

Knowledge Management and Protection Systems (KMaPS) Corso di Laurea in Ingegneria

Parte 6 - Social Media Technologies and Solutions Prof. Paolo Nesi

DISIT Lab <u>http://www.disit.dinfo.unifi.it/</u>

Department of Information Engineering, DINFO University of Florence

Via S. Marta 3, 50139, Firenze, Italy tel: +39-055-2758515, fax: +39-055-2758570

paolo.nesi@unifi.it, http://www.disit.dinfo.unifi.it/nesi



Part 6: Social Media Technologies and Solutions

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





Collaborative Systems

Collaborative systems are:

- CSCW solutions in which one or more objectives are reached on the basis of collaborations among users.
- Different paradigms according to the emphasis on one or more of the 4 axes:
 - Objectives of the collaboration
 - Interaction among users
 - Observation of the common environment
 - Assessment of results against objectives





Collaborative and Social

- A specific kind of collaborative/competitive systems are the Social Networks, that have:
 - Objective: general of the network
 - goals defined by the organizers, clear on user and content profiles
 - while: evolution is left to the users' objectives
 - Interaction: among users

DINFO

GEGNERIA

EGLI STUDI

FIRENZE

- mainly asynchronous (non real time) user collab./interac./comp.
- Observation: as user to user observation
 - reciprocal observation: profiles of friends, groups, forums,
- Assessment: to persecute the objectives
 - Controlling/assessing user behavior, via metrics
 - Gratify the users according to their objectives, egos, etc.
 - Analysis of: groups, forums, meetings, pages, messages, etc...

Assessment of the achievements with respect to the Objectives



Terminology

Social TV

A TV based on Social Networking principles, with the support of UGC, etc

Social Learning Management System, SLMS

Learning management system with SN features

Enterprise Social Network, ESN

Project management and control with SN features

Best Practice Network

Specific kind of social network devoted to the definition of best practices.

Social Media

- A set of technologies and solutions that exploit the social network related data and solutions.
- A Social Network based on media, multimedia

Social Network Analysis

- The discipline to analyze the social network in terms of user clustering and relationships, metrics for SN assessment, etc..
- It can be used to better understand motivation and rationales of success and/or problems.



Accrescimento della Conoscenza

 Knowledge Management SECI Model (Nonaka & Takeuchi, 1995)



Technological Pedagogical And Content
 Knowledge (TPACK) (Koehler & Mishra, 2009)



Creazione e accrescimento		
DAA	della cono tacita	oscenza esplicita
Incontri tacita	Socializzazione gruppi, forum, chat, meeting, workflow	Esternalizzazione pubblicazione, produczione, indicizzazione
esplicita Manuali, documenti	studio, apprendimento, newsletter, e-learning	aggregazione, associazioni, annotazioni, Natural Lang. Processing
	Interiorizzazione	

Stimoli Sociali





Sistemi Distribuiti, Univ. Firenze, Paolo Nesi 2011-2012

Part 6: Social Media Technologies and Solutions

Collaborative systems

Definition and Termin A

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





Social Network Objectives

Social network objectives

- goals defined by the organizers,
 - Business goals
 - Monitoring Business goals
 - Satisfaction of the target users
 - With the aim of user growth
- clear impact on
 - Juser profile
 - content profiles/descriptors: if any
- Assessment and Validation of the business goals achievements !!!





Social Network Aims





Scegliere !?

Where should you post your status?



UN DE FI

FIRENZE DELL'INFORMAZIONE TECHNOLOGIES LAB

Knowledge Management and Protection Systems (KMaPS): 2016-2017





Social Networks User Objectives

User Objectives

- each user persecutes its own objectives
- Identification of needs of the target users
 - Requirements of the target users
- For each target user kind they have to be
 - Guessed !
 - identified to create a service that is used
- Monitoring the user behavior
- Identification of the collective behavior
- Validation of the target user satisfactory !!!





