

Codifica di Testi - Introduzione XML Markup a.a. 2018-2019

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

CNR-ILC-LicoLab

Istituto di Linguistica Computazionale “A. Zampolli”,
6th October 2018

Contenuto della lezione

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare un documento XML e Definire uno Schema
 - Document Type Definition (DTD)
 - XSD
 - RELAXGN
- 4 Conclusioni

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Progress status

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni

I linguaggi di codifica

introduzione

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Definizione di codifica digitale del testo

Per **codifica** digitale dei testi intendiamo la *rappresentazione formale* di un **testo** ad un qualche livello descrittivo, su di un supporto digitale, in un formato utilizzabile da un elaboratore (*Machine Readable Form*) mediante un opportuno **linguaggio informatico** (F. Ciotti).

I linguaggi di codifica

Riassumendo

Impostazione teorico-pratica

- un testo è molto di più della sequenza di caratteri che lo compongono
- per mezzo della codifica vogliamo rendere esplicite le caratteristiche che vogliamo analizzare
- solo quello che è esplicito può essere interpretato ed elaborato dal computer
- vogliamo codificare il testo per quello che è, non per quello che sembra
- codifica da effettuare mediante linguaggio di markup

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Il markup

Il termine markup è stato utilizzato in passato per denotare i segni grafici che accompagnavano un testo apposti sul documento per indicare correzioni o modalità grafiche di stampa.

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

All flush left All HB Helvetica
(EXCEPT AS INDICATED) RETURN TO CUSTOMER

Key: Letterhead: 2x

Key: Business Card: 2x

Key: #10 Envelope: 2x

Key: Big Envelope: 2x

Key: Inside: 2x

Key: Your New Issue of the Corporate Criminal Liability Reporter: 2x

Key: P. O. BOX 58408
LOS ANGELES, CALIFORNIA 90058
213/583-1895
Gregory R. [redacted], Esq.
DIRECTOR

Key: P. O. BOX 58408
LOS ANGELES, CALIFORNIA 90058
213/583-1895
Gregory R. [redacted], Esq.
DIRECTOR

Key: P. O. BOX 58408
LOS ANGELES, CALIFORNIA 90058
BULK RATE
U. S. POSTAGE
Paid
PERMIT #2881
LOS ANGELES
CALIFORNIA

Key: P. O. BOX 58408
LOS ANGELES, CALIFORNIA 90058
BULK RATE
U. S. POSTAGE
Paid
PERMIT #2881
LOS ANGELES
CALIFORNIA

Key: INSIDE
Your New Issue of the Corporate Criminal Liability Reporter

7pt H.B. Helvetica

Same as ① & ② above

Same as ① above

Same as ① & ③ above

24pt English
(Fit on one)

