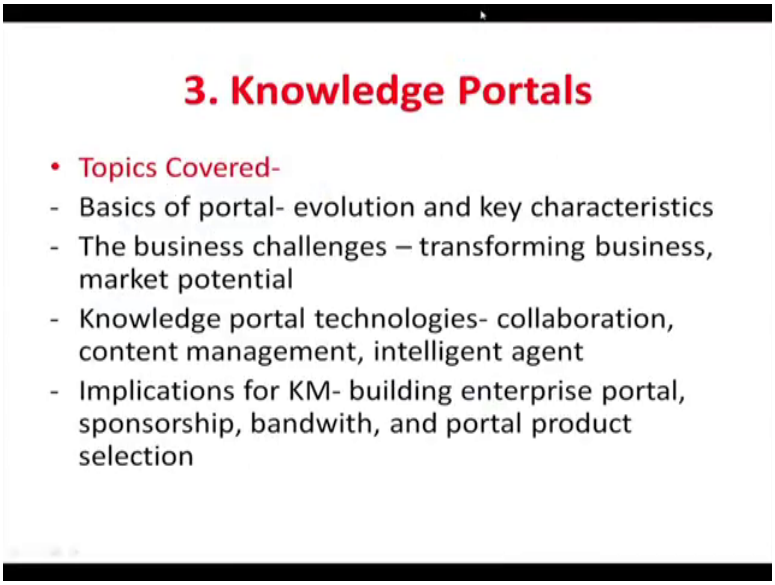
You have to transform them this raw data into some kind of knowledge right, and if you are not able to do it probably it is not going to be beneficial, because for example if you do not have correct data or if you are not going to use appropriate statistical techniques. Or if you are not going to use correct models okay or you are not going to use correct information then probably you will not be able to transform data into knowledge properly.

So, the idea is that the top management or the managers who are responsible for data mining they must consider certain things like make sure that they should be able to identify the business needs and where they are going to apply the information of the knowledge that is coming out of the data mining okay.

And whether this is going to be done inside the organization if you have business analysts then you make use of them it is cost-effective or you can go for outsourcing this kind of thing where you have external business analyst who are going to mine the data depending upon your requirement okay. And if you are going for outsourcing this kind of a services data mining then you identify a suitable vendor.

Who has experience okay, and it may be cost-effective also and then you develop a team who is responsible for taking this kind of activities on a regular basis because you know that is a cycle which goes on in a regular basis. So you need to continuously mine your data to extract useful information and that is where it is very important to see that whether you are really able to transform data into useful knowledge or not and it is a major concern for most of the managers.

(Refer Slide Time: 08:27)

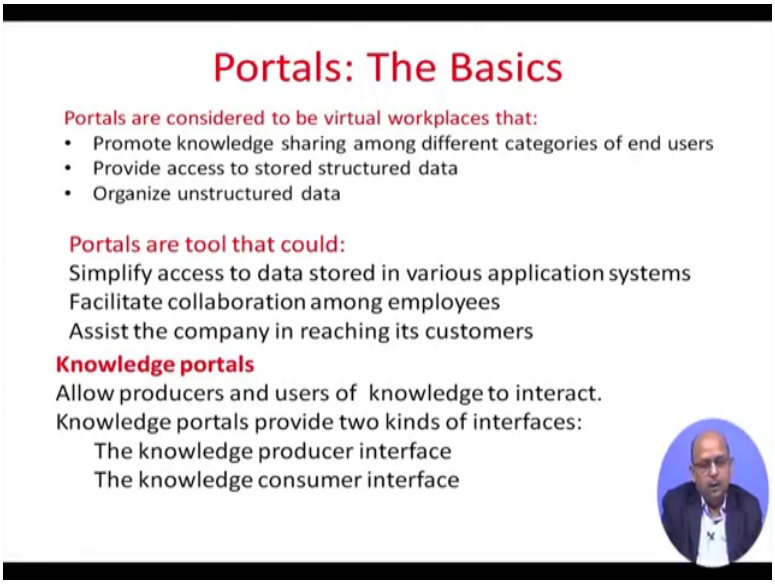


Then once we have talked about data mining activities here. Next we are going to talk about is knowledge portals. And if you look at knowledge portals they also provide a lot of the data okay. So it is also part of the data actually and then very specifically we are going to see that what kind of data is provided by the knowledge portals okay. So we will discuss some of the knowledge portals which have been created by the businesses.

And then we will look into some of the challenges in term of how you are transform business using the knowledge portals what are the market potentials and then we will also look into the applications okay, that how it go upon building enterprises portals okay. How about the sponsoring and then will also look into the certain technical issues or what to do when kind of software and hardware integration is going to be there.

And what things are you going to put on the knowledge portal right. So with this we proceed with knowledge portals okay.

(Refer Slide Time: 09:30)



Knowledge portals are nothing but virtual workplaces right, where you can share information okay, across users okay. Say for example of progressive blog or your website of the organization and if you are going to put a lot of data information order that could be used as a portal also right. And then you have access to structured data right one: kind of knowledge portals could be say for example libraries okay.

Libraries basically supports knowledge sharing across users across categories okay. And it is a storehouse of information you can say, a structure data. Now the make use of the library in a very structured manner okay. And then whatever information you are getting from different sources in library you can organize them okay. So and that is where libraries especially I am not talking about the physical libraries but I am talking about the digital libraries okay.