Social Network: User's Motivations

Creating Social relationships and contacts

- Finding new friends and/or colleagues
- Becoming a reference among friends, get higher reputation
 - Sharing content with friends
 - Writing comments and sharing experience
 - Organizing events, providing information
- Get knowledge about what other people do in their life
 - Keeping friends/colleagues in contact

Increasing Knowledge on

- specific topics, the subject of the UGC and of the SN
- how content can be created and shared
- Iife of your connected friends





VIVERSITA

DEGLI STUD

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Social Network: User's Motivations

Personal advantages for the users

- Conquering/getting a leadership
 - Increasing visibility in the community
 - be observed/recognised by a community
- Improving position in the job/life community
 - commercial purpose

Save money for the users

- Storing user content permanently and making it accessible for its own usage (making it public as side effect)
- making content public for friends
 - Saving streaming/hosting costs
- Making business among users (e.g., ebay)
 - Selling personal staff
 - Finding difficult to find products and staff



Examples of Social Networks

Creating a community to provide a service

- Motivating target users
 - Objective: share experience, collect/provide knowledge
 - knowledge production (content, comments, annotations, etc.)
 - Collaborative work with users
 - Sharing Improving community knowledge

Creating a community to make business on advertising

- Get Content for placing advertising
- Objective: increment number of users, minimizing the costs
- Stimulating viral propagation
- Sharing friendship
- Attracting new users, replacing those that abandon the SN



Social Media Landscape





Part 6: Social Media Technologies and Solutions

E)

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





User/Content of Social Network

What they are:

- Content related Items
 - Media files, web pages
 - Comments, tags, votes
 - Aggregations and links

User related information

- User profiles and relationships
- Groups profiles (professionals ...)
- Email, messages, etc.





Social Network User Classification

Lurkers: passive users,

- take and do not contribute: no content, no other users,
- can be even frequent users to read only
- they are typically invited and does not invite

Occasional users:

- sometimes they also contribute with content
- marginal active in terms of invitations

Active users:

- frequently contribute
- The first source of invitations of users and content

Pushers:

Typically active users paid to stimulate activities with content, discussions, users, mailing, etc.





The role of the Pushers

- Many SNs are promoting/pushing the most played/accessed content in the last months/days and weeks (they are most clicked content items since are most frequently presented)
 - The entrance of a content/object into those lists is a strong opportunity for marketing and promotion
 - The entrance in those lists depends on number of plays/votes, comments, etc.
 - They can be artificially created by a new human figure, that has to be a SN Users, the PUSHER

The Pusher:

DINFO

INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ Degli studi

FIRENZE

- Has to be a widely linked person
- Can put the content in his preferred, may be removing the others that he has appreciated to make the last more evident,
- Can promote/propone the content to other friends
- Can make a lot of initial play and votes to place it in a better place
- For some aspects can be robot, a bot



User Activities on Social Networks

Wikipedia (2006)

- 68000: active users
- 32 millions of lurkers
- While the 1000 more active users produced the 66% of changes.

Similar numbers in other portals:

90% lurkers

JNIVERSITÀ

degli studi FIRENZE

- 9% occasional users
- 1% active users

DINFO

INGEGNERIA DELL'INFORMAZIONE

- 90% is produced by the 1% of active users
- 10% is generated by the 9% of users including the occasional

🔊 Wikipedia (2012)

- 77000: active users
- 450 millions of unique user per month
- 22 millions of pages
- While the 1000 more active users produced the 66% of changes.











Social Network Activities meaning

- Since the 90% is managed by a small percentage of active users:
 - Votes are also produced with the same small part of the community
 - Comments, tags, annotations are also produced with the same small part of the community
 - Pushers are frequently needed to create activities and waves into the Social Networks, they create fashions and interests among the lurkers, etc....

Number of plays/accesses are produced by the whole community





Part 6: Social Media Technologies and Solutions

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication



Different User Generated Content Items

Media Content

- Classical: audio, video, images, document, animations
- Cross Media Content: interactive content, widgets, applications, procedures, courses, ….

Collections, aggregations

- Essay, courses, playlists, etc..
- Web pages, panels, Wiki

Annotations and comments on

- 🐥 media content, web pages, wiki
- Contextual on audiovisual
- Links among media and issue
- **G** Forum,
 - Forum topics
- Blogs

UNIVERSITÀ

DEGLI STUD

FIRENZE

Messages, tags

DINFO

INGEGNERIA DELL'INFORMAZIONE



static



UGC Pros

- No costs for hosting and distributing UGC
- WEB sites that host your content and provide some tools to make them accessible on web for your friends for free, if you accept to make them public or close to public
- Natural selection/emergence of better UGC items, increment of visibility for some of UGC users...
- Annotation and reuse of UGC of others users and friends
 Gratification about the promotion of UGC
- Simple search on your UGC





UGC Cons 1/2

Restricted social penetration since

only User with are ICT skilled and have a certain economical capability may access to internet and spend time to enjoy SN

Lack of privacy control, lack of DRM

- too much information is requested
- some people do not expose their true personal info
- usage data are used to profile the users' preferences in any way so that the user profile is reconstructed even with GPS locations.
- See Terms of Use

IPR problems:

