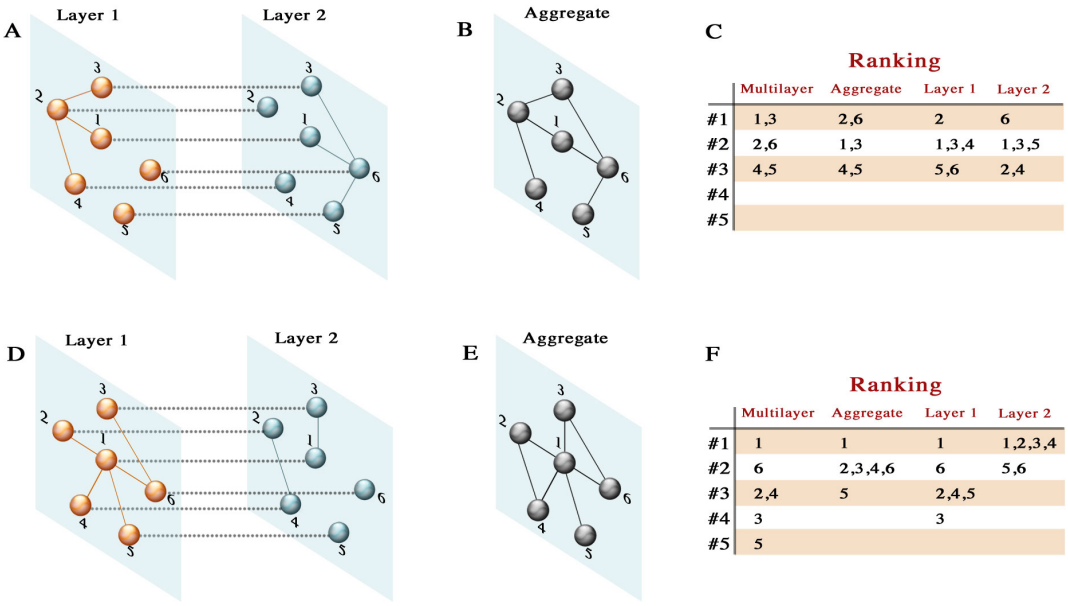
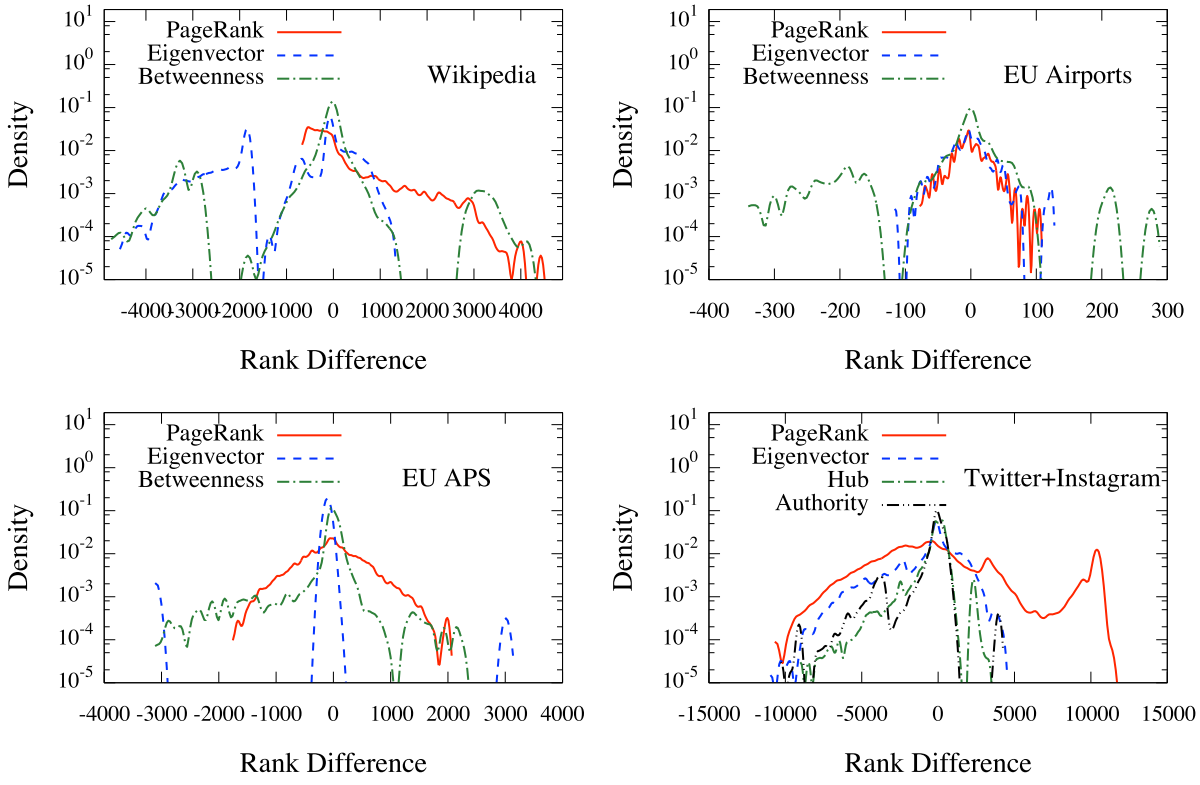
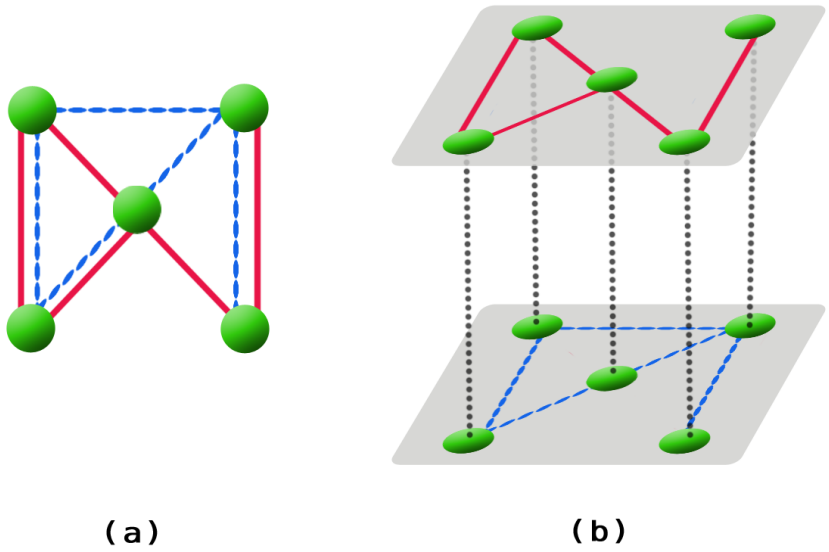


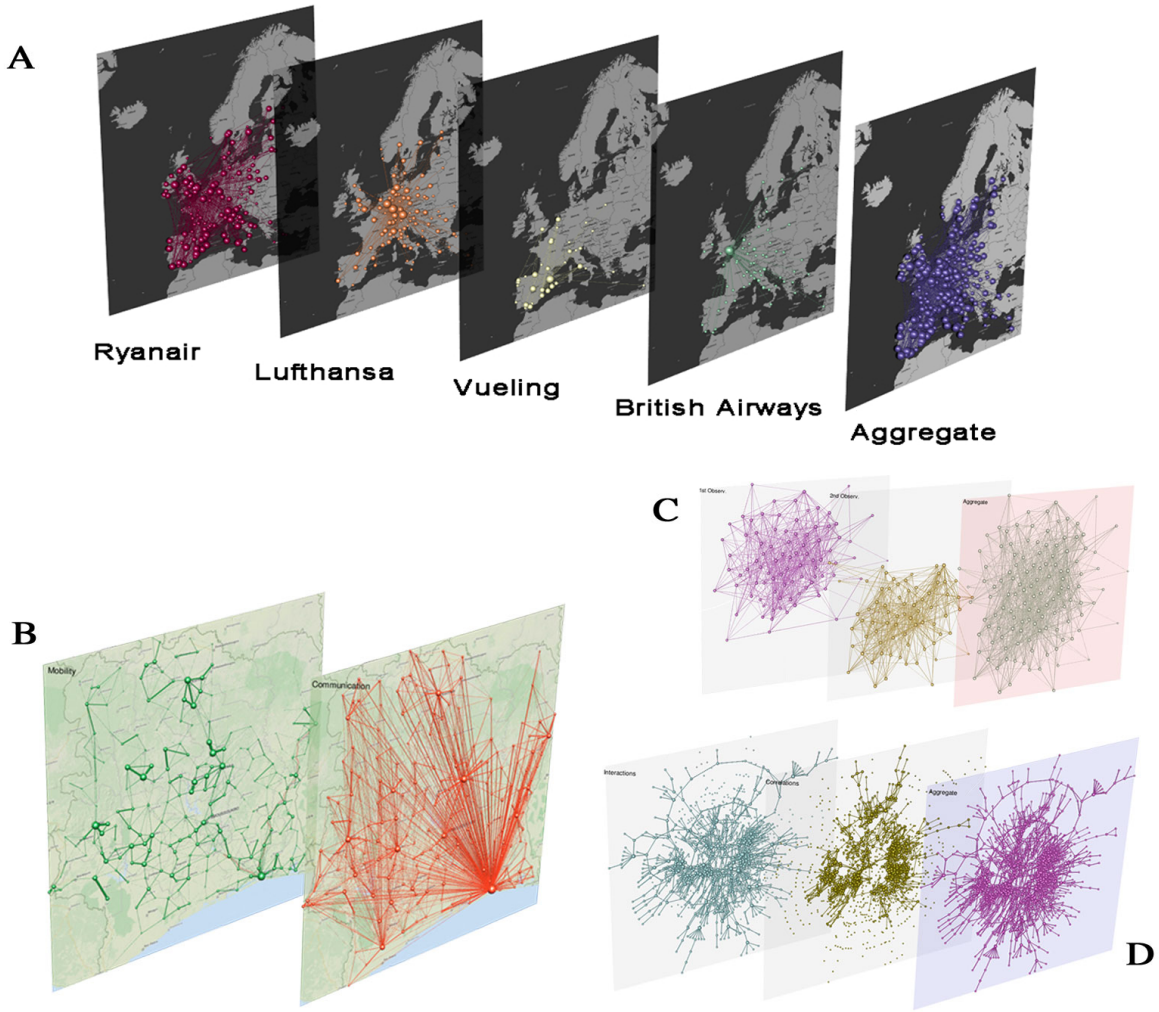
Supplementary Figure 1: Interconnected multiplex with six nodes in two layers (A and D) and corresponding aggregated networks (B and E). The nodes are ranked by their eigenvector centrality in each layer separately, in the aggregated and in the whole interconnected structure (C and F). **Case A, B and C.** Nodes 1 and 3 have a key role in the multilayer, being bridges between the two layers. In a collaboration network they would represent scientists working on two different research areas who allow information to flow from one subject to the other. While nodes 1 and 3 gain centrality from their connections to "hubs" on different layers, they also gain centrality from their own counterparts in other layers, making them important in the multilayer network. In the aggregated network their versatility disappears, because the information is washed out by projecting on a single layer, where nodes 2 and 6 are still "hubs" but it is not possible to capture the importance of nodes 1 and 3 in bridging different areas. **Case D, E and F.** This example shows how aggregating the full information on a single network introduces a spurious symmetry between nodes 2, 3, 4 and 6 that is not present in the multilayer, except for 2 and 4. The resulting score in the aggregate is not able to capture the difference between these nodes (corresponding to a degeneration in the eigenspace) while it is evident that, for instance, node 6 is more central than node 3 because of its direct connection to node 1 – the "hub" – in layer 1.



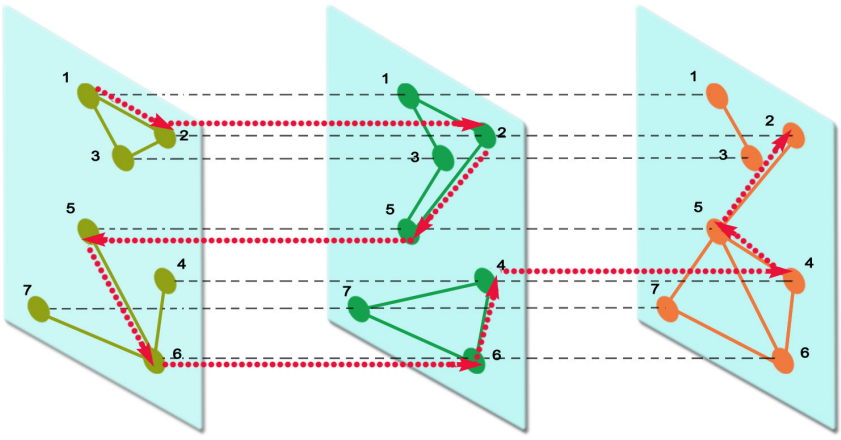
Supplementary Figure 2: Distribution of the difference between the node's rank, obtained from different centrality measures, in aggregated and multilayer network for the empirical datasets discussed in the text. If the rank difference would be negligible a distribution peaked around zero should be observed. We found useful to show HITS centrality only for Twitter+Instagram dataset, being a directed multilayer. The majority of authors in the APS dataset have eigenvector centrality not significantly different from zero, leading to the same ranking and explaining the small rank difference with the aggregated. Except for this case, the figures show that a negligible number of nodes has the same rank in the multilayer and the corresponding aggregated network in all cases.