So digital library could be considered as one of the knowledge portals, organizations website which store data they could also be considered as knowledge portals. Like if go to the website say the Ministry of labor you will find a lot of data related to the labor so which you can extract and make use of it after analyzing. So that also give you lot of information relate to the labor activities would be related to labor acts.

It could be related to labor productivity, it could be related to trading your activities all kind of data you can find out there right. So these are virtual knowledge workplaces is where you can get lot of information right. And then basically you have to see that whether you have access to these places where the data is stored in a digital form right. And what are their applications rights.

Suppose you are doing a research then you can use digital library to identify the kind of research that is being done okay. And then how this research is going to be integrate for your purpose right. It is also used for facilitating collaborations across employees. For example you are knowledge portal where you are going to share knowledge right. So suppose you have a query you can put it in the blog or the portal or knowledge sharing portal.

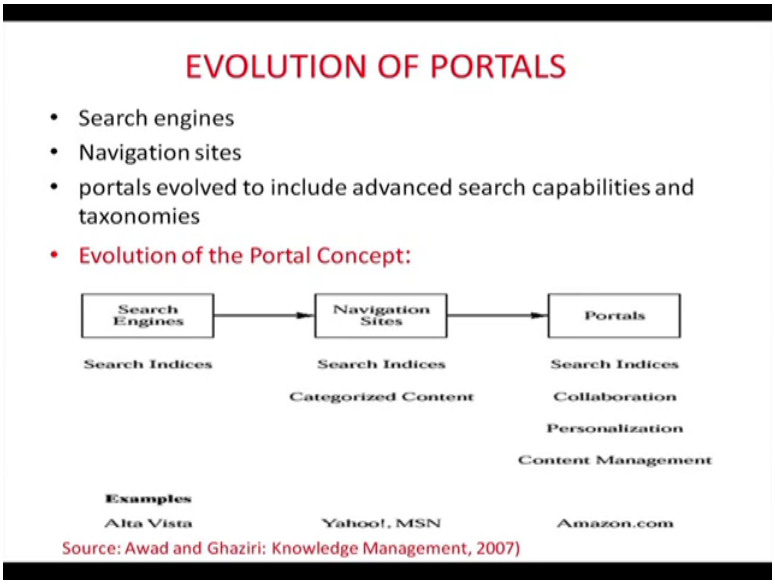
And there could be employees who are going to answer these portals okay. And that is how you can collaborate and share knowledge with each other okay. And then you can also assist the organization can assist itself in reaching out to the customers. How, because you know ultimately if you link into the customer. For example customer service centers right, virtual customer service, service centers like we are call centers today right.

And then you simply call them and get assistance from there depending upon the kind of requirements that you have okay. So knowledge portals are very important because it allows producers and users of knowledge to interact with each other okay. And then you can both of them have better interface okay, knowledge producers interface, knowledge producer interface and knowledge customers using interface okay.

Now moving further if you look at portals that have come up, as the way of development we started the search engines now. You know that we have lot of search engines, we have also lot of navigational sites like Google, we have a say these are navigation sites or Yahoo's is navigation sites right. All types they are search engines right, there are lot of search engines which are available which help you to basically retrieve useful information right.

So you start from search engines to navigation sites to portals so we are more way ahead search engines to navigational sights to portals. So portals have evolved as a way to have both the capability of a search engine and also a navigation sites right, because if you look at portals it does not only include search engines, where you can collaborate through portals and all kind of things. But it is more personalized because you can interact one-to-one with people okay.

(Refer Slide Time: 13:40)

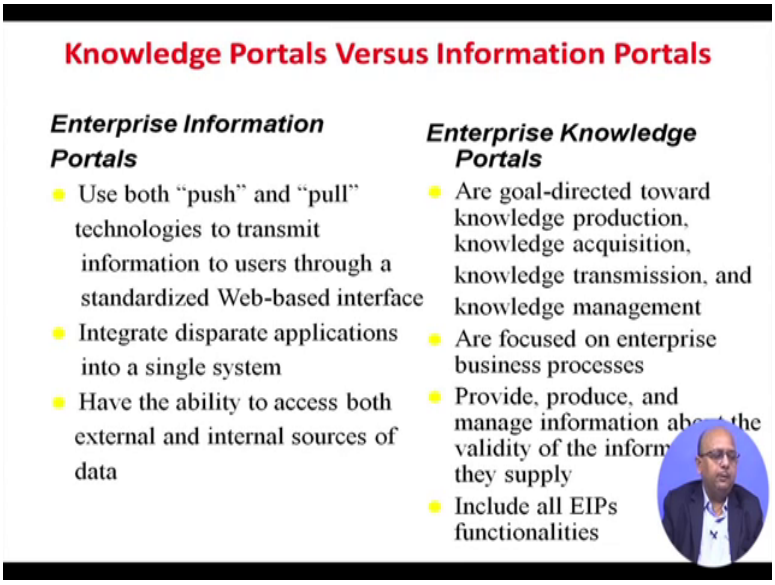


And you have lot of archive saved so far as content is concerned. So you also need some kind of architecture to manage the content that is available at the portals okay. For example Amazon.com oaky, it is the port it is interactive you can look for a product okay. You can collaborate also you can sell their also, it would be more personalized depending upon the requirement you can see what is required and then accordingly the product can be customized for you.