Handwritten notes: 7pt, 12pt, 18pt, 24pt, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 11x, 12x, 13x, 14x, 15x, 16x, 17x, 18x, 19x, 20x, 21x, 22x, 23x, 24x, 25x, 26x, 27x, 28x, 29x, 30x, 31x, 32x, 33x, 34x, 35x, 36x, 37x, 38x, 39x, 40x, 41x, 42x, 43x, 44x, 45x, 46x, 47x, 48x, 49x, 50x, 51x, 52x, 53x, 54x, 55x, 56x, 57x, 58x, 59x, 60x, 61x, 62x, 63x, 64x, 65x, 66x, 67x, 68x, 69x, 70x, 71x, 72x, 73x, 74x, 75x, 76x, 77x, 78x, 79x, 80x, 81x, 82x, 83x, 84x, 85x, 86x, 87x, 88x, 89x, 90x, 91x, 92x, 93x, 94x, 95x, 96x, 97x, 98x, 99x, 100x, 101x, 102x, 103x, 104x, 105x, 106x, 107x, 108x, 109x, 110x, 111x, 112x, 113x, 114x, 115x, 116x, 117x, 118x, 119x, 120x, 121x, 122x, 123x, 124x, 125x, 126x, 127x, 128x, 129x, 130x, 131x, 132x, 133x, 134x, 135x, 136x, 137x, 138x, 139x, 140x, 141x, 142x, 143x, 144x, 145x, 146x, 147x, 148x, 149x, 150x, 151x, 152x, 153x, 154x, 155x, 156x, 157x, 158x, 159x, 160x, 161x, 162x, 163x, 164x, 165x, 166x, 167x, 168x, 169x, 170x, 171x, 172x, 173x, 174x, 175x, 176x, 177x, 178x, 179x, 180x, 181x, 182x, 183x, 184x, 185x, 186x, 187x, 188x, 189x, 190x, 191x, 192x, 193x, 194x, 195x, 196x, 197x, 198x, 199x, 200x, 201x, 202x, 203x, 204x, 205x, 206x, 207x, 208x, 209x, 210x, 211x, 212x, 213x, 214x, 215x, 216x, 217x, 218x, 219x, 220x, 221x, 222x, 223x, 224x, 225x, 226x, 227x, 228x, 229x, 230x, 231x, 232x, 233x, 234x, 235x, 236x, 237x, 238x, 239x, 240x, 241x, 242x, 243x, 244x, 245x, 246x, 247x, 248x, 249x, 250x, 251x, 252x, 253x, 254x, 255x, 256x, 257x, 258x, 259x, 260x, 261x, 262x, 263x, 264x, 265x, 266x, 267x, 268x, 269x, 270x, 271x, 272x, 273x, 274x, 275x, 276x, 277x, 278x, 279x, 280x, 281x, 282x, 283x, 284x, 285x, 286x, 287x, 288x, 289x, 290x, 291x, 292x, 293x, 294x, 295x, 296x, 297x, 298x, 299x, 300x, 301x, 302x, 303x, 304x, 305x, 306x, 307x, 308x, 309x, 310x, 311x, 312x, 313x, 314x, 315x, 316x, 317x, 318x, 319x, 320x, 321x, 322x, 323x, 324x, 325x, 326x, 327x, 328x, 329x, 330x, 331x, 332x, 333x, 334x, 335x, 336x, 337x, 338x, 339x, 340x, 341x, 342x, 343x, 344x, 345x, 346x, 347x, 348x, 349x, 350x, 351x, 352x, 353x, 354x, 355x, 356x, 357x, 358x, 359x, 360x, 361x, 362x, 363x, 364x, 365x, 366x, 367x, 368x, 369x, 370x, 371x, 372x, 373x, 374x, 375x, 376x, 377x, 378x, 379x, 380x, 381x, 382x, 383x, 384x, 385x, 386x, 387x, 388x, 389x, 390x, 391x, 392x, 393x, 394x, 395x, 396x, 397x, 398x, 399x, 400x, 401x, 402x, 403x, 404x, 405x, 406x, 407x, 408x, 409x, 410x, 411x, 412x, 413x, 414x, 415x, 416x, 417x, 418x, 419x, 420x, 421x, 422x, 423x, 424x, 425x, 426x, 427x, 428x, 429x, 430x, 431x, 432x, 433x, 434x, 435x, 436x, 437x, 438x, 439x, 440x, 441x, 442x, 443x, 444x, 445x, 446x, 447x, 448x, 449x, 450x, 451x, 452x, 453x, 454x, 455x, 456x, 457x, 458x, 459x, 460x, 461x, 462x, 463x, 464x, 465x, 466x, 467x, 468x, 469x, 470x, 471x, 472x, 473x, 474x, 475x, 476x, 477x, 478x, 479x, 480x, 481x, 482x, 483x, 484x, 485x, 486x, 487x, 488x, 489x, 490x, 491x, 492x, 493x, 494x, 495x, 496x, 497x, 498x, 499x, 500x, 501x, 502x, 503x, 504x, 505x, 506x, 507x, 508x, 509x, 510x, 511x, 512x, 513x, 514x, 515x, 516x, 517x, 518x, 519x, 520x, 521x, 522x, 523x, 524x, 525x, 526x, 527x, 528x, 529x, 530x, 531x, 532x, 533x, 534x, 535x, 536x, 537x, 538x, 539x, 540x, 541x, 542x, 543x, 544x, 545x, 546x, 547x, 548x, 549x, 550x, 551x, 552x, 553x, 554x, 555x, 556x, 557x, 558x, 559x, 560x, 561x, 562x, 563x, 564x, 565x, 566x, 567x, 568x, 569x, 570x, 571x, 572x, 573x, 574x, 575x, 576x, 577x, 578x, 579x, 580x, 581x, 582x, 583x, 584x, 585x, 586x, 587x, 588x, 589x, 590x, 591x, 592x, 593x, 594x, 595x, 596x, 597x, 598x, 599x, 600x, 601x, 602x, 603x, 604x, 605x, 606x, 607x, 608x, 609x, 610x, 611x, 612x, 613x, 614x, 615x, 616x, 617x, 618x, 619x, 620x, 621x, 622x, 623x, 624x, 625x, 626x, 627x, 628x, 629x, 630x, 631x, 632x, 633x, 634x, 635x, 636x, 637x, 638x, 639x, 640x, 641x, 642x, 643x, 644x, 645x, 646x, 647x, 648x, 649x, 650x, 