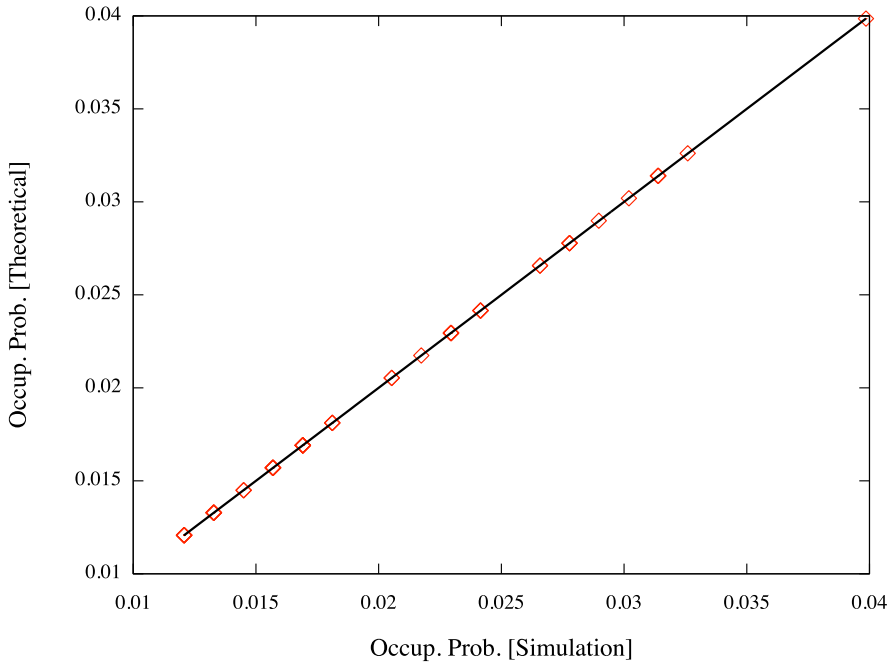
Supplementary Figure 3: a) Edge-colored graph (i.e., multiplex) representing two different types of interactions (solid and dashed edges) between 5 actors. b) An interconnected multiplex representing the same actors exhibiting the same relationships but on different levels which are separated by a cost (dotted vertical lines) to move from one layer to the other.



Supplementary Figure 4: Multilayered visualization of empirical interconnected multiplex networks, where interlayer connections are present but not shown for simplicity. (**A**) Layers correspond to flight routes operated by different air companies between European airports [1]. (**B**) Layers show the mobility and the communication network of sub-prefectures in Ivory Coast, built from mobile phone calls data (data provided during the Data for Development challenge, http://www.d4d.orange.com/). (**C**) Layers correspond to two different observations – separated by three weeks – of one ant colony [2]. (**D**) Layers correspond to the interaction network (left) of genes in *Saccharomyces cerevisiae*, obtained through synthetic genetic array methodology, and correlation-based network (middle) connecting genes with similar genetic interaction profiles [3]. To improve the visualization, we show in each layer the largest connected component of the corresponding network, where only pairs with high genetic interaction score and highly correlated genetic profiles are considered. The resultant aggregated network (right) is shown to highlight how structural information about the original layers is lost. Map tiles By Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under CC BY SA.



Supplementary Figure 5: Schematic of a random walk (dotted trajectories) in a multiplex network. A walker can jump between nodes within the same layer, or it might switch to another layer. This illustration evinces how multiplexity allows a random walker to move between nodes that belong to different (disconnected) components on a given layer.



Supplementary Figure 6: Representative example of the equivalence between the random walk occupation centrality obtained from Monte Carlo simulations and the theoretical prediction. The case of a multiplex with two layers and 50 nodes is considered.

		PageRank Centrality Ranking			Diversity	
Name	Versatility	Aggregate	Average	Consensus	Intra	Global
Edmund F. Robertson	1	1 (+0)	18 (+17)	3 (+2)	1	2
Milton Friedman	2	16 (+14)	22 (+20)	5 (+3)	1	8
Hilary Putnam	3	34 (+31)	1302 (+1299)	237 (+234)	1	4
E. O. Wilson	4	332 (+328)	996 (+992)	480 (+476)	1	8
Harold Clayton Urey	5	537 (+532)	451 (+446)	1244 (+1239)	1	8
Kurt Gödel	6	43 (+37)	325 (+319)	169 (+163)	3	8
Avicenna	7	30 (+23)	8 (+1)	41 (+34)	3	4
Ernst Mayr	8	191 (+183)	582 (+574)	68 (+60)	1	8
Herbert A. Simon	9	48 (+39)	14 (+5)	96 (+87)	3	8
Charles Stark Draper	10	1196 (+1186)	1169 (+1159)	970 (+960)	1	8
Ivan Pavlov	11	423 (+412)	56 (+45)	86 (+75)	2	6
Aristotle	12	3 (−9)	2 (−10)	19 (+7)	5	5
Paul Samuelson	13	26 (+13)	43 (+30)	15 (+2)	1	8
Immanuel Kant	14	2 (−12)	19 (+5)	4 (−10)	1	2
Norbert Wiener	15	68 (+53)	407 (+392)	245 (+230)	2	8
Chien-Shiung Wu	16	100 (+84)	599 (+583)	1315 (+1299)	1	6
George Dantzig	17	217 (+200)	1244 (+1227)	577 (+560)	2	8
Ronald Ross	18	1602 (+1584)	5472 (+5454)	5472 (+5454)	0	5
John C. Slater	19	1242 (+1223)	1627 (+1608)	2119 (+2100)	1	8
Porphyry (philosopher)	20	311 (+291)	1063 (+1043)	958 (+938)	1	3
Peter Mansfield	21	1502 (+1481)	2914 (+2893)	2682 (+2661)	1	5
Rosalyn Yalow	22	888 (+866)	1580 (+1558)	2088 (+2066)	1	7
Samuel Goudsmit	23	1364 (+1341)	1960 (+1937)	2361 (+2338)	1	8
Albert Einstein	24	10 (−14)	11 (−13)	1 (−23)	2	4
Plato	25	4 (−21)	15 (−10)	11 (−14)	2	2

Supplementary Table 1: Wikipedia dataset, ranking by PageRank.

		Eigenvector Centrality Ranking			Diversity	
Name	Versatility	Aggregate	Average	Consensus	Intra	Global
John Bardeen	1	1 (+0)	5 (+4)	1 (+0)	2	8
Enrico Fermi	2	2 (+0)	27 (+25)	9 (+7)	3	3
Luis Walter Alvarez	3	4 (+1)	90 (+87)	44 (+41)	1	6
Niels Bohr	4	3 (−1)	51 (+47)	18 (+14)	2	2
Isidor Isaac Rabi	5	7 (+2)	74 (+69)	27 (+22)	1	4
Hans Bethe	6	9 (+3)	77 (+71)	35 (+29)	1	1
Eugene Wigner	7	6 (−1)	96 (+89)	57 (+50)	2	6
Arthur Compton	8	13 (+5)	91 (+83)	51 (+43)	1	2
Yoichiro Nambu	9	5 (−4)	140 (+131)	151 (+142)	1	6
Nicolaas Bloembergen	10	8 (−2)	116 (+106)	84 (+74)	1	6
Albert Einstein	11	9 (−2)	25 (+14)	73 (+62)	2	4
Werner Heisenberg	12	14 (+2)	103 (+91)	68 (+56)	2	2
Albert Abraham Michelson	13	22 (+9)	120 (+107)	91 (+78)	1	3
James Chadwick	14	21 (+7)	124 (+110)	99 (+85)	1	3
Norman Foster Ramsey, Jr.	15	23 (+8)	130 (+115)	107 (+92)	1	3
Max Planck	16	15 (−1)	127 (+111)	115 (+99)	2	5
Paul Dirac	17	19 (+2)	110 (+93)	137 (+120)	2	2
Edward Mills Purcell	18	24 (+6)	135 (+117)	131 (+113)	1	4
Felix Bloch	19	26 (+7)	134 (+115)	124 (+105)	1	1
Chen Ning Yang	20	12 (−8)	223 (+203)	350 (+330)	1	6
J. J. Thomson	21	16 (−5)	94 (+73)	170 (+149)	3	5
Riccardo Giacconi	22	18 (−4)	187 (+165)	203 (+181)	1	6
Raymond Davis, Jr.	23	17 (−6)	122 (+99)	249 (+226)	2	7
John Robert Schrieffer	24	30 (+6)	153 (+129)	157 (+133)	1	1
John Hasbrouck Van Vleck	25	25 (+0)	158 (+133)	164 (+139)	2	3

Supplementary Table 2: Wikipedia dataset, ranking by eigenvector centrality.

		Betweenness Centrality Ranking			Diversity	
Name	Versatility	Aggregate	Average	Consensus	Intra	Global
Edmund F. Robertson	1	1 (+0)	1 (+0)	6 (+5)	1	2
Milton Friedman	2	5 (+3)	795 (+793)	133 (+131)	1	8
Albert Einstein	3	3 (+0)	102 (+99)	22 (+19)	2	4
Charles Darwin	4	7 (+3)	60 (+56)	4 (+0)	2	5
Immanuel Kant	5	9 (+4)	41 (+36)	44 (+39)	1	2
Aristotle	6	2 (−4)	4 (−2)	2 (−4)	5	5
Plato	7	10 (+3)	20 (+13)	11 (+4)	2	2
Bertrand Russell	8	8 (+0)	17 (+9)	77 (+69)	2	2
C. R. Rao	9	11 (+2)	27 (+18)	165 (+156)	1	6
Isaac Newton	10	4 (−6)	2 (−8)	1 (−9)	4	5
Paul Erdős	11	12 (+1)	6 (−5)	24 (+13)	1	1
Gottfried Wilhelm Leibniz	12	6 (−6)	8 (−4)	16 (+4)	3	3
Robert Solow	13	22 (+9)	584 (+571)	91 (+78)	1	6
Carl Friedrich Gauss	14	24 (+10)	21 (+7)	118 (+104)	2	6
Avicenna	15	26 (+11)	500 (+485)	429 (+414)	3	4
Marvin Minsky	16	33 (+17)	2396 (+2380)	328 (+312)	1	4
Kurt Gödel	17	13 (−4)	18 (+1)	101 (+84)	3	8
Benjamin Franklin	18	54 (+36)	459 (+441)	152 (+134)	3	4
Karl Marx	19	18 (−1)	35 (+16)	49 (+30)	2	3
Donald Knuth	20	16 (−4)	90 (+70)	58 (+38)	2	6
Norbert Wiener	21	19 (−2)	31 (+10)	193 (+172)	2	8
David Hume	22	23 (+1)	45 (+23)	57 (+35)	2	3
Karl Popper	23	51 (+28)	546 (+523)	654 (+631)	1	2
Kenneth Arrow	24	52 (+28)	1466 (+1442)	462 (+438)	1	6
Jerzy Neyman	25	31 (+6)	48 (+23)	230 (+205)	1	6

Supplementary Table 3: Wikipedia dataset, ranking by betweenness centrality.

	PageRank	
Name	Versatility	Aggregate
Ivan Pavlov	1	40 (+39)
Albert Einstein	2	1 (−1)
Hermann von Helmholtz	3	25 (+22)
Thomas Henry Huxley	4	14 (+10)
Charles Darwin	5	2 (−3)
David Hilbert	6	3 (−3)
Charles Wheatstone	7	43 (+36)
Felix Klein	8	5 (−3)
W. Thomson, 1st Baron Kelvin	9	24 (+15)
Karl Marx	10	4 (−6)
Karl Weierstrass	11	13 (+2)
Robert Bunsen	12	36 (+24)
Louis Agassiz	13	33 (+20)
Charles-Adolphe Wurtz	14	75 (+61)
Charles Sanders Peirce	15	8 (−7)
Marie Curie	16	20 (+4)
Lord Rayleigh	17	92 (+75)
William James	18	7 (−11)
Max Planck	19	11 (−8)
Friedrich Nietzsche	20	6 (−14)
Arthur Cayley	21	37 (+16)
James Prescott Joule	22	59 (+37)
Ludwig Boltzmann	23	79 (+56)
J. J. Thomson	24	17 (−7)
Karl Pearson	25	35 (+10)

Supplementary Table 4: Wikipedia alternative dataset, ranking by PageRank.