And then lot of the lot of content and then you can search out of this content what is required by you. Navigation sites like we have Google, yahoo and these kind of things basically you have search attributes and then you can classify and categorize the content based on the relevant information. So the portals have come a long way starting with search engines to navigation sites okay.

So if you look at portals they are more important compared to any search engines or Navigation sights okay.

(Refer Slide Time: 14:31)



Now we have two kinds of portals one is known as knowledge portal the other one is in information portal. Now if you look at information portals what all will happen, information portals only have information okay. The only thing is that you have to access this information and then you have to see that how you are going to make use of it. The source of information could be internal or external, but when I am talking about knowledge portals became knowledge portals are basically having a different kind of characteristics or trades right.

It is aimed at certain things right. You are going to make use of knowledge for certain activities right. Not only that but it is also related to different cycles of the knowledge management like acquisition, production, transmission or management also okay. So it is not only limited to say just say content management right.

For example information portal basically manage the content, but informed knowledge portals are way ahead and they compete they basically look at the entire KM cycle starting with acquisition, production, transmission and management right. And it is very important to have these kind of things because and in an enterprise portal basically gave that he is had just one system.

But knowledge portals would be having different kind of systems related to different activities of the knowledge management like knowledge sharing, knowledge discovery, knowledge acquisition, knowledge transmission knowledge right, so this difference between that.

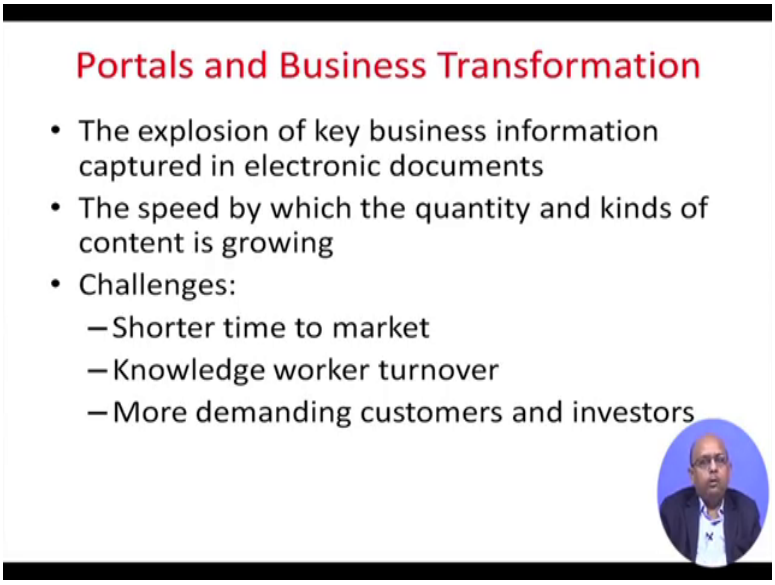
(Refer Slide Time: 16:12)

And then when you are going to have knowledge portals. You have to see that the major challenges that people are going to make use of it for business okay. So you have to make sure that you are able to optimize the use of knowledge portals by the people to make use of it in their day-to-day processes so that they are able to improve the quality and reduce the cost okay.

So it is very, very important that people are going to make use of this knowledge portal and for that what you need to do to communicate and tell them what kind of information is there in the knowledge portal and how they or what they can be used and once it is done you are going to optimize the performance operational processes and at the same time can also reduce the cost.

Then you also need to commercialize your products at lower processes, so you need to communicate to people that what are the advantages of using these knowledge portals for working effectively.

(Refer Slide Time: 17:11)



Then when you want to look at this that how the portal is going to help you to transform your business so the basically what happens in a portal the knowledge is available in a documental form electronic documents right. So this is changing very fast and the quantum of knowledge that is going to be available tomorrow is going to be huge right. So the speed at which the quality and quantity of knowledge growing is very tremendous.

So the one major problem that these portals are facing is that how to archive, classify and manage the content relate to the knowledge of a who's the huge knowledge base that is coming up okay, because you have less time to market okay. There is a turnover of the knowledge workers because they are leaving the jobs right.

And see so you have more demanding customers so you have three fronts first from the customer's fronts you need to satisfy the requirements in terms of products and services you also need to look at employees you also need to make them happy satisfied okay. And meet the expectations and challenges otherwise they would leave then there is a shorten time to market that is since product life-cycle is very short.

And you know that more and more innovative products are coming. So it has become a major business challenge for organizations to come out with new products and new services so that they are able to retain customers okay. So since they have shorten time to market shorted product life-cycle it is very, very important that how these knowledge portals are going to be used by the people to be more creative and innovative in their performance.