651x, 652x, 653x, 654x, 655x, 656x, 657x, 658x, 659x, 660x, 661x, 662x, 663x, 664x, 665x, 666x, 667x, 668x, 669x, 670x, 671x, 672x, 673x, 674x, 675x, 676x, 677x, 678x, 679x, 680x, 681x, 682x, 683x, 684x, 685x, 686x, 687x, 688x, 689x, 690x, 691x, 692x, 693x, 694x, 695x, 696x, 697x, 698x, 699x, 700x, 701x, 702x, 703x, 704x, 705x, 706x, 707x, 708x, 709x, 710x, 711x, 712x, 713x, 714x, 715x, 716x, 717x, 718x, 719x, 720x, 721x, 722x, 723x, 724x, 725x, 726x, 727x, 728x, 729x, 730x, 731x, 732x, 733x, 734x, 735x, 736x, 737x, 738x, 739x, 740x, 741x, 742x, 743x, 744x, 745x, 746x, 747x, 748x, 749x, 750x, 751x, 752x, 753x, 754x, 755x, 756x, 757x, 758x, 759x, 760x, 761x, 762x, 763x, 764x, 765x, 766x, 767x, 768x, 769x, 770x, 771x, 772x, 773x, 774x, 775x, 776x, 777x, 778x, 779x, 780x, 781x, 782x, 783x, 784x, 785x, 786x, 787x, 788x, 789x, 790x, 791x, 792x, 793x, 794x, 795x, 796x, 797x, 798x, 799x, 800x, 801x, 802x, 803x, 804x, 805x, 806x, 807x, 808x, 809x, 810x, 811x, 812x, 813x, 814x, 815x, 816x, 817x, 818x, 819x, 820x, 821x, 822x, 823x, 824x, 825x, 826x, 827x, 828x, 829x, 830x, 831x, 832x, 833x, 834x, 835x, 836x, 837x, 838x, 839x, 840x, 841x, 842x, 843x, 844x, 845x, 846x, 847x, 848x, 849x, 850x, 851x, 852x, 853x, 854x, 855x, 856x, 857x, 858x, 859x, 860x, 861x, 862x, 863x, 864x, 865x, 866x, 867x, 868x, 869x, 870x, 871x, 872x, 873x, 874x, 875x, 876x, 877x, 878x, 879x, 880x, 881x, 882x, 883x, 884x, 885x, 886x, 887x, 888x, 889x, 890x, 891x, 892x, 893x, 894x, 895x, 896x, 897x, 898x, 899x, 900x, 901x, 902x, 903x, 904x, 905x, 906x, 907x, 908x, 909x, 910x, 911x, 912x, 913x, 914x, 915x, 916x, 917x, 918x, 919x, 920x, 921x, 922x, 923x, 924x, 925x, 926x, 927x, 928x, 929x, 930x, 931x, 932x, 933x, 934x, 935x, 936x, 937x, 938x, 939x, 940x, 941x, 942x, 943x, 944x, 945x, 946x, 947x, 948x, 949x, 950x, 951x, 952x, 953x, 954x, 955x, 956x, 957x, 958x, 959x, 960x, 961x, 962x, 963x, 964x, 965x, 966x, 967x, 968x, 969x, 970x, 971x, 972x, 973x, 974x, 975x, 976x, 977x, 978x, 979x, 980x, 981x, 982x, 983x, 984x, 985x, 986x, 987x, 988x, 989x, 990x, 991x, 992x, 993x, 994x, 995x, 996x, 997x, 998x, 999x, 1000x, 1001x, 1002x, 1003x, 1004x, 1005x, 1006x, 1007x, 1008x, 1009x, 1010x, 1011x, 1012x, 1013x, 1014x, 1015x, 1016x, 1017x, 1018x, 1019x, 1020x, 1021x, 1022x, 1023x, 1024x, 1025x, 1026x, 1027x, 1028x, 1029x, 1030x, 1031x, 1032x, 1033x, 1034x, 1035x, 1036x, 1037x, 1038x, 1039x, 1040x, 1041x, 1042x, 1043x, 1044x, 1045x, 1046x, 1047x, 1048x, 1049x, 1050x, 1051x, 1052x, 1053x, 1054x, 1055x, 1056x, 1057x, 1058x, 1059x, 1060x, 1061x, 1062x, 1063x, 1064x, 1065x, 1066x, 1067x, 1068x, 1069x, 1070x, 1071x, 1072x, 1073x, 1074x, 1075x, 1076x, 1077x, 1078x, 1079x, 1080x, 1081x, 1082x, 1083x, 1084x, 1085x, 1086x, 1087x, 1088x, 1089x, 1090x, 1091x, 1092x, 1093x, 1094x, 1095x, 1096x, 1097x, 1098x, 1099x, 1100x, 1101x, 1102x, 1103x, 1104x, 1105x, 1106x, 1107x, 1108x, 1109x, 1110x, 1111x, 1112x, 1113x, 1114x, 1115x, 1116x, 1117x, 1118x, 1119x, 1120x, 1121x, 1122x, 1123x, 1124x, 1125x, 1126x, 1127x, 1128x, 1129x, 1130x, 1131x, 1132x, 1133x, 1134x, 1135x, 1136x, 1137x, 1138x, 1139x, 1140x, 1141x, 1142x, 1143x, 1144x, 1145x, 1146x, 1147x, 1148x, 1149x, 1150x, 1151x, 1152x, 1153x, 1154x, 1155x, 1156x, 1157x, 1158x, 1159x, 1160x, 1161x, 1162x, 1163x, 1164x, 1165x, 1166x, 1167x, 1168x, 1169x, 1170x, 1171x, 1172x, 1173x, 1174x, 1175x, 1176x, 1177x, 1178x, 1179x, 1180x, 1181x, 1182x, 1183x, 1184x, 1185x, 1186x, 1187x, 1188x, 1189x, 1190x, 1191x, 1192x, 1193x, 1194x, 1195x, 1196x, 1197x, 1198x, 1199x, 1200x, 1201x, 1202x, 1203x, 1204x, 1205x, 1206x, 1207x, 1208x, 1209x, 1210x, 1211x, 1212x, 1213x, 1214x, 1215x, 1216x, 1217x, 1218x, 1219x, 1220x, 1221x, 1222x, 1223x, 1224x, 1225x, 1226x, 1227x, 1228x, 1229x, 1230x, 1231x, 1232x, 1233x, 1234x, 1235x, 1236x, 1237x, 1238x, 1239x, 1240x, 1241x, 1242x, 1243x, 1244x, 1245x, 1246x, 1247x, 1248x, 1249x, 1250x, 