UNIVERSITÀ Degli studi

firenze

- Violation of IPR of third party content, free usage of UGC
- Lack of control about your own Content and UGC
- Reuse and annotation of professional content
- See Terms of Use

INGEGNERIA DELL'INFORMAZIONE



JNIVERSITÀ

DEGLI STUD

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Content Descriptors

- Static aspects: content description not changing over time. They are:
 - metadata, keywords extracted from description, comments, etc.;
 - technical description (as the Format in the following): audio, video, document, cross media, image,..;
 - content semantic descriptors such as: rhythm, color, etc.; genre, called Type in the following;
 - **groups** to which the content has been associated with;
 - taxonomies classification to which the content has been associated, taking into account also the general taxonomy;
 - dynamic aspects may be related to:
 - user's votes, user's comments;
 - number of votes, comments, downloads, direct recommendations, etc.
 - List of content played, related taxonomy ;



UNIVERSITÀ Degli studi

FIRENZE

INGEGNERIA DELL'INFORMAZIONE

Content Descriptors, 1/3

- **Classification** based on Metadata (static)
 - For example: Dublin Core, Mets, ... multi-instance and multilingual fields, taxonomy, dates, locations, etc.
 - Mainly provided by users at the upload
- Identification codes: (static)
 - 🜲 as ISAN, ISMN, ISBN, ISRC, barcodes, URI, ...
 - More diffuse on professional content
 - Provided by users at the upload and/or by the SN manager
- Technical information: (static)
 - As: Size, format, source, mime type, color, tonality, url, duration,
 - Estimated at the upload or when processed for distribution, so that when several formats are produced
 - Formalized in MPEG-7, including fingerprint,



INIVERSITÀ

DEGLI STUDI

firenze

DINFO

INGEGNERIA DELL'INFORMAZIONE

Content Descriptors, 2/3

- Geotagging: (static)
 - Grab GPS position, placement on Google map
 - Collected via: IP resolution via services, GPS signal, ..
 - Provided by users at the upload or by common users
 - Referred/associated Position (dynamic), by human

Standard Tags among those proposed by the SN (static)

- Classification as a function of the content format
 - Taxonomic classification, multilingual
 - Support of dictionary/vocabulary and/or of an ontology
- Provided by users at the upload
- □ Free Tags, such as Folksonomy (dynamic)
 - Support of dictionary and/or ontology
 - Provided by common users



Content Descriptors, 3/3

Marking: (dynamic)

preferred, uploaded, suggested, viewed

Annotations: (dynamic)

- Links to other URL
- Citations to articles in form of links
- Relationships among resources: aggregations: playlist, collections, courses
- Textual annotation as comments
- 🐥 Votes, ranks, …
- Mainly produced by humans

Contextual Annotations (dynamic)

- Description of the scene, E.g., on an image/video:
 - mark area in which a CAR is present, addition of a text
 - Mark area and time windows in which Carl and Jack talk, …
- A Mainly produced by humans, may be deduced for similarity



Votes/ranks, Comments, Preferred

- **Users** may leave on Content and Users:
 - Ranks and Votes (positive or negative)
 - Comments on content items, web pages, other comments, forum lists, groups, (positive or negative) (sentiment analysis)
 - Comments may be left as
 - Text: simple messages in a context, tags
 - Content: video, audio, images, etc.
 - →Emoticons ☺ ☺ ……
- User may mark the preferred content and users (friends)
 - Preferred content are accessible with a direct list to shortening the time for their play





NIVERSITÀ

DEGLI STUDI

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Linguistic Complexity

Multilingual:

- Speech: video, audio
- Text on: document, web pages, comments, subtitles, annotations, etc.
- Metadata (text): title, description, author, location, date, etc...

Multilingual Complexity:

- Indexing and querying
 - Translation of metadata or of queries
 - Full text of documents, frequently only one language
- Linguistic processing of text to get semantics (see part of the course on NLP)
 - extract Contextual Annotations
 - Junderstand comments: positive/negative



Content Aggregations

A content aggregation:

Creating aggregated content

- Collections, Essay
- Courses
- Playlists

Creating links and relationships among content

- Annotations: contextual, visual, etc.,
 - see you tube and Flickr annotations on video and images
- Audiovisual annotations
 - E.g., Creating links from a video to another

IPR problems of content aggregation

Aggregate means taking derivative works, you need to have the rights to do it.