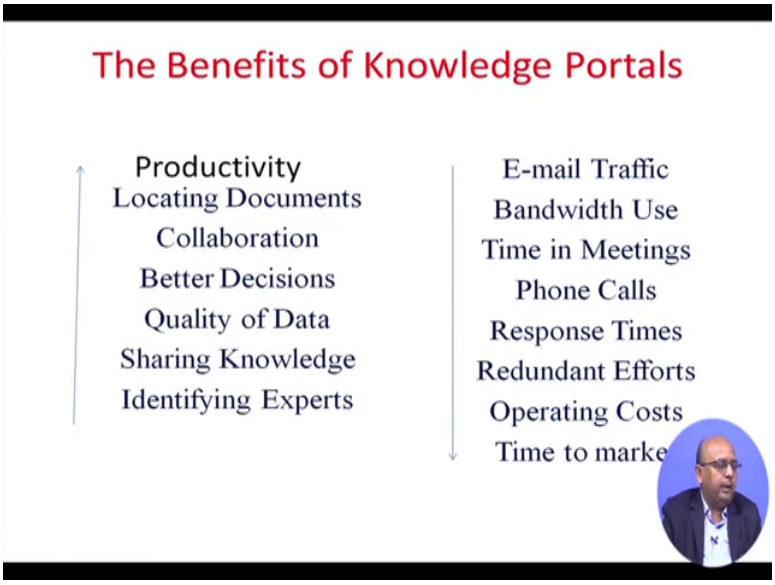
(Refer Slide Time: 18:50)

Then most of our organizations have already launched these kinds of programmers and there was survey which was done to see that what are the motives which the organizations have gone for these kind of knowledge mining programs. See most of the business objectives have one important motive that is to make profits okay. Apart from making profits and the major talent that was identified is that they want to retain key talent with the organization okay.

And they also want to retain customers, so the most important findings suggest that most the organizations have gone for KM systems to increase their market, profit, revenues right. Going for more products and services new markets, reducing their costs okay and also to retain customer services okay. But it does not act about employee except for retaining key talents and expertise because they can directly link with the profit and services right.

So if you look at the motive for organizations to launch this kind of program is very clear that they want this to make use of this KM system to make profits or better revenues retain talents right. Where customers increase their market share alright, then introducing new products and markets right and also ensure that they are able to reduce costs and improve the quality of the products right. So these are the motives for which organizations have gone for.

(Refer Slide Time: 20:26)

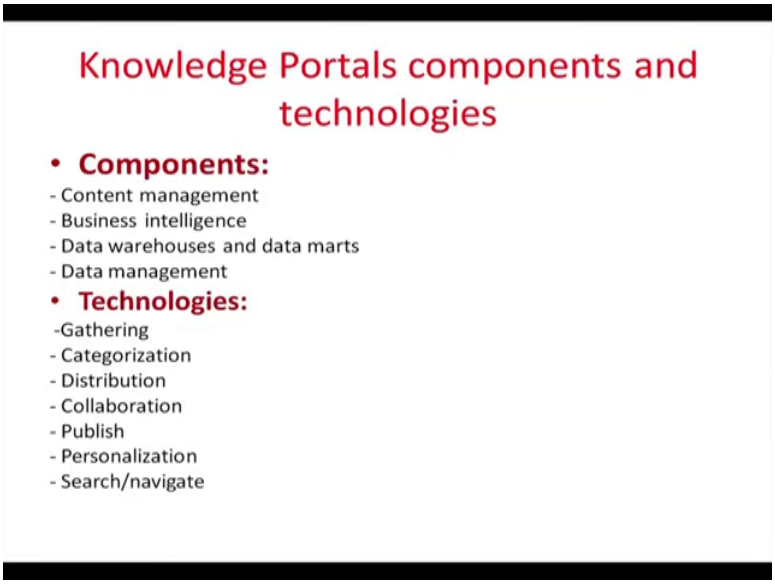


So now if you look at productivity definitely, how it is going to be done see starting with if you look at the arrows it shows that how you move up and so while as the benefits are concerned the hierarchy starting with identifying data begin over sharing knowledge then you look at the quality of the data, then how this quality data is going to help you to take better decisions through data mining okay.

And how can collaborate across verticals vertically or gently share the knowledge and also identified documents, so these are the activities are which can be done, but all this need to be linked with productivity of the organization. Now when it comes to knowledge portals what are the use, you can manage your email traffic, to have to look at your bandwidth, you to see that how much time were going spend on meetings, phone calls or what is the response times.

But therefore that you make which is useless actually what will operating cost, how much time would take to market a product right, so these are some of the issues with these knowledge portals can help you.

(Refer Slide Time: 21:33)

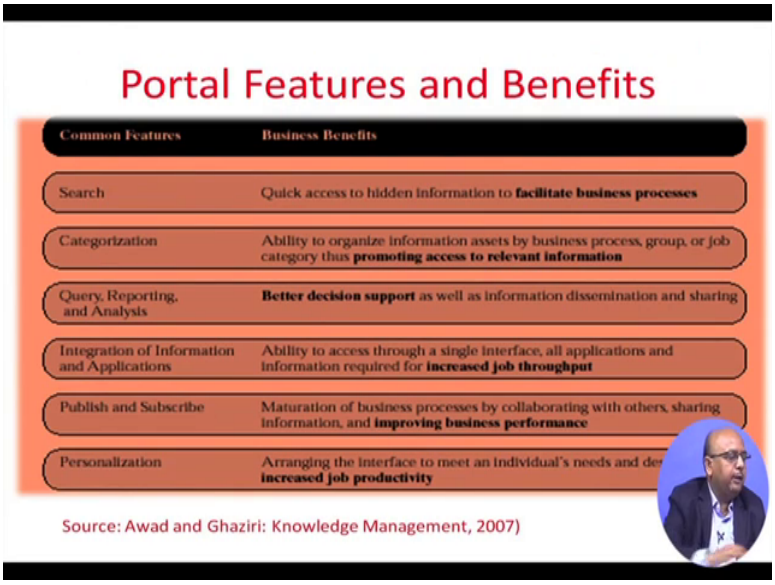


Then going further than you look at the content of a knowledge portal, what does the knowledge portal include it include content, it include information related to business intelligence that is was a mined data, converter the decisions that have been, and how to implement the decisions in the organizers knowledge and knowledge portals also include data warehouses.