1251x, 1252x, 1253x, 1254x, 1255x, 1256x, 1257x, 1258x, 1259x, 1260x, 1261x, 1262x, 1263x, 1264x, 1265x, 1266x, 1267x, 1268x, 1269x, 1270x, 1271x, 1272x, 1273x, 1274x, 1275x, 1276x, 1277x, 1278x, 1279x, 1280x, 1281x, 1282x, 1283x, 1284x, 1285x, 1286x, 1287x, 1288x, 1289x, 1290x, 1291x, 1292x, 1293x, 1294x, 1295x, 1296x, 1297x, 1298x, 1299x, 1300x, 1301x, 1302x, 1303x, 1304x, 1305x, 1306x, 1307x, 1308x, 1309x, 1310x, 1311x, 1312x, 1313x, 1314x, 1315x, 1316x, 1317x, 1318x, 1319x, 1320x, 1321x, 1322x, 1323x, 1324x, 1325x, 1326x, 1327x, 1328x, 1329x, 1330x, 1331x, 1332x, 1333x, 1334x, 1335x, 1336x, 1337x, 1338x, 1339x, 1340x, 1341x, 1342x, 1343x, 1344x, 1345x, 1346x, 1347x, 1348x, 1349x, 1350x, 1351x, 1352x, 1353x, 1354x, 1355x, 1356x, 1357x, 1358x, 1359x, 1360x, 1361x, 1362x, 1363x, 1364x, 1365x, 1366x, 1367x, 1368x, 1369x, 1370x, 1371x, 1372x, 1373x, 1374x, 1375x, 1376x, 1377x, 1378x, 1379x, 1380x, 1381x, 1382x, 1383x, 1384x, 1385x, 1386x, 1387x, 1388x, 1389x, 1390x, 1391x, 1392x, 1393x, 1394x, 1395x, 1396x, 1397x, 1398x, 1399x, 1400x, 1401x, 1402x, 1403x, 1404x, 1405x, 1406x, 1407x, 1408x, 1409x, 1410x, 1411x, 1412x, 1413x, 1414x, 1415x, 1416x, 1417x, 1418x, 1419x, 1420x, 1421x, 1422x, 1423x, 1424x, 1425x, 1426x, 1427x, 1428x, 1429x, 1430x, 1431x, 1432x, 1433x, 1434x, 1435x, 1436x, 1437x, 1438x, 1439x, 1440x, 1441x, 1442x, 1443x, 1444x, 1445x, 1446x, 1447x, 1448x, 1449x, 1450x, 1451x, 1452x, 1453x, 1454x, 1455x, 1456x, 1457x, 1458x, 1459x, 1460x, 1461x, 1462x, 1463x, 1464x, 1465x, 1466x, 1467x, 1468x, 1469x, 1470x, 1471x, 1472x, 1473x, 1474x, 1475x, 1476x, 1477x, 1478x, 1479x, 1480x, 1481x, 1482x, 1483x, 1484x, 1485x, 1486x, 1487x, 1488x, 1489x, 1490x, 1491x, 1492x, 1493x, 1494x, 1495x, 1496x, 1497x, 1498x, 1499x, 1500x, 1501x, 1502x, 1503x, 1504x, 1505x, 1506x, 1507x, 1508x, 1509x, 1510x, 1511x, 1512x, 1513x, 1514x, 1515x, 1516x, 1517x, 1518x, 1519x, 1520x, 1521x, 1522x, 1523x, 1524x, 1525x, 1526x, 1527x, 1528x, 1529x, 1530x, 1531x, 1532x, 1533x, 1534x, 1535x, 1536x, 1537x, 1538x, 1539x, 1540x, 1541x, 1542x, 1543x, 1544x, 1545x, 1546x, 1547x, 1548x, 1549x, 1550x, 1551x, 1552x, 1553x, 1554x, 1555x, 1556x, 1557x, 1558x, 1559x, 1560x, 1561x, 1562x, 1563x, 1564x, 1565x, 1566x, 1567x, 1568x, 1569x, 1570x, 1571x, 1572x, 1573x, 1574x, 1575x, 1576x, 1577x, 1578x, 1579x, 1580x, 1581x, 1582x, 1583x, 1584x, 1585x, 1586x, 1587x, 1588x, 1589x, 1590x, 1591x, 1592x, 1593x, 1594x, 1595x, 1596x, 1597x, 1598x, 1599x, 1600x, 1601x, 1602x, 1603x, 1604x, 1605x, 1606x, 1607x, 1608x, 1609x, 1610x, 1611x, 1612x, 1613x, 1614x, 1615x, 1616x, 1617x, 1618x, 1619x, 1620x, 1621x, 1622x, 1623x, 1624x, 1625x, 1626x, 1627x, 1628x, 1629x, 1630x, 1631x, 1632x, 1633x, 1634x, 1635x, 1636x, 1637x, 1638x, 1639x, 1640x, 1641x, 1642x, 1643x, 1644x, 1645x, 1646x, 1647x, 1648x, 1649x, 1650x, 1651x, 1652x, 1653x, 1654x, 1655x, 1656x, 1657x, 1658x, 1659x, 1660x, 1661x, 1662x, 1663x, 1664x, 1665x, 1666x, 1667x, 1668x, 1669x, 1670x, 1671x, 1672x, 1673x, 1674x, 1675x, 1676x, 1677x, 1678x, 1679x, 1680x, 1681x, 1682x, 1683x, 1684x, 1685x, 1686x, 1687x, 1688x, 1689x, 1690x, 1691x, 1692x, 1693x, 1694x, 1695x, 1696x, 1697x, 1698x, 1699x, 1700x, 1701x, 1702x, 1703x, 1704x, 1705x, 1706x, 1707x, 1708x, 1709x, 1710x, 1711x, 1712x, 1713x, 1714x, 1715x, 1716x, 1717x, 1718x, 1719x, 1720x, 1721x, 1722x, 1723x, 1724x, 1725x, 1726x, 1727x, 1728x, 1729x, 1730x, 1731x, 1732x, 1733x, 1734x, 1735x, 1736x, 1737x, 1738x, 1739x, 1740x, 1741x, 1742x, 1743x, 1744x, 1745x, 1746x, 1747x, 1748x, 1749x, 1750x, 1751x, 1752x, 1753x, 1754x, 1755x, 1756x, 1757x, 1758x, 1759x, 1760x, 1761x, 1762x, 1763x, 1764x, 1765x, 1766x, 1767x, 1768x, 1769x, 1770x, 1771x, 1772x, 1773x, 1774x, 1775x, 1776x, 1777x, 1778x, 1779x, 1780x, 1781x, 1782x, 1783x, 1784x, 1785x, 1786x, 1787x, 1788x, 1789x, 1790x, 1791x, 1792x, 1793x, 1794x, 1795x, 1796x, 1797x, 1798x, 1799x, 1800x