Part 6: Social Media Technologies and Solutions

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





User Profile

Static user profile aspects

- generically provided during registration
 - frequently not so much detailed in generic Social Networks,
 - users prefer to avoid filling in 'useless' forms and/or to provide false data.
- In small thematic and business oriented Social Networks the information is much more reliable.
- Dependent on the Social Network objectives

Dynamic user profile aspects

generate on the basis of the user's activities performed on the SN elements, such as the actions performed on

content, other users, on groups, on chat, etc.

estimated/inferred by assessment/analysis





Static Aspects of User profile

Static information collected during registration++

- Name, surname,
- Nationality and languages (multiple)
- Genre, age, etc..other personal info,.
- Instruction/School, work, family structure, etc.
- Personal photo
- Jobs: several different jobs with periods, etc..
- Competences: several skills
- Economical data: range, etc.
- Explicit Preferred content:
 - topics, genre, period, area, etc.
- Subscribed (slow dynamic):
 - →lists, groups, ..



Dynamic Aspects of User profile

- dynamic information collected on the basis of the activities:
 - votes and comments/annotations on:
 - contents, forums, web pages, etc.;
 - downloads and play/view/executions of content, web pages, etc.;
 - uploads and publishing of user provided content;
 - marked content as preferred/favorite;
 - recommend content/groups or users to other users;
 - chat with other users, publishing forum topic on groups;
 - queries performed on the portal, etc.;
 - create a topic in a forum or contribute to a discussion;
 - relationships/connections with other users or groups;
 - 🜲 Etc.





Dynamic aspects of user actions

- Statements written on blogs and micro blogs
 - Short Comments in a context
 - Recurrent user and statement
 - The same statement on more than one blog (pushed by pushers)
 - Dates and time, successive blog posts
 - Statements on comments
- Assessment of:
 - Market trends, market vigilance
 - Pharmavigilance





NIVERSITÀ DEGLI STUDI

FIRENZE

INGEGNERIA DELL'INFORMAZIONE

Esempio di Profilo utenti

> Informazioni statiche: •Informazioni generali: •nome, cognome, sesso, • foto, data di nascita, •descrizione personale, •località di provenienza (ISO 3166), Nazione Suddivisione Provincia •lingue parlate (ISO 369) Informazioni di contatto: •lista di contatti di instant messaging •Scuola e Lavoro: •scelta del livello scolastico. •nome della scuola. •tipo di lavoro, •nome del posto di lavoro •Interessi: •Vettore contenente la lista di valori del campo Type degli oggetti scelti dall'utente

- >Informazioni dinamiche:
 - •Lista di oggetti preferiti
 - •Lista di amici
 - •Lista gruppi
 - •Voti positivi ad oggetti
 - •Commenti ad oggetti
 - •Blog post

 Informazioni sulle preferenze sulla base delle visualizzazioni degli oggetti •Format

- •Type
- Taxonomy



UNIVERSITÀ Degli studi

FIRENZE

DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Friends and Friends

Facebook Friends







Descriptors of: user group

- Groups of users they may have specific descriptors and those inherited by the users:
- **static** aspects of the groups such as:
 - objectives, topics, web pages, keywords, taxonomy, etc.;
- **dynamic** aspects related to:
 - users belonging to the group; users may: join and leave the group, be more or less active over time;
 - content associated with the group: files, comments, etc., with their taxonomical classification, metadata and descriptors.





Groups vs Channels

Groups:

- The Group has an objective to persecute: thematic, goals, etc.
- Users belong to a group
- The users of a group may have advantages in terms of accesses at the users profiles and at services
- A group may have a distribution channel, a discussion forum, a mailing list, etc.

Channels:

- The channel is a distribution group for content.
- The channel is typically only a way to access at content,

Collective Intelligence

Modeling of almost uniform group of users with a collective profiles that represent a collectivity





Group vs Circle

- A Circle is only a collection of users without a circle coordinator/moderator.
- The concept of Circle in the context of SN has been introduced by Google+
- Main Circles kinds are:
 - Friends, family, known, following
 - Personalized circles can be created as well





Part 6: Social Media Technologies and Solutions

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





Social Network Analysis Metrics

- The SNA is mainly focused on evaluating the status of the network
- Relationships and metrics that give an idea of the evolution of the SN and of the healthy aspects user and content:
 - Which are the most important persons
 - Which are the most active people
 - Which are the critical conditions
 - Which are the major drivers of growing
 - Which are the most interested aspects/content/feature in a given period
 - Which are the most relevant topics of interest
 - Which is the most used service/functionality in the SN
 - 🐥 Etc.





User Profile Problems

Different data types:

- Numbers: age, votes, #kids, ..
- Enumerates/symbolic: language, nationality, etc.

Multiple Values / Selections:

Ianguages, nationalities, preferences, etc...