	Eigenvector	
Name	Versatility	Aggregate
Albert Einstein	1	1 (+0)
Max Planck	2	2 (+0)
J. J. Thomson	3	3 (+0)
Hendrik Lorentz	4	5 (+1)
William Henry Bragg	5	6 (+1)
Albert Abraham Michelson	6	7 (+1)
Marie Curie	7	4 (−3)
Charles Thomson Rees Wilson	8	9 (+1)
Pierre Curie	9	10 (+1)
Wilhelm Wien	10	13 (+3)
Max von Laue	11	17 (+6)
Ernest Rutherford	12	8 (−4)
Enrico Fermi	13	14 (+1)
Pieter Zeeman	14	23 (+9)
Henri Becquerel	15	16 (+1)
Walther Bothe	16	26 (+10)
Wilhelm Röntgen	17	15 (−2)
Philipp Lenard	18	28 (+10)
Johannes Stark	19	33 (+14)
Owen Willans Richardson	20	38 (+18)
Charles Glover Barkla	21	40 (+19)
Johannes Diderik van der Waals	22	31 (+9)
Jean Baptiste Perrin	23	41 (+18)
Ernest Lawrence	24	44 (+20)
Heike Kamerlingh Onnes	25	36 (+11)

Supplementary Table 5: Wikipedia alternative dataset, ranking by eigenvector centrality.

	PageRank	
Name	Versatility	Aggregate
Amsterdam Schiphol	1	1 (+0)
Brussels	2	17 (+15)
Charles de Gaulle	3	13 (+10)
Athens International	4	2 (−2)
Frankfurt	5	10 (+5)
Barcelona El Prat	6	4 (−2)
Madrid Barajas	7	3 (−4)
Munich	8	5 (−3)
Copenhagen	9	19 (+10)
Malpensa	10	24 (+14)
Vienna International	11	8 (−3)
Prague	12	21 (+9)
Istanbul Atatürk	13	7 (−6)
Stockholm Arlanda	14	22 (+8)
Oslo Gardermoen	15	11 (−4)
Warsaw Chopin	16	27 (+11)
Palma de Mallorca	17	16 (−1)
Cologne Bonn	18	28 (+10)
Fiumicino	19	12 (−7)
Budapest	20	26 (+6)
Bologna	21	54 (+33)
Zurich	22	18 (−4)
Gatwick	23	9 (−14)
London Stansted	24	6 (−18)
Heathrow	25	25 (+0)

Supplementary Table 6: EU Airports dataset, ranking by PageRank.

	Eigenvector	
Name	Versatility	Aggregate
Frankfurt	1	8 (+7)
Amsterdam Schiphol	2	1 (−1)
Brussels	3	11 (+8)
Charles de Gaulle	4	7 (+3)
Barcelona El Prat	5	5 (+0)
Malpensa	6	13 (+7)
Athens International	7	21 (+14)
Munich	8	6 (−2)
Madrid Barajas	9	2 (−7)
Copenhagen	10	12 (+2)
Vienna International	11	4 (−7)
Prague	12	19 (+7)
Fiumicino	13	3 (−10)
Palma de Mallorca	14	22 (+8)
Stockholm Arlanda	15	26 (+11)
Warsaw Chopin	16	17 (+1)
Budapest	17	16 (−1)
Oslo Gardermoen	18	24 (+6)
Zurich	19	10 (−9)
Heathrow	20	14 (−6)
Düsseldorf	21	9 (−12)
Manchester	22	35 (+13)
Hamburg	23	25 (+2)
Venice Marco Polo	24	29 (+5)
Lisbon Portela	25	18 (−7)

Supplementary Table 7: EU Airports dataset, ranking by eigenvector centrality

	Betweenness	
Name	Versatility	Aggregate
Athens International	1	2 (+1)
Oslo Gardermoen	2	1 (−1)
Amsterdam Schiphol	3	5 (+2)
London Stansted	4	4 (+0)
Tromsø	5	16 (+11)
Istanbul Atatürk	6	3 (−3)
Copenhagen	7	22 (+15)
Frankfurt	8	18 (+10)
Barcelona El Prat	9	7 (−2)
Brussels	10	19 (+9)
Madrid Barajas	11	11 (+0)
Charles de Gaulle	12	23 (+11)
Munich	13	9 (−4)
Stockholm Arlanda	14	10 (−4)
Helsinki	15	8 (−7)
Gatwick	16	6 (−10)
Vienna International	17	12 (−5)
Bodø	18	17 (−1)
Palma de Mallorca	19	20 (+1)
Prague	20	14 (−6)
Dublin	21	15 (−6)
Edinburgh	22	32 (+10)
Riga International	23	24 (+1)
Sabiha Gökçen International	24	25 (+1)
Bergamo Orio al Serio	25	34 (+9)

Supplementary Table 8: EU Airports dataset, ranking by betweenness centrality.

	PageRank	
Name	Versatility	Aggregate
Rosario Fazio	1	1 (+0)
F. M. Peeters	2	2 (+0)
Peter Hänggi	3	5 (+2)
H. T. C. Stoof	4	21 (+17)
Klaus Mølmer	5	6 (+1)
Daniel Loss	6	8 (+2)
A. Loidl	7	3 (−4)
E. V. Chulkov	8	4 (−4)
Christoph H. Keitel	9	38 (+29)
Matthias Scheffler	10	11 (+1)
Erwin Frey	11	53 (+42)
Amand Faessler	12	9 (−3)
Francesco Mauri	13	14 (+1)
Nicolas Gisin	14	12 (−2)
Klaus Richter	15	20 (+5)
Gerard Meijer	16	30 (+14)
Andrzej M. Oleś	17	7 (−10)
Ingrid Mertig	18	10 (−8)
Christoph Simon	19	23 (+4)
J. Ignacio Cirac	20	48 (+28)
Michele Lazzeri	21	32 (+11)
Ralf Stannarius	22	70 (+48)
Tilman Pfau	23	121 (+98)
Francesc Sagués	24	43 (+19)
Marjolein Dijkstra	25	28 (+3)

Supplementary Table 9: EU APS dataset, ranking by PageRank.

	Eigenvector	
Name	Versatility	Aggregate
F. M. Peeters	1	1 (+0)
M. V. Milošević	2	3 (+1)
B. Partoens	3	2 (−1)
G. R. Berdiyorov	4	4 (+0)
S. Bednarek	5	5 (+0)
V. R. Misko	6	7 (+1)
Y. Sidor	7	8 (+1)
M. D. Croitoru	8	10 (+2)
A. A. Shanenko	9	12 (+3)
I. V. Grigorieva	10	15 (+5)
T. Chwiej	11	13 (+2)
J. Adamowski	12	16 (+4)
M. Hayne	13	17 (+4)
Péter Földi	14	6 (−8)
Mihály G. Benedict	15	9 (−6)
A. K. Geim	16	19 (+3)
Orsolya Kálmán	17	14 (−3)
Egidijus Anisimovas	18	20 (+2)
A. Vagov	19	18 (−1)
A. Potenza	20	21 (+1)
S. J. Bending	21	22 (+1)
S. W. S. Apolinario	22	11 (−11)
L. A. Ponomarenko	23	24 (+1)
F. Schedin	24	25 (+1)
M. Henini	25	26 (+1)

Supplementary Table 10: EU APS dataset, ranking by eigenvector centrality.

	Betweenness	
Name	Versatility	Aggregate
Hua Wu	1	1 (+0)
Rosario Fazio	2	3 (+1)
Matthias Scheffler	3	2 (−1)
Daniel Loss	4	15 (+11)
Matthias Troyer	5	7 (+2)
Sander van Smaalen	6	5 (−1)
J. A. Mydosh	7	4 (−3)
Georg Kresse	8	12 (+4)
Ingrid Mertig	9	11 (+2)
Reinhard K. Kremer	10	6 (−4)
Oleg V. Dolgov	11	8 (−3)
Miodrag L. Kulić	12	10 (−2)
Jürgen König	13	28 (+15)
J. Ferré	14	13 (−1)
Peter Kratzer	15	9 (−6)
K. Held	16	20 (+4)
H. Bouchiat	17	33 (+16)
F. F. Assaad	18	23 (+5)
G. Faini	19	25 (+6)
Giovanna Morigi	20	32 (+12)
Gerd Schön	21	43 (+22)
Armen Sedrakian	22	18 (−4)
J. C. Cuevas	23	47 (+24)
Ulrich Hohenester	24	52 (+28)
Antti-Pekka Jauho	25	55 (+30)

Supplementary Table 11: EU APS dataset, ranking by betweenness centrality.

	PageRank	
Name	Versatility	Aggregate
Matthias Scheffler	1	1 (+0)
K. H. Bennemann	2	15 (+13)
Robert Graham	3	46 (+43)
M. A. Liberman	4	305 (+301)
Erio Tosatti	5	4 (−1)
Hans Ågren	6	2 (−4)
H. Winter	7	23 (+16)
G. Landwehr	8	8 (+0)
D. A. Ritchie	9	5 (−4)
Jakub Zakrzewski	10	7 (−3)
R. J. Needs	11	33 (+22)
Andrzej M. Oleś	12	3 (−9)
F. Perrot	13	318 (+305)
N. H. March	14	131 (+117)
J. Zhang	15	393 (+378)
C. W. J. Beenakker	16	41 (+25)
C. T. Foxon	17	9 (−8)
H. R. Ott	18	138 (+120)
P. Weinberger	19	11 (−8)
M. Müller	20	235 (+215)
L. Eaves	21	10 (−11)
Giulio Casati	22	54 (+32)
Daan Frenkel	23	38 (+15)
Michele Parrinello	24	27 (+3)
M. Pepper	25	6 (−19)

Supplementary Table 12: EU APS alternative dataset, ranking by PageRank.

	Eigenvector	
Name	Versatility	Aggregate
D. A. Ritchie	1	1 (+0)
M. Pepper	2	2 (+0)
G. A. C. Jones	3	3 (+0)
J. E. F. Frost	4	4 (+0)
J. T. Nicholls	5	5 (+0)
C. G. Smith	6	6 (+0)
H. P. Hughes	7	7 (+0)
A. J. Shields	8	8 (+0)
D. M. Whittaker	9	9 (+0)
G. Hill	10	10 (+0)
Karsten Flensberg	11	11 (+0)
J. A. A. J. Perenboom	12	12 (+0)
M. S. Skolnick	13	13 (+0)
P. Wyder	14	14 (+0)
R. T. Phillips	15	15 (+0)
J. S. Roberts	16	16 (+0)
P. C. Klipstein	17	19 (+2)
L. Eaves	18	17 (−1)
M. Henini	19	18 (−1)
G. W. Smith	20	20 (+0)
D. J. Mowbray	21	21 (+0)
P. C. Main	22	22 (+0)
P. H. Beton	23	23 (+0)
A. G. M. Jansen	24	24 (+0)
J. C. Portal	25	27 (+2)

Supplementary Table 13: EU APS alternative dataset, ranking by eigenvector centrality.

	PageRank	
Name	Versatility	Aggregate
mashable	1	1 (+0)
BreakingNews	2	3 (+1)
nerdist	3	8 (+5)
robdelaney	4	9 (+5)
feliciaday	5	7 (+2)
britneyspears	6	4 (−2)
VictoriaJustice	7	574 (+567)
Veronica	8	25 (+17)
tyrabanks	9	24 (+15)
USATODAY	10	33 (+23)
AntDeRosa	11	88 (+77)
ijustine	12	21 (+9)
chrisbrogan	13	14 (+1)
problogger	14	19 (+5)
cocorocha	15	82 (+67)
anildash	16	57 (+41)
davemcclure	17	23 (+6)
jowyang	18	17 (−1)
JoshuaDavis	19	488 (+469)
redbull	20	5 (−15)
benparr	21	48 (+27)
arrington	22	39 (+17)
chrismessina	23	71 (+48)
caro	24	171 (+147)
chrisconnolly	25	189 (+164)

Supplementary Table 14: Twitter+Instagram dataset, ranking by PageRank.