I linguaggi di codifica

Linguaggio di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso
































































































I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

		Delete		FL	Flush Left		<i>bf.ital</i>	Set in Bold Face Italic
		Insert		FR	Flush Right		<i>lf</i>	Set in Light Face
		Join			Center Horizontally		<i>wf</i>	Wrong Font
		Move closer			Center Vertically			Hyphen
		Space			Move to the next line			En Dash
		Add Space			Move to the preceding line			Em Dash
		Delete Space			Indent 1 em			Superscript
		Transpose Word			Indent 2 ems			Subscript
		Transpose Letters			Paragraph			Comma
		To separate two or more marks			All Caps			Apostrophe
		Let it Stand (ignore correction)			Small Caps			Period
		Move Left			Caps & Small Caps			Semicolon
		Move Right			Capital Letter			Colon
		Move Up			Lower Case			Quotation Marks
		Move Down			Set in Roman			Parentheses
		Align Vertically			Set in Italic			Brackets
		Align Horizontally			Set in Bold Face			

I linguaggi di codifica

Linguaggi di marcatura

Il markup

La codifica con linguaggi di marcatura (markup) è in definitiva un insieme di convenzioni, rese attraverso specifiche sequenze di caratteri, etichette, codici, (detti tags) intercalati nel testo per permettere agli elaboratori elettronici di distinguere le varie parti di un documento.

Il markup formale

Un linguaggio di markup è un sistema formale per scambiare e pubblicare informazioni in formato testo in modo strutturato.

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Il markup formale

Markup formale: costituito da un sistema ben preciso di istruzioni, ognuna delle quali è dotata di una specifica semantica e sintassi.

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Diversi tipi di markup

Esistono diversi linguaggi di markup, per rappresentare diversi tipi di documenti.

- Linguaggi procedurali (specific markup languages)
- Linguaggi dichiarativi (generic markup languages)

I linguaggi di codifica

Linguaggi di marcatura procedurale

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Linguaggi procedurali

- Orientati al documento, indicano come deve essere elaborato e disposto il testo
- Istruzioni da inserire nel testo per specificarne specifiche caratteristiche
- Font, dimensione, spaziatura del carattere, posizionamento nella pagina, colore, etc.

Esempi: TeX e LaTeX, RTF

I linguaggi di codifica

Linguaggi di marcatura procedurale

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Esempio RTF

```
{\rtf1\ansi\deff0\adeflang1025
{\fonttbl{\f0\froman\fprq2\fcharset0 Times New Roman;}
{\f1\froman\fprq2\fcharset0 Times New Roman;}
{\f2\fnil\fprq2\fcharset0 Lucida Sans Unicode;}
{\colortbl;\red0\green0\blue0;\red128\green128\blue128;}
{\stylesheet{\s1\cf0{\*\hyphen2\hyphlead2\hyphtrail2\hyphmax0}
\rtlch\af5\afs24\lang255\ltrch\dbch\af2\afs24\langfe255
\loch\f0\fs24\lang1040\snext1 Standard;}}
```

I linguaggi di codifica

Linguaggi di marcatura procedurale

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Esempio LaTeX

```
documentclass a4paper , 10pt article
```

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Il markup procedurale

L'unico utilizzo di un testo codificato tramite un linguaggio procedurale è la creazione di un output orientato alla visualizzazione.