Non-Symmetrical Distances, for instance:

Preferences: Dim ({Pref(A)}) ≠ Dim ({Pref(B)})

Dynamic information

- related computational complexity
- Different Languages of comments, descriptions,
 - Language processing and understanding
 - Dictionaries, Semantics, Taxonomy, etc.





Relevance of Users

Number of Connections with other users

- Direct connections,
- Second and third level connections,
- Etc.

Number of accesses to their

- profile page (if any)
- posted and/or preferred content
- Comments
- 🐥 groups

INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ Degli studi

firenze

Users' Activities

- Number of posted content in time
- Number of posted comments, on content, on area...
- Number of votes per content, per area, etc.
- Number of accesses to the network







NIVERSITÀ

DEGLI STUD

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Issues on Community Graphs

- Absence of not connected users that may be the majority
- Presence of a main Center of gravity
 - Presence of dense groups with leader or reference users

Number of Connections

- Distribution of connections
- Density of connections

Presences of remotely located small Groups

- Self connections among these people
- Some of these smaller remote groups are linked with the rest via 1 or few more chains of single people

Depending on their activities, the risk of losing those communities is evident



friendship propagation

User links and friendship propagation....

Mechanisms for invitation

- User A invites N Users
- Among these N Users, M Accept the invitation

Viral Indicator

- If M > N a mechanism of viral grow is started
- It can exponentially grow up or to simply produce a small pike

Users have:

firenze

- Direct Friends------ for example: 90
- Indirect Friend of different levels -----: level 1: 900
- Friends via groups (see LinkedIn) -----: 14000







Your Network of Trusted Professionals

You are at the center of your network. Your connections can introduce you to 4,397,200+ professionals — here's how your network breaks down:					
Your Connections Your trusted friends and colleagues	285				
Two degrees away Friends of friends; each connected to one of your connections	61,000+				
Three degrees away Reach these users through a friend and one of their friends	4,335,900+				
Total users you can contact through an Introduction 4,397,200+					
26,016 new people in your network since November 17					

The LinkedIn Network

The total of all LinkedIn users, who can be contacted directly through InMail.

Total users you can contact directly - try a search now!

85,000,000+





Matrix of connections

Andre Diane Beverly	Fernando	ather)(Ik	Jar ce		[<i>i][j]:</i>	mati	rix	of co	onne	ecti	ions	5
Ed	A **			Dian	Fernand	Beverl			Heath			
	Aij	Carol	Andre	е	ο	У	Ed	Garth	er	Ike	Jane	NC
	Carol	0	1	1	1	0	0	0	0	0	0	3
	Andre	1	0	1	1	1	0	0	0	0	0	4
	Diane	1	1	0	1	1	1	1	0	0	0	6
	Fernando	1	1	1	0	0	0	1	1	0	0	5
	Beverly	0	1	1	0	0	1	1	0	0	0	4
	Ed	0	0	1	0	1	0	1	0	0	0	3
	Garth	0	0	1	1	1	1	0	1	0	0	5
	Heather	0	0	0	1	0	0	1	0	1	. 0	3
	Ike	0	0	0	0	0	0	0	1	0	1	2
	Jane	0	0	0	0	0	0	0	0	1	. 0	1
	nc	3	4	6	5	4	3	5	3	2	. 1	36





Social Network Analysis Metrics

- Degree of Centrality of a node
 - Number of connections to a certain node Andre Andre
 - can be non symmetric if the relationships are not symmetric, thus the graph is oriented
 - 🐥 Diane
 - has 6 connections



- ⇒ is connected to others which are in turn connected each other.
- It is not true that to count connections is the best model to identify the most relevant node. In the above case:
 - Diana is connected to people that are in any case connected each other.
 - While Heather is central to keep lke and Jane connected to the rest of the network !!





Averaged Number of connections

- Total number of connections divided by the number of Nodes
- According to the examples above:
 - NC, Number of connections: 36
 - they are considered non bidirectional otherwise they should be 18
 - NN, Number of nodes: 10
- Averaged number of connections:
 - ANC: 36/10, 3.6 connections per node
 - **ANC**: 18/10, 1.8 connections per node



It is more similar to the user perception to say 3.6 connections rather then 1.8



Matrix of distances

D[i][j]: matrix of distances N*(N-1)/2 elements

Dij	Carol	Andre	Diane		Beverl v	Ed	Garth	Heath er	Ike	Jane	
Carol	0			1	-				3		18
Andre		0	1	1	1	2	2	2	3	4	10
Diane			0	1	1	1	1	2	3	4	13
Fernando				0	2	2	1	1	2	3	1:
Beverly					0	1	1	2	3	4	11
Ed						0	1	2	3	4	10
Garth							0	1	2	3	(
Heather								0	1	2	
lke									0	1	1
Jane										0	0
											89