And data marts and also you have to see that you are able to better manage their data in terms of quality, accuracy in this kind of things. Now what are the technologies that is used by these knowledge portals include: collect data, you are going to classify the data, distribute the data, across groups and people and associate to collaborate publish this, personalize information, using the knowledge portals.

And then how are going to search and navigate on a different kind of information or knowledge that is available with the organization.

(Refer Slide Time: 22:33)

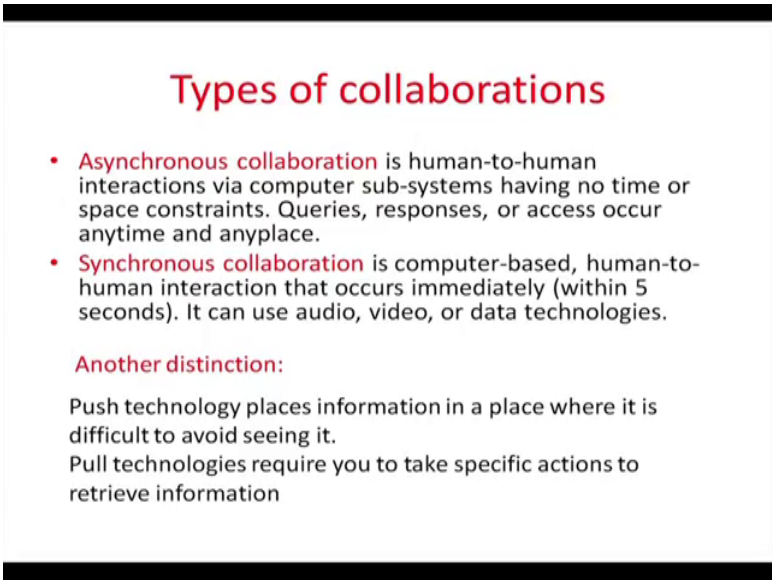


Then going further look at this likes of our example search, the weather have quick access to information to facilitate certain newspersons were not however to categorize different kind of information then when it comes to query, reporting, analysis okay, how it helps you to go for better business support right, then and not only a business decision support also howled to disseminate information and share knowledge with each other.

And then how are those integrate information and applications, you have see that how all access information, what kind of interface is there, how are going to use that interface that, especially publicized, subscribe that is how with different business process with a collaborated and shared are going to help you to improve a business that is very, very important, because ultimately what matters is bottom-line right.

And then you go for personalization also that is who the kind of interface that you have it to remove personal where one to one interactions going to be there than that helps you to increase your performance.

(Refer Slide Time: 23:46)



Then what kind of collaboration is there, then that would be human to human interaction that is asynchronous or there could be synchronous, where you are going to a computerized right, so even asynchronous collaboration with which going to be two people and then you can say that yes, you are linked with the computers, you are able to talk, were going to ask questions response with it, without any time video conference.

Like you are using is a video conferencing or Skype right, so two people can collaborate and interact with each other, without any constraint of time and space and then when asked questions and respond and then the advantage that is anywhere and then accesses there. Then synchronous collaboration is that if it is also a human to human, but computer-based but it happens immediately right.

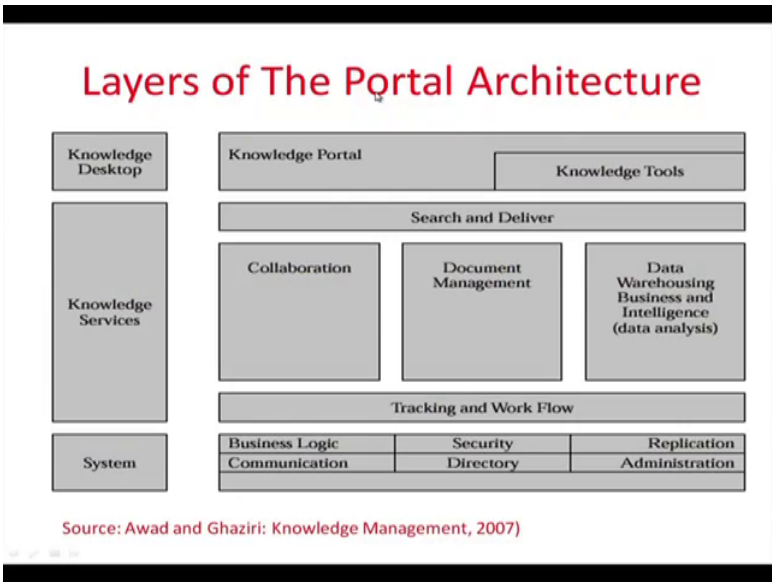
For example if you have problem then again you use the audio video or the data technologies find it right, but in asynchronous basically collaboration where you do not have any time and space right, but synchronous collaboration the only difference is that it has to be done immediately because the time and space constraints and then the job is to be the immediately and the distinction is made between an asynchronous and synchronous is that what kind of technology you are using.

Push technology or pull technology okay, so push technology where information is in a place that is difficult to avoid us a button technology requires you to take specific actions to retrieve information means the data contained in there and a new try to extract the content to get them information, was finale what happens yes it is pushing you to use it right it means

being forced by that kind information to go far right.

And in pulling what happens that is you have to go for it, so push technologies that the very attracted to go for it and pulled collect valued if you need it and then you look for it with and that is the difference between pull and push technologies.

(Refer Slide Time: 25:52)



Layers of the portal architecture that is how the portal architecture looks like a materialist support testing a prototype web of a portal architecture and so it include raising that is a knowledge extra services, systems available at knowledge is the blue flag with the knowledge sport available knowledge to the recoverable then that is our research and deliver information in knowledge services you have three kind things.

Collaboration that is how you are going to interact and relate with others, document management which is there and the how it can use and then you have facilities for data warehouses are in business and race including better right, then this is linked with tracking or flow, because there is system basically which can include all kind of things lousy communication varieties also included security, directory replication and that there are web administrators.

So this is how a portal architecture looks like and it has three parts first part, second part and third part, this is system related, this is services related, this is what like that is well-known to interact right.

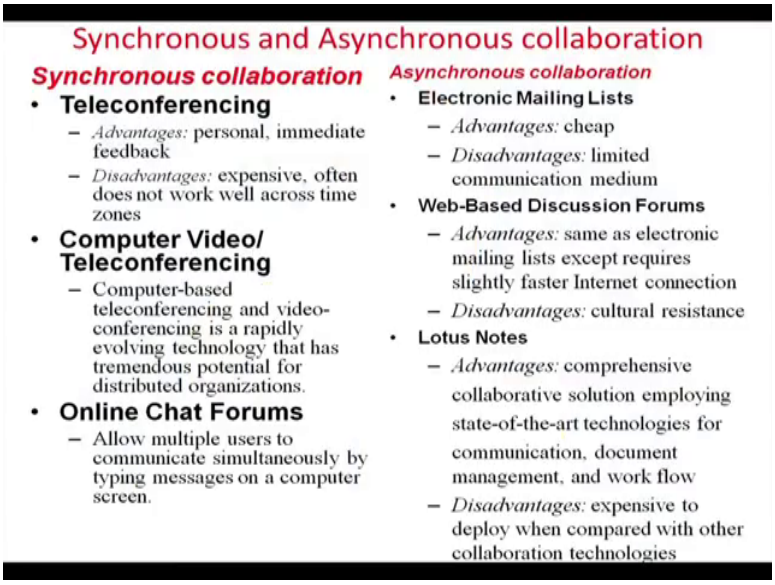
(Refer Slide Time: 27:04)

Now if you look at collaboration tools you have two safe you are using a lot of collaboration tools like web browsers, emails, right search functions are there, you have multipurpose data and then that you do have collaborative services also when you can have access and easy-to-use data and Web services, you have indexing services also for full-text search of documents, right.

For example if you go to library databases right and then run different kind of databases are posted by these libraries, so if you go to these particular database to click their and then you provide a full list of our databases or the data relate to different journals or research papers and the nuclear comeback in them and get all text documents right, so that is what indexing service.

Ad it is well organized central storage, location was your central library where it is located and from where you interact and relate yourself right.

(Refer Slide Time: 28:02)



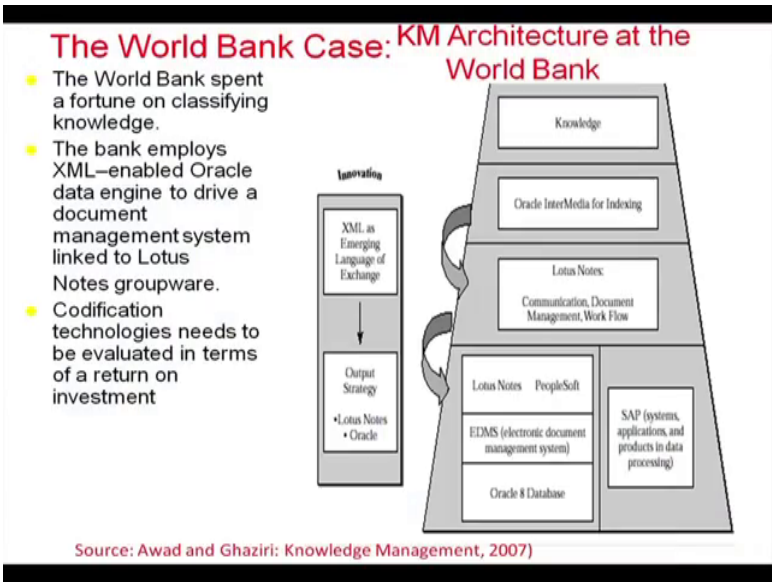
Now if you look at that synchronous basically, synchronous collaboration happens through as I told you through teleconferencing, video conferencing and online, chat Skype, most of the online things are happening like teleconferencing the advantages that it is immediate, the disadvantages are yes it expensive if it does not work well then you have a problem with the conferencing is again and says computer connected actually.

And now you have a technology but you need to see that you have a system in place which is meant to help you to for these kind of things right, for example you need to have a good connection, good bandwidth at the end of an them have online chat, forums okay, you have Tweeters, Facebook, forums are available so you can communicate in number with people depending upon the requirement using a computer.

But only requirement is that the systems requirements must, then you have email is like an example I want to send a mail to a lot of people like, our institute we have a Gmail system and information that of those as simulated to the people into all the people, then you have blessed with decision discussion forums, we have a portal where people can complain and give advice and are also now make complaints are endless, suggestions also at its what needs to be done.

Then you have Lotus Notes from IBM and its very complacent and is basically provides you the complaints of solutions on because the technology is very, very strong here and is used the secure accommodation as well as document management that what is that these kind of fire and techniques are very expensive when it is compared to synchronous collaboration.

(Refer Slide Time: 29:49)

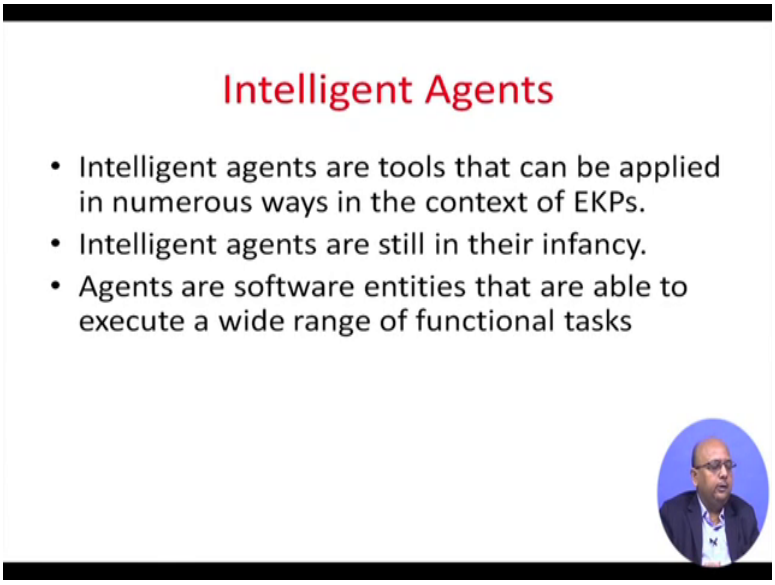


Then now if you look at the case, that how this KM architecture over this example from the World Bank okay, that the World Bank uses are the knowledge management system in a big way and they have information for classifying the knowledge so they are basically, using Excel is Oracle learned it ideal that is XML to drive document management system and that is linked to lotus notes that look at this XML is in the language exchange that is used.

And then it leads to output strategy that is Lotus Notes, Oracle of this hierarchy will architecture of the knowledge management architecture of the World Bank knowledge's their then you have a rattling or indexing your Lotus Notes for the American document management down the line you have software is like a SATA, PeopleSoft, electronic document management systems and databases.

So basically codification technologies are used to that our American on investment to see that how much is being used what is the benefit of having this kind of knowledge management system in the archives now more into this.

(Refer Slide Time: 30:57)

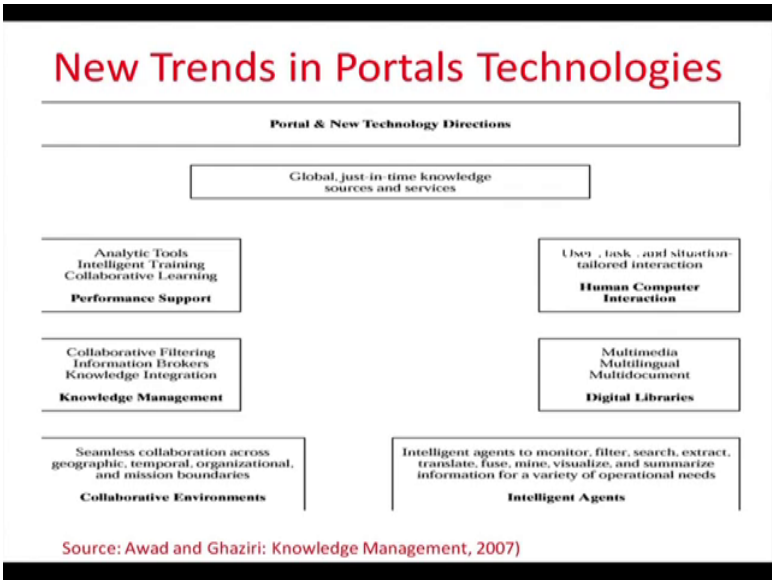


We also have intelligent agents are tools that can be applied in the context of enterprise knowledge portals okay, but it is not a devlop up much basically the software is an institute huge amount of data for different activities but it is still are in the beginning stages like my fellow practice they can offer customized service, service assistance on through online services.

For example of a query related customer service you call them it automatically routes it okay gives you reply, is still in the development phase, you can go for customer profile profiling or you can integrate to further profile of the customers with the marketing activities and you can also predict requirements of the customer using data mining techniques, or you can go for the negotiation by sitting prices or payment schedules based on your requirement.

And also financial connections and from the customers we have, they are lot of benefits but still it in the infancy stage.

(Refer Slide Time: 31:59)

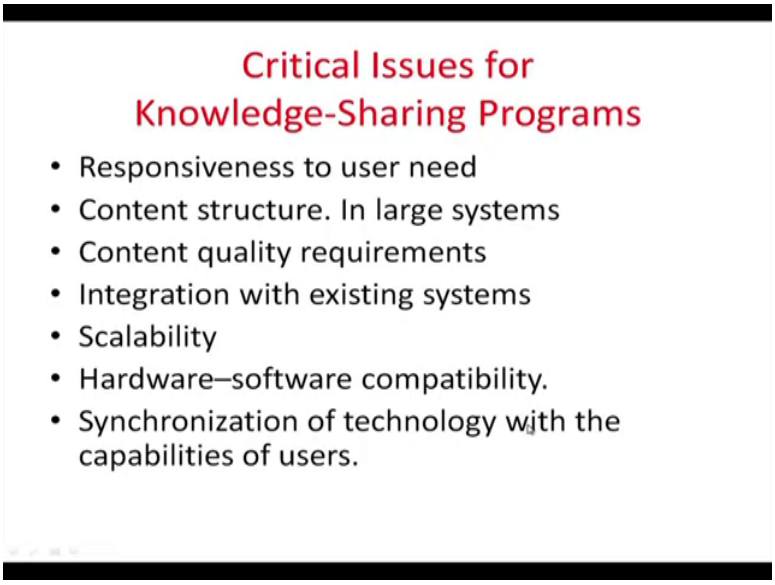


So new revenue trends in the portal technologies which has, of late just in time of solutions and services showing you have to say certain things like you have better computer, so you have more interactions, to be more situation tailored based, interaction between the people and the system similarly you are going to see that if you have an analytical tools for training and collaboration and that is very get better performance support.

Then have a knowledge management systems and also digital library right and ultimately were going to have need to create more collaborative environment and intelligent agents, these intelligent agents basically helps you to monitor, filters, search, extract, translate, mine, visualize, summarize information for different requirements.

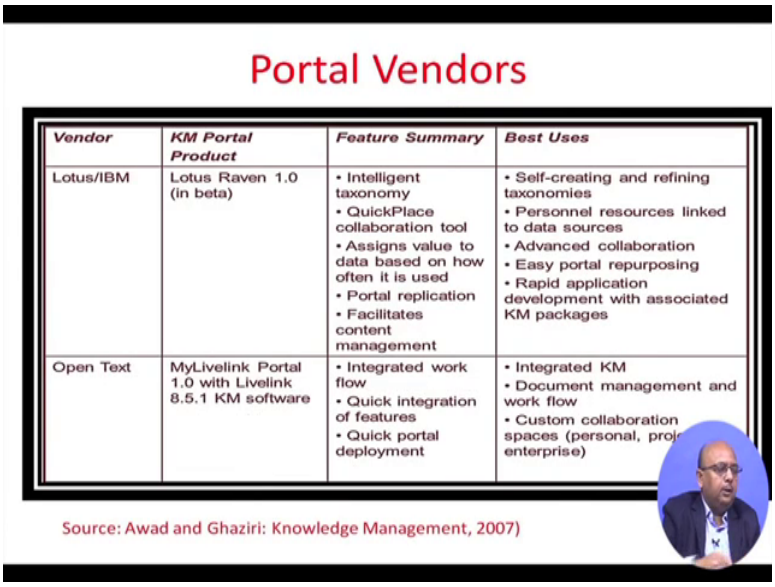
So it is a more integrated approach to look at the things in a more holistic way, not only monitoring and filtering data, mining and visualizing data but also giving a lot of information that can be used to buy you different activities.

(Refer Slide Time: 33:04)



Moving to next is that is knowledge sharing is a big issue to make sure that the content is structured in such a way the quality requirements are met right, it is integrated with the system, security issues are taken care of the synchronized with the technology so that you can use these knowledge portals for better knowledge sharing.

(Refer Slide Time: 33:25)



Similarly to see that was a different kind of portals which are available like Lotus is there, open text in their and the characteristics and the uses also define here, for example gobal Lotus Notes user self creating refining taxonomies use is more personal resource link to data sources it goes far advanced collaboration and you have and the need go for rapid application of integration and you can really associate with knowledge manufacturers.

Like open text are it is also integrate knowledge management system but it used for

knowledge man document management basically an customer of collaboration, similarly also have Plum tree which is very easy in extensive content application integration ,the advantages that its performance and scalability it has better security features or data taxonomies are taken care of, you have different than a bit accesses and can be customized also depending upon the requirement.

Then you have the webmeta engine 1.0 and is at these are the new developments which have taken place in terms of its usability, tracking site statistics, contain streaming to our list services also the revenue things which is coming up, without Internet you can may use certain applications and perform certain tasks, so these are the new developments which is coming up in and that is how we find that these knowledge portal are to be very, very useful, thank you very much.