	Eigenvector	
Name	Versatility	Aggregate
mashable	1	9 (+8)
BarrettAll	2	7 (+5)
JonahLupton	3	1 (−2)
MarshaCollier	4	3 (−1)
cspenn	5	5 (+0)
mayhemstudios	6	2 (−4)
JessicaNorthey	7	8 (+1)
awakeningaimee	8	4 (−4)
chrisbrogan	9	39 (+30)
BarbaraDuke	10	10 (+0)
Garin	11	13 (+2)
bsainsbury	12	17 (+5)
AnthonyGemma	13	6 (−7)
jkcallas	14	16 (+2)
Britopian	15	14 (−1)
jowyang	16	34 (+18)
problogger	17	62 (+45)
AskAaronLee	18	19 (+1)
bryankramer	19	32 (+13)
cc_chapman	20	15 (−5)
BillHibbler	21	37 (+16)
CharityIdeas	22	26 (+4)
BobWarren	23	21 (−2)
chrisguillebeau	24	60 (+36)
BrettGreene	25	35 (+10)

Supplementary Table 15: Twitter+Instagram dataset, ranking by eigenvector centrality.

		Authority
Name	Versatility	Aggregate
mashable	1	1 (+0)
chrisbrogan	2	3 (+1)
JonahLupton	3	8 (+5)
BarrettAll	4	9 (+5)
cspenn	5	4 (−1)
MarshaCollier	6	6 (+0)
JessicaNorthey	7	7 (+0)
mayhemstudios	8	2 (−6)
awakeningaimee	9	15 (+6)
problogger	10	5 (−5)
jowyang	11	10 (−1)
Garin	12	16 (+4)
BreakingNews	13	12 (−1)
bsainsbury	14	30 (+16)
BarbaraDuke	15	26 (+11)
jkcallas	16	18 (+2)
Britopian	17	14 (−3)
AnthonyGemma	18	23 (+5)
AskAaronLee	19	13 (−6)
bryankramer	20	37 (+17)
cc_chapman	21	11 (−10)
BrettGreene	22	22 (+0)
chrisguillebeau	23	24 (+1)
BillHibbler	24	53 (+29)
CHRISVOSS	25	17 (−8)

Supplementary Table 16: Twitter+Instagram dataset, ranking by authority centrality.

	Hub	
Name	Versatility	Aggregate
JonahLupton	1	1 (+0)
AnthonyGemma	2	3 (+1)
awakeningaimee	3	5 (+2)
MarshaCollier	4	4 (+0)
BarbaraDuke	5	7 (+2)
mayhemstudios	6	2 (−4)
bsainsbury	7	13 (+6)
cspenn	8	12 (+4)
BarrettAll	9	10 (+1)
carece	10	6 (−4)
Garin	11	17 (+6)
BrandYou	12	19 (+7)
7onashoestring	13	9 (−4)
jkcallas	14	18 (+4)
JessicaNorthey	15	16 (+1)
cammipham	16	14 (−2)
CHRISVOSS	17	11 (−6)
BillHibbler	18	39 (+21)
Britopian	18	20 (+2)
AskAaronLee	20	30 (+10)
CharityIdeas	21	40 (+19)
DashBurst	22	35 (+13)
cc_chapman	23	15 (−8)
DvinMsM	24	8 (−16)
cfleury	25	24 (−1)

Supplementary Table 17: Twitter+Instagram dataset, ranking by hub centrality.

Supplementary Note 1. Multilayer, multiplex and interconnected networks

It is important to discuss the difference between the topological structure which represents the core of this study, namely interconnected multilayer networks [4, 5, 6, 7, 8, 9, 10], and other multilayer structures which have been named *multiplexes* in the past and have been the subject of recent studies [11, 12, 13, 14, 15]. Note that interconnected multilayer networks are not simply a special case of or equivalent to interdependent networks [16]: in multilayer systems, many or even all of the nodes have a counterpart in each layer, so one can associate a vector of states to each node. This feature has no counterpart in interdependent networks, which were conceived as interconnected communities within a single, larger network [17, 18].

Historically, the term *multiplex* has been adopted to indicate the presence of more than one relationship between the same actors of a social network [19]. This type of network is well understood in terms of "coloring" (or labeling) the edges corresponding to interactions of different nature. For instance, the same individual might have connections with other individuals based on financial interests (e.g., color red) and connections with the same or different individuals based on friendship (e.g., color blue). This type of network is represented by a *non-interconnected multiplex*.

Conversely, in other real-world systems, like the transportation network of a city, the same geographical position can be part, for instance, of the network of subway or the network of bus routes, simultaneously. In this specific case, an edge-colored graph would not capture the full structure of the network, because it is missing information about the cost to *move* from the subway network to the bus route. This cost can be economic or might account for the time required to physically commute between the two layers. Therefore, the interconnected multilayer topology presented in this section provides a better representation of the system. In Supplementary Fig. **3** is shown an illustration of an edge-colored graph (Supplementary Fig. **3**a) and an interconnected multiplex (Supplementary Fig. **3**b). It is evident that a simple projection of the latter – mathematically equivalent to sum up the corresponding adjacency matrices – would provide a network where the information about the colors is lost. On the other hand, an edge-colored graph can not account for interconnections, keeping irreconcilable the two structures in Supplementary Fig. **3** which should be used to represent very different networked systems.

For further details about the classification of such multilayer networks we refer to [20] and references therein.

A real-world example of a multiplex network is provided by transportation network of a city, where the same geographical position can be part, for instance, of the network of subway or the network of bus routes, simultaneously. We show in Supplementary Fig. **4**A the case of flight routes operated by different air companies

between European airports. For instance, layers encode the human mobility and the mobile communication network of different geographical areas (Supplementary Fig. **4**B) or physical contacts over time between ants in a colony (Supplementary Fig. **4**C). In other real-world systems, like genetic networks, two genes might exhibit different interactions (e.g., allelic or non-allelic), or be related because of their chemical interaction or their functional role (Supplementary Fig. **4**D).

Supplementary Note 2. Tensorial notation

Edge-colored graphs can be represented by a set of adjacency matrices [1, 12, 13, 14]. However, standard matrices, used to represent networks, are inherently limited in the complexity of the relationships that they can capture, i.e., they do not represent a suitable framework in the case of interconnected multiplexes. This is the case of increasingly complicated types of relationships – that can also change in time – between nodes. Such a level of complexity can be characterized by considering tensors and algebras of higher order [7].

A great advantage of tensor formalism also relies on its compactness. An adjacency tensor can be written using a more compact notation that is very useful for the generalization to *multilayer* networks. In this notation, a row vector $\mathbf{a} \in R^N$ is given by a covariant vector a_α ($\alpha = 1, \ldots, N$), and the corresponding contravariant vector a^α (i.e., its dual vector) is a column vector in Euclidean space. A canonical vector is assigned to each node and the corresponding interconnected multilayer network is represented by a rank-4 adjacency tensor.

However, in the majority of applications, it is not necessary to perform calculations using canonical vectors and tensors explicitly. Consequently, a classical single-layer network represented by a rank-2 mixed adjacency tensor W^α_β [7] can be simply indicated by W^i_j, where the "abuse of notation" consists in interpreting the indices i and j as nodes and W^i_j would indicate intensity of the relationship between them. Hence, W^i_j represents the well-known adjacency matrix of a graph and the classical notation for the weight w_{ij} of the link between i and j corresponds to W^i_j. The "abuse of notation" also consists in treating W^i_j as a rank-2 tensor, although it explicitly indicates the entry of a matrix, while keeping the algebraic rules governing covariant and contravariant tensors. This "abuse of notation" dramatically reduces the complexity of some tensorial equations, although it is worth remarking that it should be used only when calculations do not involve canonical tensors explicitly.

To distinguish simple networks from the more complicated situations (e.g., interconnected multiplex networks) that we use in this paper, we will use the term *monoplex networks* to describe such standard networks, which are time-independent and possess only a single type of edge that connects its nodes.

In general, there might be several types of relationships between pairs of nodes and a more general system represented as a multilayer object – in which each type of relationship is encompassed in a single *layer* α ($\alpha = 1, 2, \ldots, L$) of a system – is required. Note that α has no more the same meaning of the index in the adjacency tensor discussed above. To avoid confusion, in the following we refer to nodes with Latin letters and to layers with Greek letters, allowing us to distinguish indices that correspond to nodes from those that correspond to layers in tensorial equations.

We use an *intra-layer adjacency tensor* for the 2^{nd}-order tensor $W_j^i(\alpha)$ that indicates the relationships between nodes within the *same* layer α. We take into account the possibility that a node i from layer α can be connected to any other node j in any other layer β. To encode information about relationships that incorporate multiple layers, we introduce the 2^{nd}-order *inter-layer adjacency tensor* $C_j^i(\alpha\beta)$. Note that $C_j^i(\alpha\alpha) = W_j^i(\alpha)$.

It has been shown that the mathematical object accounting for the whole interconnected multilayer structure is given by a 4^{th}-order (i.e., rank-4) *multilayer adjacency tensor* $M_{j\beta}^{i\alpha}$. This tensor might be simply thought as a higher-order matrix with four indices. It is the direct generalization of the adjacency matrix in the case of monoplexes, encoding the intensity of the relationship (which may not be symmetric) between a node i in layer α and a node j in layer β [7]. This object is very general and can be used to represent structures where an actor is present in some layers but not in all of them. This is the case, for instance when considering a network of online social relationships, of an individual with an account on Facebook but not on Twitter. The algebra still holds for these situations without any formal modification. In fact, one simply introduces "empty nodes" and assigns the value 0 to the associated edges, although the calculations of network diagnostics should carefully account for the presence of such nodes (for instance, for a proper normalization) [7].

Often, to reduce the notational complexity in the tensorial equations, the Einstein summation convention is adopted. It is applied to repeated indices in operations that involve tensors. For example, we use this convention in the left-hand sides of the following equations:

$$A_i^i = \sum_{i=1}^{N} A_i^i \,, \tag{1}$$

$$A_j^i B_i^j = \sum_{i=1}^{N}\sum_{j=1}^{N} A_j^i B_i^j \,, \tag{2}$$

$$A_{j\beta}^{i\alpha} B_{i\gamma}^{k\beta} = \sum_{i=1}^{N}\sum_{\beta=1}^{L} A_{j\beta}^{i\alpha} B_{i\gamma}^{k\beta} \,, \tag{3}$$

whose right-hand sides include the summation signs explicitly. It is straightforward to use this convention for the product of any number of tensors of any order. In the following, we will use the t-th power of rank-4 tensors, defined by multiple tensor multiplications:

$$(A^t)_{j\beta}^{i\alpha} = (A)_{j_1\beta_1}^{i\alpha} (A)_{j_2\beta_2}^{j_1\beta_1} \dots (A)_{j\beta}^{j_{t-1}\beta_{t-1}} \tag{4}$$

Repeated indices, such that one index is a subscript and the other is a superscript, is equivalent to perform a tensorial operation known as a *contraction*. Moreover, one should be very careful in performing tensorial calculations. For instance, using traditional notation the product $a^i b^j$ would be a number, i.e., the product of the components of two vectors. However, in our formulation, the same calculation denotes a Kronecker product

between two vectors, resulting in a rank-2 tensor, i.e., a matrix.

An interesting network that can be derived from the interconnected structure is the aggregated network, where the edges between two actors are summed up across all layers. The superposition of the different layers is equivalent to summing up the adjacency tensor of each layer. The corresponding aggregated network G_j^i is a monoplex and is obtained by contracting the layer indices of the multilayer adjacency tensor, i.e., $G_j^i = M_{j\alpha}^{i\alpha}$. This aggregation loses the information about inter-layer connections. If such an information is important for the application of interest, then the tensor should be contracted with the 1-tensor u_α^β (the rank-2 tensor with all components equal to 1), i.e., $\bar{G}_j^i = M_{j\beta}^{i\alpha} u_\alpha^\beta$.

This formalism is extremely useful to put in evidence how topological descriptors of interconnected networks differ from the ones corresponding to their aggregated graphs [7, 21]. Moreover, it is particularly suitable to perform compact calculations.

As a representative example, let us consider the number of paths of length 2 from a node in a certain layer to any other node in any other layer of the system. Taking advantage of the extended algebra, it is straightforward to show that the resulting rank-4 tensor accounting for such paths is given by $H_{j\beta}^{i\alpha} = M_{k\gamma}^{i\alpha} M_{j\beta}^{k\gamma}$. If only the number of paths between any pair of nodes is required, regardless of the layer, then the corresponding rank-2 tensor of paths is simply obtained by contracting with the 1-tensor u_α^β, i.e., $X_j^i = H_{j\beta}^{i\alpha} u_\alpha^\beta$. Conversely, in the case of the aggregate, we first contract the multilayer adjacency tensor to obtain the aggregation $J_j^i = M_{j\beta}^{i\alpha} u_\alpha^\beta$, where inter-layer connections are included as self-loops, and then square the resulting tensor to obtain $Y_j^i = J_k^i J_j^k$. Of course, a similar argument can be used to calculate the number of longer paths. From these tensorial equations it is evident that the aggregated graph can not be considered, in general, a good proxy of the interconnected topology.

Summarizing, the tensorial formulation provides a suitable framework for several real-world networked systems, from transportation networks to social ones. It is also worth noting that special cases of multilayer adjacency tensors are time-dependent (i.e., "temporal") networks [7, 20]. More specifically, in the case of social sciences the multilayer adjacency tensor can be used, for instance, to model the structural changes of a social network over time, or to define the topology of actors involved in several different levels of relationships and for whom it is indispensable to define an inter-connection between such levels. For these networked systems, it is desirable to adopt descriptors (e.g., clustering coefficient, modularity, *etc*) that are the natural extension of their well-known counterparts in monoplex networks.

Supplementary Note 2.1. On the tensorial nature of adjacency tensors

Although we have already shown in Ref. [7] the advantages of using the tensor formalism to deal with multilayer networks, the assignment of the indices as covariant or contravariant may seem arbitrary. The problem arises from the absence of natural basis transformations which could guide us in this decision. The idea is that, if we perform a change of basis governed by a matrix Q^{α}_{β}, each contravariant index of any tensor is transformed using Q, while covariant indices change with Q^{-1}, the inverse of Q. Thus, an object with three indices which transforms with two Q and one Q^{-1} is bounded to be 1–covariant and 2–contravariant. However, these transformation are not the origin but the consequence of the "meaning" of the object. For example, inner products, metric tensors and symplectic forms must be 2–covariant since they are bilinear functions which assign two vectors to a number, while linear transformations are 1–covariant and 1–contravariant because they have to convert a vector (or 1–form) in another vector (or 1–form).

In the case of monoplex networks, the adjacency tensor may be viewed as a linear transformation which, given a vector (or 1–form) representing a node, returns the set of their adjacent nodes. Thus, the only acceptable representation for the monoplex adjacency object is a 1–covariant and 1–contravariant tensor. Likewise, the multilayer adjacency tensor transforms a node in one layer into the set of adjacent nodes, keeping also the information of which layer they belong to, thus a 2–covariant and 2–contravariant tensor is needed.

Once we know the order of the adjacency tensor, its transformation under a change of coordinates is completely determined. First we show how this works for a single-layer network and, afterwards, for a full multilayer network.

By following Ref. [7], the adjacency tensor W^{α}_{β} of a network can be represented as a linear combination of tensors of the canonical basis by

$$W^{\alpha}_{\beta} = \sum_{i,j=1}^{N} w_{ij} e^{\alpha}(i) e_{\beta}(j) = \sum_{i,j=1}^{N} w_{ij} E^{\alpha}_{\beta}(ij), \tag{5}$$

where $E^{\alpha}_{\beta}(ij) \in \mathbb{R}^{N\times N}$ indicates the tensor of the canonical basis corresponding to the tensorial product of the canonical vectors $\mathbf{e}(i)$ and $\mathbf{e}^{\dagger}(j)$ (defined in $\mathbb{R}^N$) assigned to nodes i ($e^{\alpha}(i)$) and j ($e_{\beta}(j)$), respectively. Let

$$Q^{\alpha}_{\beta} = \sum_{i=1}^{N} e'^{\alpha}(i) e_{\beta}(i) \tag{6}$$

be the change of basis tensor which transforms the basis vector set $\{e^{\alpha}(i), i = 1, \ldots, N\}$ into a second set $\{e'^{\alpha}(i), i = 1, \ldots, N\}$. Here, Q^{α}_{β} is expressed in terms of the basis vectors from both bases and it is straightforward to show that $e'^{\alpha}(i) = Q^{\alpha}_{\beta} e^{\beta}(i)$ and $e'_{\beta}(i) = e_{\alpha}(j)(Q^{-1})^{\alpha}_{\beta}$. By remarking that a change of basis should not

affect the intensity of the relationship between nodes n_i and n_j, by following the above prescription, we obtain:

$$\begin{aligned} W'^{\gamma}_{\delta} &= \sum_{i,j=1}^{N} w_{ij} e'^{\gamma}(i) e'_{\delta}(j) = \sum_{i,j=1}^{N} w_{ij} Q^{\gamma}_{\alpha} e^{\alpha}(i) e_{\beta}(j) (Q^{-1})^{\beta}_{\delta} \\ &= Q^{\gamma}_{\alpha} \left[\sum_{i,j=1}^{N} w_{ij} e^{\alpha}(i) e_{\beta}(j) \right] (Q^{-1})^{\beta}_{\delta} = Q^{\gamma}_{\alpha} W^{\alpha}_{\beta} (Q^{-1})^{\beta}_{\delta}, \end{aligned} \tag{7}$$

providing the desired tensor transformation law.

In the following we use the same notation of Ref. [7] to avoid confusion. In the same spirit, we introduce the vectors $e^{\tilde{\gamma}}(k)$ $(\tilde{\gamma}, k = 1, \ldots, L)$ of the canonical basis in the space $\mathbb{R}^L$, where the greek index indicates the components of the vector and the latin index indicates the k-th canonical vector. Therefore, it is straightforward to build the 2^{nd}-order tensors $E^{\tilde{\gamma}}_{\tilde{\delta}}(hk) = e^{\tilde{\gamma}}(h) e_{\tilde{\delta}}(k)$ representing the canonical basis of the space $\mathbb{R}^{L \times L}$.

The representation of the multilayer object $M^{\alpha\tilde{\gamma}}_{\beta\tilde{\delta}}$ in terms of the Kronecker product of canonical vectors is given by [7]

$$M^{\alpha\tilde{\gamma}}_{\beta\tilde{\delta}} = \sum_{i,j=1}^{N} \sum_{h,k=1}^{L} w_{ij}(hk) e^{\alpha}(i) e_{\beta}(j) e^{\tilde{\gamma}}(h) e_{\tilde{\delta}}(k). \tag{8}$$

Proceeding as in the case of a single-layer network, we obtain

$$\begin{aligned} M'^{\alpha\tilde{\gamma}}_{\beta\tilde{\delta}} &= \sum_{i,j=1}^{N} \sum_{h,k=1}^{L} w_{ij}(hk) Q^{\alpha}_{\rho} e^{\rho}(i) (Q^{-1})^{\sigma}_{\beta} e_{\sigma}(j) \tilde{Q}^{\tilde{\gamma}}_{\tilde{\phi}} e^{\tilde{\phi}}(h) (\tilde{Q}^{-1})^{\tilde{\epsilon}}_{\tilde{\delta}} e_{\tilde{\epsilon}}(k) \\ &= Q^{\alpha}_{\rho} \tilde{Q}^{\tilde{\gamma}}_{\tilde{\phi}} M^{\rho\tilde{\phi}}_{\sigma\tilde{\epsilon}} (Q^{-1})^{\sigma}_{\beta} (\tilde{Q}^{-1})^{\tilde{\epsilon}}_{\tilde{\delta}}, \end{aligned} \tag{9}$$

providing the desired transformation law of the multilayer adjacency tensor under a change of coordinates.

Supplementary Note 3. Centrality in Interconnected Networks

In this section, we focus on the definition of node centrality in a multilayer network. We obtain these properties using algebraic operations involving the multilayer adjacency tensor, canonical vectors, and canonical tensors, achieving the natural extension of the concept of centrality in single-layer networks. We refer to [7] for other multilayer network diagnostics.

In practical applications one is often interested in assigning a global measure of importance to each node, aggregating the information obtained from the different layers. A naive choice could be to combine the centrality

of the nodes – obtained from the different layers separately – according to some heuristic choice. This is a viable solution when there is no interconnection between layers, i.e., in the case of edge-colored graphs [15, 22]. However, the main drawback of applying this approach to interconnected multilayer networks is that the measure will depend on the choice of the heuristics and might not evaluate the real importance of nodes. Conversely, our approach capitalizes on the tensorial formulation of interconnected multilayer networks and accounts for the higher level of complexity of such systems without relying on external assumptions and naturally extending the well-known centrality measures adopted for several decades in the case of monoplexes.

Random walk occupation centrality. A random walk is the simplest dynamical process that can occur on a monoplex network, and random walks can be used to approximate other types of diffusion [23, 24]. Random walks on monoplex networks [23, 25, 24] have attracted considerable interest because they are both important and easy to interpret. They have yielded important insights on a huge variety of applications and can be studied analytically. For example, random walks have been used to rank Web pages [26] and sports teams [27], optimize searches [28], investigate the efficiency of network navigation [29, 30], characterize cyclic structures in networks [31], and coarse-grain networks to illuminate meso-scale features such as community structure [32, 33, 34]. Another interesting application of random walks is to calculate the centrality of actors in complex networks when they have not knowledge of the full topology but only local information is available. In such cases, centrality descriptors based on shortest-paths, e.g., betweenness and closeness centrality, should be substituted by centrality notions based on random walks [25, 35].

In this paper, we consider a discrete-time random walk. As we illustrate in Supplementary Fig. **5**, a random walk on a multilayer network induces nontrivial effects because the presence of inter-layer connections affects its navigation of a networked system [36]. Let $T^{i\alpha}_{j\beta}$ denote the tensor of transition probabilities for jumping between pairs of nodes and switching between pairs of layers, and let $p_{i\alpha}(t)$ be the time-dependent tensor that gives the probability to find a walker at a particular node in a particular layer. Hence, the covariant master equation that governs the discrete-time evolution of the probability from time t to time $t+1$ is $p_{j\beta}(t+1) = T^{i\alpha}_{j\beta} p_{i\alpha}(t)$.

The steady-state solution of this equation, i.e., for $t \longrightarrow \infty$, is given by $\Pi_{i\alpha}$, quantifying the probability to find a walker in the node i of layer α. In the case of monoplexes, the steady-state solution can be obtained by solving the eigenvalue problem for the rank-2 transition tensor and calculating the leading eigenvector corresponding to the unitary eigenvalue. Similarly, in the case of multilayer networks, the solution can be obtained by calculating the leading *eigentensor*, solution of the higher-order eigenvalue problem

$$T^{i\alpha}_{j\beta}\Pi_{i\alpha} = \lambda \Pi_{j\beta}. \tag{10}$$

We refer to 4 for the mathematical details to solve this problem.

The probability $\Pi_{j\beta}$, that we define *random walk occupation centrality*, accounts for the full interconnected structure of the multilayer network. Although different exploration strategies can be adopted to walk in a multilayer network [36], here we focus on the natural extension of well-known random walks in monoplex networks [25]. In this process, the walker in node i and layer α might jump to one of its neighbors $j \neq i$ – within the same layer – with uniform probability, or might switch to its counterpart i in a different interconnected layer $\beta \neq \alpha$. It is worth remarking that the inter-layer connection is treated as an edge that can be chosen randomly among all edges traversing the node.

In the more general case of weighted networks, the jumping probability is proportional to the weight of the edges. Let us indicate with $s_{i\alpha}$ the strength of node i in layer α, including the inter-layer connections. The multi-strength vector, whose components indicate the strength of each node accounting for the full multilayer structure, is given by summing up its strengths across all layers, i.e., by $S_i = s_{i\alpha}u^{\alpha}$, where u^{α} is the 1-vector, namely a vector with all components equal to 1. We indicate with $D^{i\alpha}_{j\beta}$ the strength tensor whose entries are all zeros, except for $i = j$ and $\alpha = \beta$ where the entries are given by $s_{i\alpha}$. This tensor represents the multilayer extension of the well-known diagonal strength matrix in the case of monoplexes. Therefore, the transition tensor is given by $T^{i\alpha}_{j\beta} = M^{k\gamma}_{j\beta}\tilde{D}^{i\alpha}_{k\gamma}$, where $\tilde{D}^{i\alpha}_{j\beta}$ is the tensor whose entries are the inverse[1] of the non-zero entries of the strength tensor. For this classical random walk, it can be easily shown that $\Pi_{i\alpha} \propto s_{i\alpha}$ [36].

This centrality, as others in the rest of the paper, assigns a measure of importance to each node in each layer, accounting for the full interconnected structure of the multilayer network. However, in practical applications one is often interested in assigning a global measure of importance to each node, aggregating the information obtained from the different layers. The choice of the aggregation method is not trivial, it strongly influences the final estimation and might lead to wrong results.