Carol

Diane

Fernando

Heather

Ike

Jane

Andre

Beverly



UNIVERSITÀ Degli studi

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Averaged shortest path from one person to another





Eccentricity of a node

Eccentricity:

- the max distance of a certain node with respect to all other nodes of the network
 - *→Ecc(Jane)* = 4
 - →Ecc(Fernando) = 3
- See Jane column on right side of table of Distance Matrix







Centrality of a node

- **Eccentricity**:
 - Ecc(Jane) = 4
 - Ecc(Fernando) = 3



- Centrality Ce():
 - Ce(Jane) = 1/4
 - Ce(Fernando) = 1/3
 - Ce(Heather) = 1/2



Heather is the node with max Ce() since it can reach all the nodes with max 2 hops.





Closeness Centrality of a node

- Closeness Centrality: the reciprocal of the sum of all the distances of that node with respect to the other nodes
 - See column on right of the distance matrix
- Fernando and Gary have a lower number of connections with respect to Diane, on the other hand they have the best position to access to all the other nodes.



- They have the best view on what happen in the network.
- Cc(Fernando) = Cc(Garth) = 0,071

$$C_C(v) = \frac{1}{\sum_{u \in V, u \neq v} d(v, u)}$$

INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ Degli studi

FIRENZE

				Fernan	Beverl			Heath				
Dij	Carol	Andre	Diane	do	у	Ed	Garth	er	Ike	Jane		
Carol	0	1	1	1	2	2	2	2	3	4	18	0,056
Andre	1	0	1	1	1	2	2	2	3	4	17	0,059
Diane	1	1	0	1	1	1	1	2	3	4	15	0,067
Fernando	1	1	1	0	2	2	1	1	2	3	14	0,071
Beverly	2	1	1	2	0	1	1	2	3	4	17	0,059
Ed	2	2	1	2	1	0	1	2	3	4	18	0,056
Garth	2	2	1	1	1	1	0	1	2	3	14	0,071
Heather	2	2	2	1	2	2	1	0	1	2	15	0,067
Ike	3	3	3	2	3	3	2	1	0	1	21	0,048
Jane	4	4	4	3	4	4	3	2	1	0	- 29	0,034
											89	



UNIVERSITÀ Degli studi

FIRENZE

DINFO

INGEGNERIA DELL'INFORMAZIONE

Betweenness Centrality of a node

- Betweenness Centrality: control degree of a node about the information flowing among other nodes
 - the ratio between the number of shortest paths between vertex s,t in which the node (v) is involved

$$C_B(v) = \sum_{s \neq t \neq v \in V} \frac{S_{st}(v)}{Total(S_{st})}$$

- No shortest path passes via Carol, Ed and Jane to connect a couple of other nodes:
 - Sij() for them is Zero
- Heather is important since without it:
 - Ike and Jane would be cut out.

	Sij	Cb(v)
Carol	0	0,000
Andre	2	0,034
Diane	7	0,121
Fernando	11	0,190
Beverly	2	0,034
Ed	0	0,000
Garth	11	0,190
Heather	16	0,276
Ike	9	0,155
Jane	0	0,000
total	58	





Clustering Coefficient

- In SN or in any large group of related data it very important to identify clusters.
 - They can be independent or overlapped.



- Clustering Coefficient: the ration between the number of connections with its neighborhoods (neigh) and the max number of possible connections among them.
- In the above presented example the results are reported in this slide
- It can be estimated for all the network as
 - (Num links)/(max num links)
 - Our case: 18*2/(10(10-1))=0,04
 - 🐥 Average: 0,49

	Neigh	Clust Size	Ci()
Carol	3	4	4,50
Andre	4	5	6,40
Diane	6	7	10,29
Fernando	5	6	8,33
Beverly	4	5	6,40
Ed	3	4	4,50
Garth	5	6	8,33
Heather	3	4	4,50
Ike	2	3	2,67
Jane	1	2	1,00





università degli studi FIRENZE

INGEGNERIA DELL'INFORMAZIONE

Clustering Coefficient, Ci()



of links to neighbors

max # links among neighbors

Cimax()=(N-1) / (N(N-1)/2) = 2/N

Ci()=2/5=0,40

Maximum for full connected.