I linguaggi di codifica

Linguaggi di marcatura procedurale

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Scrivere la tesi di laurea con $\text{\LaTeX} 2_{\epsilon}$

Dipartimento di Ingegneria Meccanica,
Nucleare e della Produzione
Università di Pisa
56126 Pisa PI

Sommario

Lo scopo del presente articolo è fornire gli strumenti per scrivere una tesi di laurea utilizzando $\text{\LaTeX} 2_{\epsilon}$. Tale obiettivo è conseguito analizzando i problemi tipici incontrati durante la stesura della tesi e le possibili soluzioni; si pone particolare attenzione ai pacchetti da usare nelle varie circostanze. I singoli argomenti non vengono approfonditi nei dettagli ma si rimanda alla letteratura specifica o ad i manuali dei pacchetti suggeriti, ove necessario.

*Ringrazio in primo luogo Fabiano Busdraghi che ha collaborato alla scrittura delle sezioni riguardanti le figure e gli oggetti flottanti. Ringrazio inoltre tutti coloro che mi hanno consigliato durante la stesura e la revisione di questo documento ed in particolare Claudio Beccari, Gustavo Cevolani, Massimo Guiggiani, Maurizio Himmelmann, Lorenzo Pantieri e Emiliano Vavassori.

I linguaggi di codifica

Linguaggi di marcatura dichiarativi

Linguaggi dichiarativi

Orientati al testo, annotano la struttura, la funzione ed il significato degli elementi costitutivi del testo, tralasciandone l'aspetto.

- La posizione che il brano in questione occupa all'interno del documento (markup strutturale)
- Peculiarità del testo stesso (markup semantico)
- I fogli di stile definiscono la formattazione dell'output
- Molteplici usi del medesimo testo

Esempio: famiglia SGML, XML

I linguaggi di codifica

Linguaggi di marcatura dichiarativi

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition (DTD)
XSD
RELAXGN

Conclusioni

Markup dichiarativi: contenuto e presentazione

La separazione tra contenuto e presentazione non solo è intenzionale, ma è la caratteristica principale di questi sistemi di marcatura: essa permette di concentrarsi sull'annotazione logica-semantica per funzioni di ricerca e di analisi, lasciando ad altro (ai fogli di stile) la resa grafica.

Unico testo più usi

In questo modo si ha inoltre la possibilità di utilizzare uno stesso testo codificato con finalità o formattazioni differenti, a seconda delle varie esigenze.

I linguaggi di codifica

Markup dichiarativi: esempio SGML

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Standard Generalized Markup Language

```
1 <!DOCTYPE testo [  
2 <!ELEMENT testo (titolo?, paragrafo+)>  
3 <!ELEMENT titolo (#PCDATA)>  
4 <!ELEMENT paragrafo (#PCDATA)>  
5 ]>  
6 <testo>  
7   <titolo> Questo è il titolo del documento</titolo>  
8   <paragrafo> Questo è un paragrafo </paragrafo>  
9 </testo>  
10
```

I linguaggi di codifica

Markup dichiarativi vs Markup procedurali

resa a video della frase

Le *Guidelines for Electronic Text Encoding and Interchange* sono *molto* complete e descrivono uno standard di *markup* del testo basato su XML.

Le `\textit{Guidelines for Electronic Text Encoding and Interchange}` sono `\textit{molto}` complete e descrivono uno standard di `\textit{markup}` del testo basato su XML.

`<titolo>`Le Guidelines for Electronic Text Encoding and Interchange`</titolo>` sono `<enfasi>`molto`</enfasi>` complete e descrivono uno standard di `<linguastraniera>` markup`</linguastraniera>` del testo basato su XML.

LaTeX vs SGML

I linguaggi di codifica

Linguaggi di marcatura

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

linguaggi semi-dichiarativi e/o semi-procedurali

Esistono anche linguaggi che possono essere definiti semi-procedurali, o semi-dichiarativi, che come si intuisce utilizzano le istruzioni sia per una codifica di tipo procedurale, sia per una codifica di tipo descrittivo o dichiarativo.

HTML

HTML ha tra le sue etichette istruzioni di tipo procedurale per indicare come devono essere rese determinate porzioni di testo, e istruzioni di tipo dichiarativo che hanno una base semantica.

Progress status

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XXXXXXXXXXXXX

XML has its roots in Standard Generalized Markup Language (SGML), a language introduced in the 1980s that describes the structure and content of any machine-readable information.

XXXXXXXXXXXXXXXXXXx

XML can be thought of as a lightweight version of SGML. Like SGML, XML is a language used to create vocabularies for other markup languages

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XXXXXXXXXXXXXXXXXXXX

It is a set of rules for defining custom-built markup languages

XXXXXXXXXXXXXXXXXXXX

XML was originally created to structure, store, and transport information. Like SGML, XML can be used to create XML applications or vocabularies, which are markup languages tailored to contain specific pieces of information.