However, this is not case for the framework discussed in the present study. In fact, the centrality $\Pi_{i\alpha}$ is calculated by inherently accounting for the interconnected structure of the whole system. We do not require to combine arbitrarily the information from different separate measures. In our framework, the most intuitive type of aggregation, i.e., summing up over layers, represents the unique and correct choice. Let $\pi_i = \Pi_{i\alpha}u^{\alpha}$ be the random walk centrality measure obtained by aggregating over layers. Here, π_i indicates the probability of finding the walker in node i, regardless of the layer. It is worth noting that this probability is proportional to $s_{i\alpha}u^{\alpha}$, i.e., the multi-strength of node i. Therefore, in this specific case, the computation of the centrality by means of the aggregated network would provide the same result of the calculation accounting for the interconnected multiplex, if inter-layer edges are accounted for as self-loops. In the more specific case that the inter-layer edges

[1]It is worth remarking that, in general, this is different from the inverse of a tensor $A^{i\alpha}_{j\beta}$, that is defined as the tensor $B^{i\alpha}_{j\beta}$ such that $A^{i\alpha}_{k\gamma}B^{k\gamma}_{j\beta} = \delta^{i\alpha}_{j\beta}$, where $\delta^{i\alpha}_{j\beta} = \delta^{i}_{j}\delta^{\alpha}_{\beta}$.

have the same strength for all nodes, the random walk centrality will be just proportional to the degree of the aggregated network, without necessity of accounting for the self-loops. Unfortunately, this is no more the case for the centrality measures discussed in the rest of this study, where calculating the diagnostics from the aggregate might lead to wrong conclusions.

A ground-truth for this diagnostics can be obtained from numerical simulations of the random walk process in the multilayer network, where the larger the number of times the walker hits a node larger the random walk centrality of that node. In Supplementary Fig. **6** we show the comparison between π_i obtained from simulation and its theoretical prediction. As expected, the agreement is excellent and this equivalence holds regardless of the number of nodes in the network, the number of layers or their topology.

PageRank centrality. We capitalize on this result to extend to interconnected networks a widely adopted measure of centrality, i.e., the PageRank [26]. A recent study in this direction has been reported in [22], in the case of edge-colored graphs where the authors, exploiting the random walk interpretation of PageRank centrality, define the PageRank of a multiplex network by means of a random walk subjected to teleportation. In that study, the PageRank for nodes in the first layer is computed using the standard definition for a monoplex [26], whereas the PageRank for nodes in the second layer is computed using the centrality information obtained from the first one. It is worth noting that this definition is limited to edge-colored graphs with only two layers, being any extension to a larger number of layers possible but very complicated from the mathematical point of view.

Here, we exploit the fact that PageRank centrality can be seen as the steady-state solution of the equation $p_j(t+1) = R^i_j p_i(t)$ in the case of monoplexes, where R^i_j is the rank-2 transition tensor (i.e., the transition matrix) of a random walk where the walker jumps to a neighbor with rate r and teleport to any other node in the network with another rate r'. For simplicity, we assume that $r' = 1 - r$ in the following. In the case of interconnected multilayer networks, the teleportation might occur to any other node in any layer. Depending on the application of interest, the walker can be teleported to other nodes with a rate that is specific to each layer. However, to keep the study as simple as possible, we consider the case with the same teleportation rate for all layers. Let $R^{i\alpha}_{j\beta}$ be the corresponding transition tensor, where the walker jumps to a neighbor with rate r and teleport to any other node in the network with rate $1 - r$. This rank-4 tensor is given by

$$R^{i\alpha}_{j\beta} = rT^{i\alpha}_{j\beta} + \frac{(1-r)}{NL} u^{i\alpha}_{j\beta}, \tag{11}$$

where $u^{i\alpha}_{j\beta}$ is the rank-4 tensor with all components equal to 1. The steady-state solution of the master equation corresponding to this transition tensor provides the PageRank centrality for interconnected multiplex networks. It is worth noting that the above definition is valid for all multiplexes where all nodes have out-going edges. If

this is not the case, as in several real-world networks, Eq. (11) reduces to $R^{i\alpha}_{j\beta} = \frac{1}{NL} u^{i\alpha}_{j\beta}$ for all nodes i with no out-going connections, ensuring the correct normalization of the transition tensor $R^{i\alpha}_{j\beta}$. For the analysis reported in the present study, we use $r = 0.85$ as in the classical PageRank algorithm.

To compute the aggregate centrality of a node, accounting for the whole interconnected topology, we proceed as for the random walk occupation centrality previously discussed. Let $\Omega_{i\alpha}$ be the eigentensor of the transition tensor $R^{i\alpha}_{j\beta}$ (see 4 for details), denoting the steady-state probability to find the walker in node i and layer α. The multilayer PageRank is obtained by simply contracting the layer index of the eigentensor with the 1-vector: $\omega_i = \Omega_{i\alpha} u^{\alpha}$, i.e., by summing up over layers.

Eigenvector centrality. Among the numerous notions of centrality introduced to quantify the importance of nodes (and other components) in a network [37], eigenvector centrality is among the oldest ones. A node i has a high eigenvector centrality if its neighbors also have high eigenvector centrality, and the recursive nature of this notion yields a vector of centralities that satisfies an eigenvalue problem.

In the case of monoplexes, the *eigenvector centrality vector*, whose components are the centralities of nodes according to [38, 39], is a solution of the tensorial equation $W^i_j v_i = \lambda_1 v_j$, where λ_1 is the largest eigenvalue of W^i_j and v_i indicates the eigenvector centrality of node i.

A naive approach for the calculation of the importance of each node might be to project the interconnected topology to an aggregated monoplex, and to associate to each node the centrality he or she has in such an aggregated network. The main drawback of this approach is that it mixes the information from all layers with uncontrollable effects, as shown in the Supplemental Material for both synthetic and empirical networks.

Another attempt to extend this calculation to the case of multilayer networks might be to calculate the eigenvector centralities for each layer separately, to build the tensor $\bar{V}_{i\alpha}$ encoding the centrality of each node in each layer. The successive step would be to choose an heuristic aggregation of such centralities to assign a unique centrality measure to each node, regardless of the layer. However, the tensor $\bar{V}_{i\alpha}$ is not the solution of a unique eigenvalue problem but the combination of the solutions of L different eigenvalue problems treated separately, therefore it is not the natural extension of the notion of eigenvector centrality to the realm of interconnected multilayer networks.

Instead, according to [7], this descriptor can be obtained as the solution of the tensorial equation

$$M^{i\alpha}_{j\beta} \Theta_{i\alpha} = \lambda_1 \Theta_{j\beta}, \tag{12}$$

where λ_1 is the largest eigenvalue and $\Theta_{i\alpha}$ is the corresponding *eigentensor* encoding the centrality of each

node in each layer when accounting for the whole interconnected structure. The eigentensor can be obtained by means of an iterative procedure, as the power method in the case of monoplexes. A proof of the existence of such eigentensor is provided in 4. Thus, the multilayer generalization of Bonacich's eigenvector centrality [38, 39] is given by $\Theta_{j\beta} = \lambda_1^{-1} M^{i\alpha}_{j\beta} \Theta_{i\alpha}$ [7].

As already pointed out in the previous sections, the overall centrality of each node can be simply obtained by contracting over layers the centrality of each node in each layer, i.e., by $\theta_i = \Theta_{i\alpha} u^{\alpha}$.

At variance with the eigenvector centrality calculated from the monoplex aggregated *before* the calculation and the one calculated by heuristically aggregating the centralities obtained separately, our measure is obtained from the mathematical extension of the original definition. The aggregation performed at the end of the calculation does not require any heuristic choice, because it is already accounting for the whole interconnected topology and, as we have previously shown, it is enough to contract the resulting eigentensor.

Katz centrality. It is a well-known fact that eigenvector centrality can lead to wrong results in the case of directed networks. In fact, nodes with only outgoing edges have an eigenvector centrality of 0 if Bonacich's definition is adopted. Moreover, in this case there are two leading eigenvectors, for in-going centrality and out-going centrality, requiring to distinguish between covariant and contravariant calculations. The Katz centrality [40] attempts to solve the above problem by assigning a small amount b of centrality to each node before calculating centrality. For monoplexes, the Katz centrality is given by $v_j = [(\delta - aW)^{-1}]^i_j u_i$, where a must be smaller than the largest eigenvalue and often one chooses $b = 1$.

Following a similar idea, we define the centrality tensor for each node in each layer as the solution of the tensorial equation

$$\Phi_{j\beta} = a M^{i\alpha}_{j\beta} \Phi_{i\alpha} + b u_{j\beta}, \tag{13}$$

corresponding to the natural extension of the equation proposed by Katz to the case of interconnected multilayer networks. The solution is given by $\Phi_{j\beta} = [(\delta - aM)^{-1}]^{i\alpha}_{j\beta} U_{i\alpha}$, where $\delta^{i\alpha}_{j\beta} = \delta^i_j \delta^{\alpha}_{\beta}$. As for the eigentensor centrality, this *Katz centrality tensor* accounts for the whole interconnected topology and it is enough to contract it with the 1-vector to obtain the Katz centrality for each node, i.e., $\phi_i = \Phi_{i\alpha} u^{\alpha}$.

HITS centrality. Similarly to the PageRank, another approach was introduced to rank Web sites with respect to their importance for users. This approach considers two different descriptors for each node, namely hub and authority [41]. In fact, Web pages that point to an important page generally also point to other important pages, building a structure similar to a bipartite topology where relevant pages – i.e., authorities – are pointed

by special Web pages – i.e, hubs. It follows that nodes with high authority centrality are linked by nodes with high hub centrality while very influent hubs point to nodes which are very authoritative. Such a mechanism is described by two coupled equations which reduce to the two eigenvalue problems $(WW^\dagger)^i_j v_i = \lambda_1 v_j$ and $(W^\dagger W)^i_j z_i = \lambda_1 z_j$, where $W^\dagger$ denotes the transpose of the adjacency tensor, λ_1 indicates the leading eigenvalue while v_i and z_i indicate hub and authority scores, respectively. The natural extension of the equations proposed by Kleinberg to the case of interconnected multilayer networks is given by

$$(MM^\dagger)^{i\alpha}_{j\beta}\Gamma_{i\alpha} = \lambda_1 \Gamma_{j\beta}, \tag{14}$$

$$(M^\dagger M)^{i\alpha}_{j\beta}\Upsilon_{i\alpha} = \lambda_1 \Upsilon_{j\beta}, \tag{15}$$

$$\tag{16}$$

where $\Gamma_{i\alpha}$ and $\Upsilon_{i\alpha}$ indicate hub and authority centrality, respectively. It is worth remarking that for undirected interconnected multiplexes, hub and authority scores are the same and equal to the corresponding eigenvector centrality. The hub and authority tensors should be contracted with the 1-vector to obtain the scores corresponding to each node regardless of the layer, i.e., $\gamma_i = \Gamma_{i\alpha}u^\alpha$ and $\upsilon_i = \Upsilon_{i\alpha}u^\alpha$, respectively.

Centrality measures based on shortest path. For sake of completeness, in this paragraph we briefly extend centrality measures based on shortest paths, namely betweenness and closeness.

Equivalently to the case of a monoplex, we define a path $\ell_{[o\sigma\to d\gamma]} \in \mathcal{P}_{[o\sigma\to d\gamma]}$, in the interconnected multilayer network, as an ordered sequence of nodes which starts from node o in layer σ and finishes in node d in layer γ. We require that there exists an edge between all pairs of consecutive nodes in ℓ. Here, $\mathcal{P}_{[o\sigma\to d\gamma]}$ indicates the set of all possible paths between node o in layer σ and node d in layer γ. For every path $\ell_{[o\sigma\to d\gamma]}$ it is possible to define a cost function $c(\ell_{[o\sigma\to d\gamma]})$, usually depending on the weight of the edges the path traverses and on the application of interest, to account for the "goodness" of the path. Hence, the shortest path from node o in layer σ to node d in layer γ is the path

$$\ell^*_{[o\sigma\to d\gamma]} = \min_{\ell'_{[o\sigma\to d\gamma]}\in\mathcal{P}_{[o\sigma\to d\gamma]}} c(\ell'_{[o\sigma\to d\gamma]}) \tag{17}$$

which minimizes the cost function. Using Eq. (17) we define the shortest path from node o to node d, regardless of the layer, as

$$\ell^*_{[o\to d]} = \min_{\sigma,\gamma\in\{1,2,\ldots,L\}} \ell^*_{[o\sigma\to d\gamma]}. \tag{18}$$

The centrality $\hat{\tau}_j$ of node j is defined to be proportional to the number of times that node j, regardless of the layer, belongs to the set $\ell^*_{[o\to d]}$ for every possible origin-destination pair (o, d).

The extension of the shortest-path betweenness centrality, defined in the case of monoplex networks in [42],

is obtained by counting the number of shortest paths between any pair of *origin* and *destination* nodes (o, d), that go though node j in the interconnected structure.

On the other hand, in the same spirit of monoplex networks, we define the shortest-path closeness centrality of a node j in an interconnected multilayer topology as the average of the inverse of the cost of the shortest paths which start from any other node o in the network. Thus, given the cost of a shortest path $c(\ell^*_{[o\to i]})$ between node i and node o, the shortest-path closeness centrality $\hat{\xi}_i$ can be easily computed by considering all possible origin nodes o.

Supplementary Note 4. Eigenvalue problem with tensors

The eigenvalue problem for a rank-2 tensor, i.e., a standard matrix, is defined by $W^i_j v_i = \lambda v_j$. The extension of this problem to rank-4 tensors leads to the equation

$$M^{i\alpha}_{j\beta} V_{i\alpha} = \lambda V_{j\beta}. \tag{19}$$

To solve this problem, it is worth noting that any tensor can be *unfolded* to lower rank tensors [43]. For instance, a rank-2 tensor like W^i_j, with N^2 components, can be flattened to a vector w_k with N^2 components. In the case of the rank-4 multilayer adjacency tensor $M^{i\alpha}_{j\beta}$, although any unfolding is allowed, it is particularly useful for some applications to choose the ones flattening to a squared rank-2 tensor $\tilde{M}^k_l$ with $NL \times NL$ components, where L indicates the number of layers [5]. In fact, this unfolding produces as many block adjacency matrices, named *supra-adjacency matrices* in some applications [5, 36, 21], as the number of permutations of diagonal blocks of size N^2, i.e., $L!$. However, such unfoldings do not alter the spectral properties of the resulting supra-matrix and can be used to solve the eigenvalue problem for rank-4 tensors. In fact, the solution of the eigenvalue problem

$$\tilde{M}^k_l \tilde{v}_k = \tilde{\lambda}_1 \tilde{v}_l, \tag{20}$$

is a *supra-vector* with NL components which corresponds to the unfolding of the eigentensor $V_{i\alpha}$.

Supplementary Note 5. Rankings and consensus

In this section we describe the details of the calculation of the rankings of nodes from their centralities, and the different consensus of layer rankings which have been explored.

The easiest way to obtain a ranking from any kind of centrality simply consists in sorting the centralities in descending order, and then assign a rank to each node according to its position in the sorted list; rankings range from 1 (highest centrality) to N (lowest centrality). However, it must be taken into account that several nodes may share the same centrality, and in this case we should assign to all of them the same rank; we have chosen to give them a rank equal to the smallest position in the list among the repeated centralities, e.g. if positions 85 to 90 have centralities 0.4, 0.3, 0.3, 0.3, 0.2 and 0.1, the corresponding rankings would be 85, 86, 86, 86, 89 and 90.

Another aspect which affects the final ranking of the nodes is the precision of the calculation of the centralities. Since the accuracy of the algorithms to find eigenvectors is limited, we have rounded all the centralities to the sixth significant digit, thus avoiding spurious differences in ranking.

For the comparison of versatility with classical centralities which do not consider the multilayer structure, there are two main options:

- *Aggregate network*: Build the aggregate network, and calculate the centralities of the nodes and their ranks.
- *Layer networks*: Calculate the centralities at each layer, considered as separate networks, and combine them (or their ranks) to obtain a new ranking of the nodes.

In the second case, it is necessary to define how to combine the layer centralities or rankings. We have used the following methods:

- *Average*: For each node, calculate the average of its centralities in each layer, and create a ranking from this list of average centralities.
- *Consensus*: Rank each layer independently and apply a consensus method to obtain the final rank.
  - *Unweighted consensus*: For each node, we sort its rankings from best (lowest) to worst (highest). We then sort the nodes according to their best rankings; for nodes with equal best ranking, the second best ranking is used; and if they are still equal, we consider the third, fourth, fifth, and so on, until all ranks are consumed.
  - *Weighted (by links) consensus*: This is similar to the unweighted consensus, but giving more importance to the ranks in layers with more links. For each node, instead of having a list containing one value of the rank for each layer, we clone it as many times as number of links has that layer, thus

generating a list for each node of length equal to the number of intra-layer edges in the multilayer network.

– *Weighted (by nodes) consensus*: Equivalent to the weighted (by links) consensus, but using as weights the number of non-isolated nodes in each layer.

The aggregation of rankings is an old field of research which has been analyzed from many points of view, from consensus rules like the ones exposed above, to approaches based on distances between rankings (e.g., the metric for weak orders by Kemeny and Snell [44]) or using probabilistic models for the rankings and trying to optimize the expected loss [45]. The main conclusion here is the absence of a unique way to combine rankings, thus making the comparison between layer-based rankings and versatility rather arbitrary.

Supplementary Note 6. Supplementary tables

Wikipedia dataset

We considered biologists, chemists, computer scientists, economists, inventors, mathematicians, philosophers and physicists in Wikipedia, and we built an interconnected multilayer network where each layer represents a subject and two people are connected if a hyperlink exists between their pages. The weight of each link (i.e., the number of hyperlinks between two Web pages) is redistributed across all layers where the corresponding nodes exist. If a pair of nodes does not exist in the same layer, a inter-layer edge between them is created. For simplicity, we considered only the largest connected component and people with at least four links.

This dataset is only a proxy of the "true" network where biologists, chemists, computer scientists, economists, inventors, mathematicians, philosophers and physicists are linked according to their role in one or more disciplines. For this reason, it should be taken into account that some nodes, central in the Wikipedia dataset, could not play any role in the true network: we identified a few cases of this type. For instance, in the following table the top ranked node is Edmund F. Robertson. We found that the reason is that he is one of the creators of the MacTutor History of Mathematics Archive, a Web site containing detailed biographies on many mathematicians in our network, whose corresponding pages point to this Web site and the Wikipedia page of Edmund F. Robertson. This kind of bias is very difficult to eliminate automatically at large scale and we believed that, for the purpose of our study, we could just ignore this node in the ranking presented in the paper, without excluding him from the network. The ranking discussed in the paper have been readjusted accounting this shift of one unit.

Moreover, a very few spurious nodes as "Indian astronomy" have been found in the dataset by a later manual inspection out of thousands of pages. Such nodes have not been removed because the aim of the present work is not to perform a detailed analysis of the most influent people in the history of science and philosophy but rather to show the differences between ranking by accounting for the whole interconnected structure of a network and ranking by neglecting such information. Of course, researchers interested in performing more meaningful analyses of this dataset should take care and be aware of this kind of biases.

The aggregation of this multilayer network to a weighted single-layer network is performed by accounting for weight of inter-layer links that are not between nodes' replica. For example, if there exists a hyperlink between two scientists (a and b) and these happen to satisfy three disciplines simultaneously (e.g., physicists, philosophers, and chemists), then we sum up the weights of the 3 intra-layer directed edges between a and b, one for each layer, obtaining the observed number of hyperlinks between a and b. If a and b do not share at least one layer, and directed inter-layer edges are made between all pair of layers where a and b exist, we sum up the weights of these inter-layer edges to obtain the aggregated weight corresponding, again, to the observed number of hyperlinks between a and b. Finally, inter-layer edges between all nodes' replicas are not accounted for in the aggregation process.

In Tables **1** to **3** we show the top 25 ranked persons according to PageRank, eigenvector and betweenness multilayer centralities, compared to their corresponding rankings obtained from the aggregated network, and from the average of the layer centralities. We also computed intra- and global diversity measures: Intra-layer diversity accounts for the number of layers in which a node has intra-layer links, while global diversity is the number of layers that have any kind of links from or to the corresponding node. Note that layer centralities lose all the information about interlayer links, thus it is not surprising that their average leads to the most distant rankings.

Wikipedia alternative dataset

Here we focused our attention on a more homogenous set of individuals considering the subset of people whose date of death was between 1860 and 1960. The choice of this cut ensures a fair comparison between philosophers, physicists, *etc*, being active in their field during the same period.

EU Airports dataset

Here, the network is built from flight routes operated by 37 different air companies (layers) between 450 European airports (nodes) [1].

The aggregation of this multilayer network to a weighted single-layer network is performed by summing up the weights of intra-layer links and by discarding inter-layer edges between all nodes' replicas.

EU APS dataset

We collected all papers published in the journals of the American Physical Society journals (Physical Review Letters, Physical Review and Review of Modern Physics) between 2005 and 2009, focusing on scientists working in European institutions. Meta-data in the dataset provided information about the topic of the papers through the specification of the assigned "Physics and Astronomy Classification Scheme" (PACS) code, developed by the American Institute of Physics (AIP) and used in Physical Review since 1975 to identify fields and sub-fields of physics. We restricted the analysis only to papers with at most ten authors, to avoid biases due to the papers of experimental high-energy physics in which all the project collaborators are listed as co-authors, and to authors who published at least 5 papers in the considered time period. Then we considered only the giant connected component of the co-authorship network built starting from those authors, resulting in 3171 authors.

We exploited this information to build an interconnected multilayer network in which each layer represents a sub-field of physics, as defined by the PACS systems:

- *General*
- *The Physics of Elementary Particles and Fields*;
- *Nuclear Physics*, *Atomic and Molecular Physics*;
- *Electromagnetism, Optics, Acoustics, Heat Transfer, Classical Mechanics, and Fluid Dynamics*;
- *Physics of Gases, Plasmas, and Electric Discharges*;
- *Condensed Matter: Structural, Mechanical and Thermal Properties*;
- *Condensed Matter: Electronic Structure, Electrical, Magnetic, and Optical Properties*;
- *Interdisciplinary Physics and Related Areas of Science and Technology*;

- *Geophysics, Astronomy, and Astrophysics.*

Two authors are connected in the aggregated network if they co-authored at least one paper in the dataset. The weight of the corresponding link is equal to the number of papers that they co-authored. In the multilayer network the contribution of value 1 of each paper is redistributed across all involved layers. For example, if two authors a and b co-authored a paper characterized by three (or more) PACS codes belonging to three different sub-fields, then the following links are created: three intra-layer links between a and b (one on each of the three layers), three inter-layer links between a on each layer and the replica of itself in each of the two other layers, three inter-layer links between b on each layer and the replica of itself in each of the two other layers, and six inter-layer links between a on each layer and b in each of the two other layers, totaling 15 links. The weight of each link is equal to $1/15$, so that the sum of the weights between two nodes across all layers corresponds to the number of papers co-authored by the corresponding authors, exactly like in the aggregate.
Let us also highlight that two PACS codes belong to two different sub-fields only if they belong to two different general categories defined above. Therefore a paper could have three PACS codes but the corresponding authors could be linked only in one or two layers according to how many different categories the three PACS codes belong to. For example a paper characterized by the codes 89.75.2k, 87.23.Ge, and 05.70.Ln only refers to two layers: *Interdisciplinary Physics and Related Areas of Science and Technology* and *General*.

The aggregation of this multilayer network to a weighted single-layer network is performed similarly to the case of the Wikipedia dataset, by accounting for weight of inter-layer links that are not between nodes' replica. The resulting aggregated network is exactly what one would obtain by observing the number of co-authored papers among authors in the dataset.

EU APS alternative dataset

We built an alternative dataset in the same fashion described in the previous section, but restricting the analysis only to the decade starting in 1990. In particular we selected only the authors who published at least ten papers in this time interval, and were active in at least seven out of the ten years considered, resulting in a network composed of 535 nodes. In this way we could check our results on a more homogeneous subset.

Online Social Multiplex dataset

The same user might use different nicknames in different social networks and this is generally the main obstacle to build a multiplex social network, because it is not possible to retrieve information about the identity corresponding to each nickname. However, many recent online ranking platforms provide a suitable environment to collect such an information, because each user (with a unique identifier on such platforms) provides his or her nicknames in the different social networks.

In November 2013, we gathered publicly available information from one of such online ranking platforms about approximately 32000 users of different online social networks. We focused our attention on Twitter and Instagram because they provide Application User Interfaces (API) easy to use and fast enough to allow us to reconstruct the underlying social graph (directed and unweighted). We found 13297 users in our dataset with an account on both Twitter and Instagram, with an edge overlap between the two networks of about 25%.

The follower/friend relationships between all users of Twitter in our sub-sample have been obtained by querying Twitter through the official API (https://dev.twitter.com/rest/public) and we used this information to build the first layer of directed and unweighted social relationships. The relationships between all users of Instagram in our sub-sample have been obtained by querying Instagram through the official API (http://instagram.com/developer/endpoints/) and we used this information to build the second layer of directed and unweighted social relationships. Bidirectional inter-layer edges are made only between a user and his or her replica on the other layer, if he or she has an account in both online social networks.

The aggregation of this multilayer network to a weighted single-layer network is performed by summing up the weights of intra-layer links and by discarding inter-layer edges between all nodes' replicas.

Supplementary Note 7. Covering and congestion analysis

Random walks covering analysis

Random walk dynamics between vertices and layers of a multilayer network can be described by four fundamental transition rules [36] that can be represented in one compact tensorial formulation [7]. Let $\mathcal{P}_{i\alpha}^{j\beta}$ be the tensor accounting for the probability to move from node i in layer α to node j in layer β within one step.

Let $p_{i\alpha}(t)$ the probability that the random walker is visiting node i in layer α at time t. The master equation describing the probability of finding the walker in vertex j and in layer β at time $t + \Delta t$ is given by

$$p_{j\beta}(t + \Delta t) \quad = \quad \mathcal{P}_{j\beta}^{i\alpha} p_{i\alpha}(t). \tag{21}$$

We indicate with $\mathbf{p}_\alpha$ the row vector with N components $p_{i\alpha}$ with respect to layer α, with $\mathbf{P} \equiv (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_L)$ the *supra-vector* with NL components, and we rewrite Eq. 21 as $\dot{\mathbf{P}}(t) = -\mathbf{P}(t)\mathcal{L}$, hereafter referred to as the random walker equation. In this equation, $\mathcal{L}$ is the $NL \times NL$ *normalized supra-Laplacian* matrix [5, 36], a flattening of $\mathcal{P}_{j\beta}^{i\alpha}$.

The *coverage* $c(t)$ is defined as the average fraction of distinct vertices visited at least once in a time less than or equal to t, regardless of the layer, assuming that walks started from any other vertex in the network. The time evolution of the supra-vector of probabilities $\mathbf{P}$ for a walker starting from vertex j is given by $\mathbf{P}(t) = \mathbf{P}_j(0)e^{-\mathcal{L}t}$. From eigen-decomposition of the normalized supra-Laplacian, we obtain

$$\mathbf{P}(t) = \mathbf{P}_j(0)e^{-\mathcal{L}t} = \mathbf{P}_j(0) \sum_{\ell=1}^{NL} e^{-\lambda_\ell t} \mathcal{V}_\ell, \tag{22}$$

where each supra-matrix $\mathcal{V}_\ell$ is obtained from products of eigenvectors and λ_ℓ are the corresponding eigenvalues. Finally, assuming the general case that the normalized supra-Laplacian matrix has more than one null eigenvalue, the coverage is given by

$$(t) \quad = \quad 1 - \frac{1}{N^2} \sum_{i,j=1}^{N} \delta_{i,j}(0) \times \exp\left[- \sum_{\ell \in \Lambda^0} \mathcal{C}_{i,j}(\ell) t - \sum_{\ell \in \Lambda^+} \mathcal{C}_{i,j}(\ell) \frac{e^{-\lambda_\ell t} - 1}{-\lambda_\ell} \right],$$

where Λ^0 and Λ^+ indicate the sets of all null and positive eigenvalues of the normalized supra-Laplacian, respectively; $\delta_{i,j}(0) = 0$ for $j = i$ (i.e., the random walker starts in vertex i and the probability of not finding it in the same vertex must be zero) and 1 otherwise; $\mathcal{C}_{i,j}(\ell) = \mathbf{P}_j(0)\mathcal{V}_\ell \mathcal{P} \mathbf{E}_i^\dagger$ are constants depending on the vertex, the transition matrix, the eigen-decomposition and the initial conditions [36].

Congestion analysis

To simulate the airport transportation network traffic, we extended to multilayer networks a previous model proposed by some of the authors [46] . The modeling of the airports transportation network is as follows. To each airport, in each layer (airline), we assign a first-in-first-out queue where airplanes will wait to be routed. Since our analysis is intended to describe the relative increase of the queue, the maximum capacity of a node in terms of the queue size is unlimited. However, each node has a limited capacity to route planes. We set this capacity represented by η to one plane per unit of time. The model however can deal with any value of η. Given that each layer corresponds to an airline, at each time step, each company (node in a given layer) can route a maximum of τ airplanes from each airport. Note that changing layer corresponds to fly with a flight operated by a different company. With respect to the dynamics (walkers), at each time step, to each airport, we inject $L \times \rho$ (where L is the number of layers) airplanes with an airport destination chosen at random (a given node independent on the layer). After injection and during the following time steps the airplanes travel to their destinations using shortest-path routes. It is possible that several shortest paths exist between the two airports, some may traverse using a single layer and some may use several layers and so the full multiplex structure. From these possible paths each airplane chooses one at random, to reach its destination airport. Since each airport can be represented in several layers it is common that an airplane starts its trip in one layer and ends its trip in another layer.

The phenomena of congestion occurs when a company in an airport is not able to handle all the incoming traffic and the amount of airplanes waiting to be routed increases proportional to time [46]. We consider an airport is congested if any of the companies operating in it is congested. To obtain the maximum injection rate (maximum ρ value) each airport is able to handle before it reaches a congestion state we conducted a series of Monte Carlo experiments with different values of ρ starting from 0.1 and ending at 1 in steps of 0.001. For each experiment, we analyzed how the size of the queue of all companies operating in each airport behave with respect to time during 9000 time steps (with a warm up of 1000 time steps), we consider the airport is congested if the linear approximation of the queue's size with respect to time exceeds an slope of 0.01.

It is known that the node betweenness and the phenomena of congestion are closely related [46]. In particular, it can be use to accurately compute the critical injection rate that will congest the first airport in single layer networks. This airport corresponds to the one with higher betweenness. Here, we use a similar approach extended to multilayer networks [47] and we analyze the order in which airports congest in the simulations with respect to the order given by its betweenness centrality on the aggregated network and its betweenness versatility on the multiplex network. For the betweenness versatility, as in the simulations, we consider an

airport is congested if any of the companies operating in it is congested, so the critical injection rate of the airport is chosen to be the minimum obtained in the different layers. Results are shown in figure 3 (panel B) of the main document. Low ranked airports arrive to congestion earlier. We see that the betweenness versatility is able to accurately predict the first airport that arrive to congestion state. This can be seen in the figure where the airport ranked zeroth by the simulations is also ranked zeroth by the betweenness. For the other airports, we see that the betweenness versatility nicely correlates with the order obtained in the simulations. The results obtained with the betweenness centrality also show that it is able to detect the first airport to arrive to congestion (even though this results is in general not true). However, we see that the ranking obtained with it poorly correlates with the simulations.

Supplementary References

[1] Cardillo, A. *et al.* Emergence of network features from multiplexity. *Scientific reports* **3**, 1344 (2013).

[2] Blonder, B. & Dornhaus, A. Time-ordered networks reveal limitations to information flow in ant colonies. *PloS one* **6**, e20298 (2011).

[3] Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327**, 425–431 (2010).

[4] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).

[5] Gómez, S. *et al.* Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701 (2013).

[6] Radicchi, F. & Arenas, A. Abrupt transition in the structural formation of interconnected networks. *Nature Physics* **9**, 717–720 (2013).

[7] De Domenico, M. *et al.* Mathematical formulation of multi-layer networks. *Phys. Rev. X* **3**, 041022 (2013).

[8] Sole-Ribalta, A. *et al.* Spectral properties of the laplacian of multiplex networks. *Phys. Rev. E* **88**, 032807 (2013).

[9] Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys. Rev. Lett.* **111**, 128701 (2013).

[10] García, R. S., Cozzo, E. & Moreno, Y. Dimensionality reduction and spectral properties of multiplex networks. *arXiv:1311.1759* (2013).

[11] Lee, K.-M., Kim, J. Y., Cho, W.-K., Goh, K.-I. & Kim, I.-M. Correlated multiplexity and connectivity of multiplex random networks. *New J. Phys.* **14**, 033027 (2012).

[12] Nicosia, V., Bianconi, G., Latora, V. & Barthelemy, M. Growing multiplex networks. *Phys. Rev. Lett.* **111**, 058701 (2013).

[13] Bianconi, G. Statistical mechanics of multiplex networks: Entropy and overlap. *Phys. Rev. E* **87**, 062806 (2013).

[14] Battiston, F., Nicosia, V. & Latora, V. Metrics for the analysis of multiplex networks. *arXiv:1308.3182* (2013).

[15] Sola, L. *et al.* Eigenvector centrality of nodes in multiplex networks. *Chaos* **3**, 033131 (2013).

[16] Gao, J., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Networks formed from interdependent networks. *Nature Physics* **8**, 40–48 (2011).

[17] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).

[18] Dickison, M., Havlin, S. & Stanley, H. E. Epidemics on interconnected networks. *Phys. Rev. E* **85**, 066109 (2012).

[19] Padgett, J. F. & Ansell, C. K. Robust action and the rise of the medici, 1400-1434. *American journal of sociology* 1259–1319 (1993).

[20] Kivelä, M. *et al.* Multilayer networks. *Journal of Complex Networks* **2**, 203–271 (2014).

[21] Cozzo, E. *et al.* Clustering coefficients in multiplex networks. *arXiv:1307.6780* (2013).

[22] Halu, A., Mondragon, R. J., Pansaraza, P. & Bianconi, G. Multiplex PageRank (2013). `arXiv:1306.3576`.

[23] Chung, F. R. K. *Spectral Graph Theory* (American Mathematical Society, Providence, RI, 1997), 2nd edn.

[24] Newman, M. *Networks: An Introduction* (Oxford University Press, Inc., New York, NY, USA, 2010).

[25] Noh, J. D. & Rieger, H. Random walks on complex networks. *Phys. Rev. Lett.* **92**, 118701 (2004).

[26] Brin, S. & Page, L. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the Seventh International Conference on World Wide Web 7*, WWW7, 107–117 (Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, 1998).

[27] Callaghan, T., Mucha, P. J. & Porter, M. A. Random walker ranking for NCAA Division I-A football. *American Mathematical Monthly* **114**, 761–777 (2007).

[28] Viswanathan, G. *et al.* Optimizing the success of random searches. *Nature* **401**, 911–914 (1999).

[29] Yang, S.-J. Exploring complex networks by walking on them. *Phys. Rev. E* **71**, 016107 (2005).

[30] da Fontoura Costa, L. & Travieso, G. Exploring complex networks through random walks. *Phys. Rev. E* **75**, 016102 (2007).

[31] Rozenfeld, H. D., Kirk, J. E., Bollt, E. M. & Ben-Avraham, D. Statistics of cycles: how loopy is your network? *J. Phys. A* **38**, 4589 (2005).

[32] Gfeller, D. & De Los Rios, P. Spectral coarse graining of complex networks. *Phys. Rev. Lett.* **99**, 038701 (2007).

[33] Rosvall, M. & Bergstrom, C. T. An information-theoretic framework for resolving community structure in complex networks. *PNAS* **104**, 7327–7331 (2007).

[34] Lambiotte, R., Delvenne, J.-C. & Barahona, M. Laplacian dynamics and multiscale modular structure in networks. *arXiv preprint arXiv:0812.1770* (2009).

[35] Newman, M. E. A measure of betweenness centrality based on random walks. *Social networks* **27**, 39–54 (2005).

[36] De Domenico, M., Solé-Ribalta, A., Gómez, S. & Arenas, A. Navigability of interconnected networks under random failures. *PNAS* **111**, 8351–8356 (2014).

[37] Wasserman, S. & Faust, K. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences (Cambridge University Press, Cambridge, UK, 1994).

[38] Bonacich, P. Technique for analyzing overlapping memberships. *Sociological Methodology* **4**, 176–185 (1972).

[39] Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology* **2**, 113–120 (1972).

[40] Katz, L. A new status index derived from sociometric analysis. *Psychometrika* **18**, 39–43 (1953).

[41] Kleinberg, J. M. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)* **46**, 604–632 (1999).

[42] Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry* 35–41 (1977).

[43] Kolda, T. G. & Bader, B. W. Tensor Decompositions and Applications. *SIAM Review* **51**, 455–500 (2009).

[44] Kemeny, J. M. & Snell, J. L. *Mathematical Models in the Social Sciences* (Blaisdell Publishing Company, Waltham, 1962).

[45] Truchon, M. Aggregation of rankings: A brief review of distance-based rules. *CIRPEE Working Paper* 05–34 (2007).

[46] Guimerà, R., Diaz-Guilera, A., Vega-Redondo, F., Cabrales, A. & Arenas, A. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.* **89**, 248701 (2002).

[47] Solé-Ribalta, A., De Domenico, M., Gómez, S. & Arenas, A. Centrality rankings in multiplex networks. In *Proceedings of the 2014 ACM Conference on Web Science*, 149–155 (ACM, 2014).