A lower value means to have less connections to neighbors and thus the needs of clustering

MIT:	0.22
Stanford:	0.21



Distribuzione dei colleghi (facebook)







http://www.eclap.eu/drupal/?q=graphviewer

User Metrics





53.33 - 64

Connection

Group Connectivity

Groups Graph







Anonymity vs Relationships

- Given a certain user, the graph of his relationships with other users and action make it quite unique in the network.
- For this reason, SN have difficulties at providing data related to SN users even if names, email, location, etc. are removed.
- There is a business in de-anonymize the social network data for data intelligence and user behavior discovering
- Therefore, there are also algorithms for obfuscating the data for making the Anonymity process more effective without destroying the meaning of the data.





Part 6: Social Media Technologies and Solutions

2

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication





FIRENZE

INGEGNERIA DELL'INFORMAZIONE



Knowledge Management and Protection Systems (KMaPS): 2016-2017


UNIVERSITÀ

degli studi FIRENZE DINFO

INGEGNERIA DELL'INFORMAZIONE DISIT

Top Web in Italia

2015 in italia 2013 in italia 2014 in italia ➔ 1 Google.it 1 Google \rightarrow 1 Google.it → 2 facebook.com 2 facebook > 2 facebook.com → 3 Google.com 3 Google.com → 3 Google.com 4 YouTube.com 4 YouTube 4 YouTube.com 5 Yahoo 5 amazon.it 5 amazon.it 6 Wikipedia.org 6 Wikipedia 6 Yahoo.com 7 Yahoo.com 7 Libero 7 Libero.it -8 ebay 8 Wikipedia.org 8 ebay.it 9 Libero.it 9 ebay.it 9 repubblica 10 repubblica.it 10 Amazon 10 repubblica.it 11 Corriere.it 11 Corriere 11 subito.it 12 subito.it 12 Live.com 12 live.com 13 live.com 13 Linkedin.com 13 Corriere.it 14 twitter.com 14 Virgilio 14 virgilio.it 15 Linkedin.com 15 Linkedin.com

WORLD MAP OF SOCIAL NETWORKS June 2009



credits: Uincenzo Cosenza: www.uincos.it

license CC-BY-NC

sources: Google Trends for Websites/Alexa

WORLD MAP OF SOCIAL NETWORKS June 2013



Knowledge Management and Protection Systems (KMaPS): 2016-2017

WORLD MAP OF SOCIAL NETWORKS July 2014



WORLD MAP OF SOCIAL NETWORKS August 2015



WORLD MAP OF SOCIAL NETWORKS January 2016



Knowledge Management and Protection Systems (KMaPS): 2016-2017



source: SimilarWeb/Alexa



Estimated Unique Monthly Visitor

In 2012

- Facebook: 750,000,000
- Twitter: 250,000,000
- LinkedIn 110.000.000
- MySpace: 70.500.000
- GooglePlus 65.000.000
- DevianArt: 25.500.000
- LiveJournal: 20.500.000
- Tagged: 19.500.000
- Orkut: 17.500.000
- PinInterest: 15.500.000
- ÷ ...

INGEGNERIA DELL'INFORMAZIONE

UNIVERSITÀ Degli studi

FIRENZE

Badoo: 2.500.000

- 🤊 In 2014
 - 🐥 Google: 1.100 M
 - 🐥 YouTube: 1.000 M
 - 🐥 FaceBook: 900 M
 - 🐥 Yahoo: 750 M
 - 🐥 Amazon: 500 M
 - 🐥 Wikipedia: 475 M
 - 🐥 Twitter: 290 M
 - 🜲 Bing: 285 M
 - 🜲 eBay: 285 M
 - 🜲 MSN 280 M
 - 🐥 Microsoft: 270 M
 - 🐥 LinkedIn 260 M
 - 🐥 PinInterest: 250 M





Age distribution on social networks & online communities

Average based on the 24 sites included in this survey.





Sorted by average age, lowest at the top.

UN DE FI

Data source: DoubleClick Ad Planner (Google), U.S. demographics, June 2012.

Average age on social networks & online communities

Hacker News Orkut Github Reddit Tumblr Hi5 Tagged LiveJournal MySpace Stack Overflow Twitter Quora StumbleUpon Last.fm Slashdot Goodreads WordPress.com Blogger Pinterest Flickr Facebook Yelp LinkedIn

UNIVERSITÀ

DEGLI STUDI

FIRENZE

DINFO

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DISIT

DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB



Gender distribution on social networks & online communities

Male

Female

Pinterest Goodreads Blogger Tagged Tumblr MySpace Facebook Twitter WordPress.com Yelp Flickr Hi5 deviantART StumbleUpon LinkedIn LiveJournal Last.fm Quora Reddit Orkut Github Stack Overflow **Hacker News** Slashdot 0%

DEGLI STUDI

FIRENZE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE DELL'INFORMAZIONE DELL'INFORMAZIONE 10%

20%

Data source: DoubleClick Ad Planner (Google), U.S. demographics, June 2012.

30%

40%

Knowledge Management and Protection Systems (KMaPS): 2016-2017

50%

60%

70%

80%

90% 100%

www.pingdom.com



Part 6: Social Media Technologies and Solutions

ج

Collaborative systems

Definition and Terminology

Social Network

- Forrester Trend for Social Networking
- Motivations for Social Networking
- Application, classification of Social Networking
- Examples of Social Networks
- factors of Social Networks

User/Content Social Network

- User classification and behavior
- User Generated Content, UGC
- Content descriptors
- User and group descriptors

Measures of Social Networks

- User profile problems
- Measures of Social Networks
- Metrics and examples: Centrality, Clustering,
- Direct measures of user actions

Business of Social Networks

- Penetration of social networks
- Numbers of Social Networks

interoperability and standards

- Social icons
- Embedding
- Authentication



Interoperability among Social Networks

- SN may be interoperable with other portals and SN
 Allowing:
 - posting comments and contributions via the so called Social Icon interface
 - importing user registration/profile and info or directly with some SSO
 - exporting SN content in other portals, for example via some API.
 - hosting SN players into other WEB portal pages, via some HTML segment to be copied
 - hosting widgets/applications into the WEB pages of the Social Network, via some programming model





Interoperability and standards

- Interoperability for user profiles
 - migration/interchange
 - Authentication

Lack of standards to cope with these aspects
 Possible future standards coming from W3C





Posting Content via Social Icon

- Social Share/ Social Icon/
- A portal may:
 - expose some Social Icon



- Call the SN with a rest call with a set of information: image link, title, description, etc.
- If you are logged into the SN
 - the post is directly included asking you some information to complete the post: groups, comments, etc.
- If you are not logged into SN
 - A login dialog is presented to ask you log into the SN, then...





Interoperability for Users

- Interchange of user profiles
 - OpenID: user identity standard, to allow user profile and credential interoperability among portals, is SSO method
 - OAuth: delegation protocol for accessing to credentials, user authentication
 - OpenSocial (by Google with MySpace): exchange of user profile.
 - Many big Social Networks have joined the OpenSocial API movement, including hi5, LinkedIn, Netlog, Ning, Plaxo, Orkut, Friendster, Salesforce, Yahoo, Ning, SixApart, XING, etc.
 - Facebook Connect is in competition with OpenSocial
 - Other technologies:
 - XUP of W3C, XMPP, FOAF, XFN, etc..





Exporting SN content

- Most of the Social Networks allow to access their content and information via specific API,
 - 🐥 http or WS
 - YouTube and Flickr
- The interoperability API allow to:
 - Make queries (see Europeana)
 - Get metadata and statistical results about the number of plays
 - Get/post content (in some case)
 - make play from other remote pages to create related applications by using their content repositories and servers
 - Create other applications that can exploit their SN infrastructure
 - Get open data (see dbPedia, Europeana)





Hosting SN players

- Some of the SNs present their own API on the basis of which third party portals can access to their players.
- For example the players of YouTube can be embedded into third party web portals by taking the HTTP segment from the You Tube page
 - Esempio di Media Player Embedding from ECLAP
 - <iframe src='http://www.eclap.eu/drupal/?q=en-US/embed&axoid=urn%3Aaxmedis%3A00000%3Aobj%3A6e73a b5c-be18-4523-9ca5-916d5505e7fa' width='300' height='200' frameborder='0'></iframe>
- From YouTube

DINFO

INGEGNERIA DELL'INFORMAZIONE

JNIVERSITÀ Degli studi

FIRENZE

<iframe width="420" height="315" src="http://www.youtube.com/embed/f6rx6TLBXAY" frameborder="0" allowfullscreen></iframe>



Hosting Widget

- Twitter, Facebook and MySpace allow to create applications directly shown into their pages, we call them Widgets
- The users may select these applications to use and to be shown in their preferred pages.
 - The SN promote the widget to the user via several different kinds: I like, I share, etc.
- The Widgets have to be created by using a specific standard.
 - The widget are typically iFrames generated by a third party portal and embedded into the SN.
- Once the user install an application/widget, the SN ask to the user if the widget can access to its own information. The accessed information can be passed to the third party portal.