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XXXXXXXXXXXXXXXXXXXX

XML is a markup language that is extensible

XXXXXXXXXXXXXXXXXXXX

developed and maintained by the World Wide Web Consortium (W3C), an organization created in 1994 to develop common protocols and standards for sharing information on the World Wide Web.

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XXXXXXXXXXXXXXXXXX

XML, or eXtensible Markup Language, is a specification for storing information. It is also a specification for describing the structure of that information. And while XML is a markup language (just like HTML), XML has no tags of its own

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

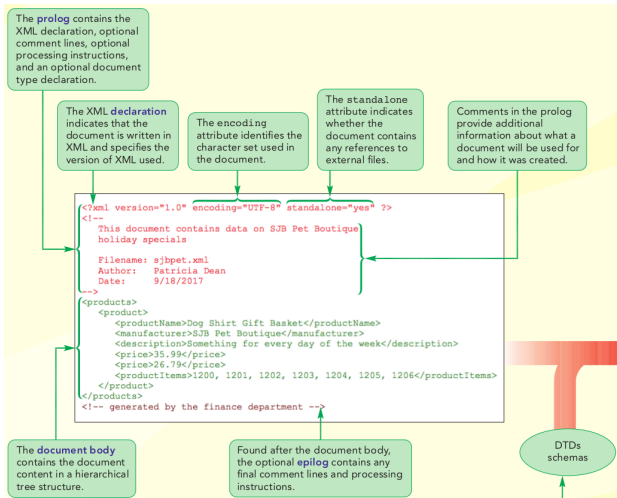


immagine dal libro *New Perspectives on XML, 3rd Edition*

Fondamenti XML

eXtensible Markup Language: Syntax Rules

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

- Every
- XML
- XML
- Every
- XML
- Some
- Comments

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XML vista ad albero

XML is based on hierarchical trees in which order is significant.
In XML, hierarchy and sequence are the main methods used to represent information.

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

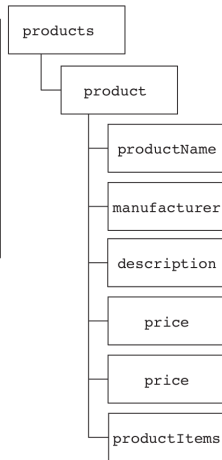
Conclusioni

XML document

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!--
  This document contains data on SJB Pet Boutique
  holiday specials

  Filename: sjbpet.xml
  Author:   Patricia Dean
  Date:    9/18/2017
-->
<products>
  <product>
    <productName>Dog Shirt Gift Basket</productName>
    <manufacturer>SJB Pet Boutique</manufacturer>
    <description>Something for every day of the week</description>
    <price>35.99</price>
    <price>26.79</price>
    <productItems>1200, 1201, 1202, 1203, 1204, 1205, 1206</productItems>
  </product>
</products>
<!-- generated by the finance department -->
```

Hierarchy tree structure



Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

TEI-XML vocabulary

To meet the need of textual scholars,

XML Vocabularies

XML vocabularies

Fondamenti XML

eXtensible Markup Language: Esempio TEI

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

```
<div type="narrative" n="6">
  <head>Sixth Narrative</head>
  <head>contributed by Sergeant Cuff</head>
  <div type="fragment" n="6.1">
    <opener>
      <dateline>
        <name type="place">Dorking, Surrey,</name>
        <date>July 30th, 1849</date>
      </dateline>
      <salute>To <name>Franklin Blake, Esq.</name> Sir, -</salute>
    </opener>
    <p>I beg to apologize for the delay that has occurred in the
      production of the Report, with which I engaged to furnish you.
      I have waited to make it a complete Report ...</p>
    <closer>
      <salute>I have the honour to remain, dear sir, your
        obedient servant </salute>
      <signed>
        <name>RICHARD CUFF</name> (late sergeant in the
          Detective Force, Scotland Yard, London). </signed>
      </closer>
    </div>
  </div>
```

immagine dal sito TEI Guide Lines

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Manutenibilità

Because XML focuses on communicating the data, the overall structure is simple and easy to design and maintain.

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Documento ben formato (well-formed)

A well-formed document contains no syntax errors and satisfies the general specifications for XML code as laid out by the W3C. At a minimum, an XML document must be well formed or it will not be readable by programs that process XML code.

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Parti principali di un documento XML

An XML document consists of three parts—the prolog, the document body, and the epilog.

Fondamenti XML

eXtensible Markup Language: Esempio TEI

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following document is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονόσ' ἔτη. πᾶσι γὰρ δὴ πού τοις φιλολόγοις γνώριμα
              τὰ περὶ τῆς φιλοπονίας τάνδρὸς ἱστορούμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσαν τήνδε "κατέβην χθές
              εἰς Πειραιᾶ μετὰ Γλαύκωνος τοῦ Ἀριστάντος
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>

<!-- This document is not completed and was cut without a special meaning -->
```

Parti principali di un documento XML

The prolog includes the following parts:

- XML declaration: indicates that the document is written in the XML language
- Processing instructions (optional): provide additional instructions to be run by programs that read the XML document
- Comment lines (optional): provide additional information about the document contents
- Document type declaration (DTD) (optional): provides information about the rules used in the XML document's vocabulary

Fondamenti XML

eXtensible Markup Language

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Parti principali di un documento XML

The document body,

Fondamenti XML

eXtensible Markup Language: Prologo

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

XML declaration

```
<?xml version="version number" encoding="encoding  
type" standalone="yes|no" ?>
```

Progress status

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

- Document Type Definition (DTD)
- XSD
- RELAXGN

4 Conclusioni

Elementi per la definizione degli schemi xml

principi

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

we need to make sure that the data that we receive follows a certain XML structure and should contain values which are coherent.

Your function needs to make sure that the caller passes correct XML data. You could make use of an XML Schema to perform this validation.

Performing such validations without the help of a SCHEMA will be extremely difficult most of the time.

Elementi per la definizione degli schemi xml

principi

Make sure that the XML document is structured exactly the way your function expects it to be.

We need an XML schema when we need to make sure that the XML document that we need to work with is in the expected format. Make sure that the values of elements and attributes are within the accepted range.

When data is managed and exchanged in XML format, there needs to be clear agreement about the structure of the XML document.

There needs to be a contract between the caller and the callee about the XML document being exchanged.

validate the XML document to make sure that it adheres to the format defined in the contract.

Elementi per la definizione degli schemi xml

principi

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

A Schema provides such the contract.
It defines the structure of the XML document.
It defines rules to validate the value of elements and attributes
as well as their formats.
Once a schema is defined, a Schema Validator can validate an
XML document against the rules defined in the Schema.

Elementi per la definizione degli schemi xml

Tipi di formalismi per definire schemi XML

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

DTD, XDR, SOX, Schematron, DSD, DCD, DDML, RELAX
NG

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Esempio DTD

An XML document may have a reference to an external DTD file or can have the DTD embedded as part of the XML file. The XML document given below has embedded DTD information.

```
<?xml version="1.0"?>
<!DOCTYPE Employee [
  <ELEMENT Name {#PCDATA}>
  <ELEMENT First {#PCDATA}>
  <ELEMENT Middle {#PCDATA}>
  <ELEMENT Last {#PCDATA}>
]>
<Employee>
  <Name>
    <First>Jacob</First>
    <Middle>Sebastian</Middle>
    <Last>
      <Name>
        <First>Jacob</First>
        <Middle>Sebastian</Middle>
        <Last>Sebastian</Last>
      </Name>
    </Employee>
  </Employee>
</Employee>
```

The example given below shows an XML document that refers to an external DTD file.

```
<?xml version="1.0"?>
<!DOCTYPE Employee SYSTEM "employee.dtd">
<Employee>
  <Name>
    <First>Jacob</First>
    <Middle>Sebastian</Middle>
    <Last>
      <Name>
        <First>Jacob</First>
        <Middle>Sebastian</Middle>
        <Last>Sebastian</Last>
      </Name>
    </Employee>
  </Employee>
</Employee>
```

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Document Type Definition (DTD)

A document type definition describes the rule for an xml document structure, declaring the elements, attributes and entities that are part of the xml document

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Document Type Definition (DTD)

For an xml document to be valid, it must be well formed and also satisfies the document type definitions specified by the document type declaration.

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Document Type Definition (DTD)

For an xml document to be valid, it must be well formed and also satisfies the document type definitions specified by the document type declaration.

a

ny document that lacks a document type declaration may be well formed but cannot be valid.

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Document Type Definition (DTD)

Document validation also aids file sharing since different applications that are aware of the generally agreed upon document type definition can produce xml document of similar structure and therefore easily communicate with each other through data exchange.

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Document Type Definition (DTD)

included in the document' prolog

Root element and content

declare the root element and its content model (children elements)

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Element Declaration (con figli)

```
<!ELEMENT element-name (child-element1,  
child-element-2 ...)>
```

Element Declaration (solo testo)

```
<!ELEMENT element-name (#PCDATA)>
```

Parsed character contents are plain texts with no child elements

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

Modificatori

Optional modifiers can also be used in element declarations to specify the number of times a child element may appear.

Modificatori

- + One or more occurrences
- ? Zero or one occurrence
- * Zero or more occurrences

Elementi per la definizione degli schemi xml

Principi Document Type Definition

Modificatori

```
<!ELEMENT element-name (B, C)+ >
```

```
<!ELEMENT element-name (B+, C) >
```

```
<!ELEMENT element-name (B, C+) >
```

```
<!ELEMENT element-name (B+, C+) >
```

Modificatori

If a child element must appear once then leave out the modifier.

Progress status

Codifica di
Testi -
Introduzione
XML Markup
a.a.
2018-2019

A.M. Del
Grosso

I linguaggi di
codifica

Fondamenti
del linguaggio
XML

Validare un
documento
XML e
Definire uno
Schema

Document Type
Definition
(DTD)
XSD
RELAXGN

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni